Responding to the Zika Epidemic: Preparation of a Neurodevelopmental Testing Protocol to Evaluate Young Children in Rural Guatemala

Amy K. Connery,^{1,2*} Gretchen Berrios-Siervo,^{1,3,4} Paola Arroyave,⁵ Desiree Bauer,⁵ Sara Hernandez,⁵ Alejandra Paniagua-Avila,⁵ Guillermo Antonio Bolaños,⁵ Saskia Bunge-Montes,^{5,6} Hana M. El Sahly,⁷ Mirella Calvimontes,⁵ Daniel Olson,^{1,3,6} Flor M. Munoz,^{7,8} and Edwin J. Asturias^{1,3,6}

¹Children's Hospital Colorado, Aurora, Colorado; ²Department of Physical Medicine and Rehabilitation, University of Colorado School of Medicine, Aurora, Colorado; ³Department of Pediatrics, University of Colorado School of Medicine, Aurora, Colorado; ⁴Department of Neurology, University of Colorado School of Medicine, Aurora, Colorado; ⁵Center for Human Development, Fundacion para la Salud Integral de los Guatemaltecos, Retalhuleu, Guatemala; ⁶Center for Global Health and Department of Epidemiology, Colorado School of Public Health, Aurora, Colorado; ⁷Department of Molecular Virology and Microbiology, Baylor College of Medicine, Houston, Texas; ⁸Department of Pediatrics, Baylor College of Medicine, Houston, Texas

Abstract. The ongoing Zika virus (ZIKV) epidemic in Latin America presented a unique opportunity to develop a neurodevelopmental assessment protocol for children in a lower middle-income country. Although studies of neurodevelopment in young children have taken place in many diverse global settings, we are not aware of any study that has provided a high level of detail about how a measure was selected and then specifically translated and adapted in a low-resource setting. Here, we describe considerations in measurement selection and then the process of translation and adaptation to assess neurodevelopmental outcomes of infants and young children with postnatal exposure to ZIKV in rural Guatemala. We provide a framework to other research teams seeking to develop similar assessment models across the globe.

INTRODUCTION

It is estimated that 43% of children less than 5 years living in low- and lower middle-income countries (LMICs) fail to meet basic developmental milestones in either cognitive or social-emotional development, which is linked to persistent neurodevelopmental problems and poorer performance on cognitive measures later in life.^{1–3} Stunting and poverty are commonly used as a proxy for risk as appropriate neurodevelopmental assessment measures are often not available in LMICs.¹ Most standardized and well-validated tools from high-income countries have significant limitations of score interpretation and feasibility when used in resource-constrained settings.

The urgency presented by the recent Zika virus (ZIKV) epidemic in Latin America and the associated neurodevelopmental sequelae from congenital infection highlighted the limitations in the neurodevelopmental assessment of young children in LMICs. Zika virus was introduced into Brazil as early as 2013,^{4,5} but it was not until microcephaly cases were reported in 2016 that the connection was made between congenital ZIKV infection and poor neurodevelopmental outcomes.⁶ Since then, a growing body of evidence supports the causal role of ZIKV in “congenital Zika syndrome,” a term that encompasses the numerous sequelae that have thus far been associated with this disease process, including fetal loss, growth restriction, microcephaly, hearing loss, ophthalmologic abnormalities (retina and optic nerve), neurologic abnormalities (seizures and irritability), motor/limb abnormalities (arthrogryposis), and others (dysphagia, constipation, and pneumonia).^{7–13} More concerning, perhaps, is the possibility that additional neurodevelopmental sequelae will only be revealed as affected children reach school age, when in utero infection does not result in the typical congenital Zika syndrome. In addition, the risks of postnatal infection in early life are not yet known.

To understand any potential neurodevelopmental sequelae of ZIKV infection globally, children should undergo systematic evaluation. Although there is a large literature on the translation and adaptation of neurodevelopmental tests for use in diverse settings,^{14–23} there is limited information on how to apply this knowledge to select, translate, and adapt neurodevelopmental tests for specific research projects while addressing various forms of potential bias.²⁴ Here, we describe the process of selection, translation, and adaptation of a neurodevelopmental test used to assess outcomes of young children living in a rural area of Guatemala endemic to dengue and ZIKV, to inform other researchers developing similar projects and help close the gap in this area.

METHODS

Setting. The study site is located in the coastal lowlands of southwest Guatemala, at the intersections of the departments of San Marcos, Quetzaltenango, and Retalhuleu, encompassing 22 rural communities with approximately 30,000 residents. These communities are monolingual Spanish-speaking and do not identify as indigenous. The population suffers from high rates of food insecurity and child undernutrition, diarrheal disease, maternal depression, maternal and child morbidity, and mortality.²⁵ Recent studies have shown that the area is highly endemic for dengue, chikungunya, and ZIKV.^{25,26}

Research design. In 2017, we launched a prospective, natural history study of postnatally acquired ZIKV infection in infants and young children at the University of Colorado Center for Human Development research and clinic site in Guatemala. The study was funded by the National Institutes of Health through the Baylor College of Medicine Vaccine and Treatments Evaluation Unit. The study protocol was reviewed and approved by the Institutional Review Board at Baylor College of Medicine, the Colorado Multiple Institutional Review Board, and the Ethics Review Committee of the Ministry of Health in Guatemala.

* Address correspondence to Amy K. Connery, Children's Hospital Colorado, Aurora, CO 80045-7106. E-mail: amy.connery@childrenscolorado.org

The study aimed to enroll children from birth to 5 years of age ($n = 1,200$) to follow them up for at least 1 year to determine the incidence of postnatally acquired symptomatic and asymptomatic ZIKV infection. Neurodevelopmental assessments would be conducted serially at the following time points for each participating child: baseline (enrollment), 6 months, and 1 year for infants, and baseline and 1 year for older children (age 1–5 years). We planned to characterize neurodevelopmental outcomes and describe potential neurodevelopmental differences in those infected with ZIKV versus noninfected children by age of the child and by timing of infection.

RESULTS

Developing the neurodevelopmental assessment plan.

We reviewed several resources to guide the process of selection, translation, and adaptation of neurodevelopmental measures and how to use them in the context of the research protocol.^{14–22,27–29} Figure 1 provides information on key decisions to be made in test selection when a culturally and linguistically appropriate measure is not available.

In the selection of specific tests, we recognized that very few tests have been standardized and validated for use in LMICs and none for the Guatemalan population and our study age range.^{23,30} As the study aimed to evaluate children while

the ZIKV epidemic was ongoing, the development of a new measure, which can take several years, was impractical because of time constraints.^{17,19} Therefore, a multidisciplinary team composed of psychologists and physicians with content, cultural, and local expertise selected the most appropriate test from existing tools based on several factors.^{23,30} We considered tests that assessed all domains commonly included in a comprehensive, performance-based assessment of young children: gross and fine motor, expressive and receptive language, and general cognitive/problem-solving skills.^{23,31–34} Importantly, we took into account the relevance of these developmental domains in the local community (i.e., was assessment of these particular domains the best way to measure the neurodevelopment of children in this community).^{21,35} Although we recognized that cultural factors can impact the rate of developmental skill attainment, consensus was clear that these skill areas were relevant for the study of early neurodevelopment in Guatemala.^{36,37}

The feasibility of test administration was a key aspect of our test selection.²³ We planned to enroll 1,200 children ages 0–5 years and perform the test multiple times during the length of the study (we expected approximately 2,900 evaluations within 24 months). Therefore, we sought an instrument that covered the entire age range of our population (to be able to compare results across subjects and time) and that was not

FIGURE 1. What to do when no test is available: Decision tree for neurodevelopmental test selection.

overly intensive to administer. The Mullen Scales of Early Learning (MSEL) was selected as our primary performance-based developmental assessment measure.³³ First, the MSEL is validated for children from birth to 68 months, allowing us to use a single test for the entire age span of children targeted by the study. Second, its administration time is approximately two to three times shorter than that of other commonly used performance-based measures,^{31,38} and finally, it covers all the developmental domains of interest to our study. Although the MSEL has not been widely used globally, it has been successfully adapted and administered in other settings, including South Africa and Benin.^{39,40}

Last, we considered the important step of establishing psychometric equivalence to the original version of the test and the lack of available regional norms for our population on the MSEL.^{41–44} Children at the study site are exposed to many neurodevelopmental risk factors that go beyond ZIKV infection.²⁶ Therefore, establishing equivalency to the original test would have entailed obtaining a “healthy,” lower risk sample of children, likely from another part of the country or region, which was not feasible given time constraints. In addition, as the objective of the study was to observe any potential adverse impact of ZIKV over and above the presence of the many other neurodevelopmental risk factors highly prevalent in the community, we determined that norming would occur within the obtained sample of children and we would not be establishing equivalency to the original, American norms or making cross-cultural comparisons with children living in other settings at this time for the purposes of this study. This process is commonly used when tests need to be adapted to non-U.S. populations by conducting group comparisons and factoring variables that may also contribute to poor child development, such as gender, home environment, and maternal education.^{40,45}

Process of translation and adaptation of the MSEL. The first step was the selection of a team composed of content and topic specialists, local experts, and persons native to the target and source languages, to ensure cultural relevance and sensitivity.^{14,18,27} In doing so, we avoided the vulnerability of Western bias and lack of expertise in the constructs of the study potentially introduced when the work is solely carried out by American researchers or local professional translators, respectively. Two bilingual neuropsychologists from the United States (A. K. C. and G. B. S.) and three Guatemalan psychologists (P. A., D. B., and S. H.) worked in tandem conducting focus groups of community members and local nurses until consensus was achieved.

We closely followed the theoretical principles outlined by Peña⁴⁶ to ensure equivalence between the original and adapted measures in cross-cultural research. Peña noted that although many research teams consider linguistic equivalence, or direct translation of a measure for use in another setting, functional, cultural, and metric equivalence are often not considered and can present threats to the validity of using the measure for clinical or research purposes. Although several studies have used these guiding principles in test adaptation,^{39,47–49} Table 1 provides details on how this process was undertaken by our team in the context of each of the four concepts of equivalence in the translation and adaptation of the MSEL.

In the last step of our process with the MSEL, we focused our efforts on addressing potential problems with method

bias, which may arise from administration procedures.^{21,50} The first was to ensure high-quality training of local examiners and fidelity in assessments. The three Guatemalan psychologists were trained at the site by one of the neuropsychologists from the United States (A. K. C.). Then, videos of pilot assessments and some ongoing assessments once the study began were sent to the neuropsychologists in the United States for review, and feedback and further training occurred in weekly video conferencing and by e-mail. The translation and adaptation were continually reviewed for any potential problems and modified by consensus during this time. A detailed manual with test administration and scoring instructions was created in Spanish, as was a translated test protocol to help ensure fidelity in methods. A quality checklist of procedures was also created, which included checks on scoring and test administration. This checklist was used extensively at the beginning of the study and then periodically so that the psychologists were able to audit each other.

We also addressed method bias by paying close attention to introducing the spaces and tasks to children and adapting the order of events to help with lack of familiarity of the testing environment, and stimulus familiarity. We developed several options of spaces in which to work to support children who were shy or uncomfortable, flexibility on the order in which subtests could be administered, and we allotted time to allow children to play with extra toys not used during the test they may have been seeing for the first time before the formal assessment began. We frequently asked an older sibling, for whom the situation was often less novel because of school attendance, to participate in test activities that we were not planning with the target child to model engagement with the evaluation process.

We supported caregivers unfamiliar with the testing situation, as well. At first, we only provided general explanations and guidance to the caregivers about testing procedures. We soon recognized that we needed to provide specific details not only about the assessment process, but also about what a caregiver should say or not say to help support and motivate her child while maintaining standardized testing procedures. For example, many caregivers required direct guidance on how to encourage their child without giving him or her a direct answer, additional assistance, or a threat for poor performance.

When testing was completed, we provided a visual to caregivers explaining how their child performed in relationship to the average child in the community. We then provided information around the importance of engaging in developmentally stimulating activities with the child. In response to caregivers' concerns that they did not have time during the day to engage in these activities with their child, we developed visual examples of how this could occur in the context of a typical day. These visuals are included in our Supplemental Materials.

DISCUSSION

To address key public health and research gaps during the ZIKV epidemic, we successfully and rapidly (in less than 3 months) developed a neurodevelopmental assessment plan for a rural area of Guatemala. Here, we review how decisions were made regarding test selection, how we addressed the translation and adaptation of specific test items to minimize potential bias and preserve the constructs being assessed,

TABLE 1
Translating and adapting: addressing four types of equivalence

Type of equivalence	Examples
<i>Linguistic:</i> Translation of test instructions, stimuli, and protocol forms from the original into the target language	<p>Translation of the MSEL into Spanish (appropriate for Guatemala) was undertaken by the neurodevelopmental team with permission from Pearson, the company that owns the MSEL copyright.</p> <p>Local nurses were consulted when questions about specific regional language differences arose.</p> <p>The final version was back-translated by a third-party professional translator not involved in the initial translation.</p> <p>The back-translation was reviewed and approved by Pearson.</p>
<i>Functional:</i> A translated test item should generate the same type of response as the original test item and, therefore, measure the same construct. That is, if a strict translation into the target language alters the meaning of the question or the concept being assessed and, therefore, results in a different pattern of responding, then changes to the language and wording may be necessary	<p>For an item attempting to elicit the response “car,” the English question is, “What do we ride in?” The direct translation of the verb “to ride” in Spanish refers to mounting or riding a horse or other animal. Therefore, we changed the verb to the Spanish “to drive” and used the most common locally used equivalent Spanish verb for the local population “manejar,” so children would give the same response as English-speaking children.</p> <p>There were instances in which an English word translated to two different words with somewhat different meanings in Spanish. In these cases, we chose the word in Spanish that would result in a similar answer to the English question. For example, in English, the word “fire” can refer to flames and to something that is burning (e.g., a fire in the woods). In Spanish, “fuego” (fire) refers only to flames and “incendio” (fire) refers to something that is burning. There is a question that asks, “What should you do if you wake up in the middle of the night and there is a fire inside of your house?” If we used the word “fuego,” children would not think of a house on fire, but could interpret that flames come from the wood burning stove. Therefore, the word “incendio” was selected to preserve the meaning of the question.</p> <p>The word “letter” has two meanings in English, allowing for two different responses (i.e., a symbol of the alphabet and a written communication). In Spanish, there are two different words for each of those definitions, “letra” and “carta,” respectively. It was decided by consensus that “letra,” as in letter of the alphabet, better represented the item level of difficulty at this section of the assessment.</p>
<i>Cultural:</i> An item or test question should be understood in a comparable manner, given local customs and experiences, between both the original and adapted versions	<p>One item involved placing a spoon and a doll on the table and asking the child to hand the doll to the evaluator. Because toys are so much more unfamiliar in this population of children, in our pilot testing, children did not want to hand over the more novel, interesting toy (the doll). Therefore, we changed the item to include two equally appealing and novel objects: a doll and a duck.</p> <p>On an item assessing motor movement and reactivity in babies, the evaluator or the caregiver tries to stimulate the baby with noise, movement, and facial expressions in an attempt to elicit vigorous movement. In our pilot testing, no babies were reacting to these attempts apart from eye contact. We hypothesized that, culturally, caregivers in this area are often not actively engaging with their babies in animated ways; therefore, babies were unfamiliar with this type of stimulation. We, therefore, proceeded to adapt this item to observe for any vigorous movement observed throughout the entire assessment and not only during this item.</p> <p>An item assessing language understanding asks the child why it would not be safe to swim when there was not a lifeguard present. Because these types of water safety controls are not present in Guatemala, most children would not know what a lifeguard is or their purpose. In addition, the direct translation for “lifeguard” can also include “lifejacket” in Spanish. Therefore, we adapted the question to ask why it would not be safe to swim without an adult present.</p>
<i>Metric:</i> Translated test items are equivalent in difficulty to the original test	<p>On items requesting the child to repeat various sentences, we had to ensure that sentence length (as measured by syllables) was not changed in translation of the item, thereby making the item either more or less difficult. So, while we addressed content so that familiarity would not interfere with task difficulty (i.e., cultural equivalence), we also adjusted translations to achieve exact equivalence in length, ensuring metric equivalency. For example, “In the winter we go sliding down the hill on our sled” was changed to “En el verano bajo la montaña en bici,” which directly translates to “In the summer I go down the hill on a bike.”</p> <p>For a more difficult task of visual reception, the subject is asked to match a target word to one of three choices. Although we did not expect most children in the study to be able to read, matching unfamiliar letter patterns would be harder than matching familiar letter patterns. Therefore, we changed the stimulus words to Spanish target words that matched the original English words in length, vowel placement, and rhyming when present. For example, we replaced “coat, boat, goat, coat” with the Spanish “masa, tasa, casa, masa”.</p> <p>We sometimes chose a slang word over a direct translation to maintain metric equivalency. For example, the direct translation of “faucet” is “grifo,” which is not a commonly used word in Guatemala, making it less likely a Guatemalan child would be able to define it. Therefore, we used the slang term (“chorro”) rather than a direct translation for the question.</p>

and how administration methods were adapted to the local population.

Although most health-care providers, researchers, and policy makers recognize the importance and urgency of assessment of early childhood development in LMICs, the scarcity of appropriate tests that can be administered by local trained staff is a barrier. In a review of available tests for the assessment of developmental problems in children in LMICs, only 20 tools were identified.⁵¹ Of these, child development experts did not rate any test as meeting all feasibility criteria of application in LMICs and only three tests met most criteria. There was “expert bias” toward the psychometric performance of an instrument, reducing possible false positives in individual subjects, ability of the results to be understood by health workers and caregivers, and linkage of results to early interventions.⁵¹

The norms on which neurodevelopmental test results are based are also frequently inadequate and present an additional challenge to drawing conclusions regarding neurodevelopmental test performance and comparison of information across studies. Norms should represent the demographic characteristics of the child being evaluated. However, overly narrow norms (e.g., drawn from a small local population) may obscure the effects of true etiologic risk factors that have negatively impacted a large portion of children in a community (e.g., poverty and malnutrition), and therefore, may not be sufficient for the evaluation of typical and atypical neurodevelopment. Consequently, national or regional norms are typically preferable.²³ Several organizations are working toward the goal of universal or regional norms or the development of measures that have global application.^{52–55} Akin to the World Health Organization growth reference charts created in 2006, a universal neurodevelopmental reference should ideally be a statistical summary of developmental milestones from a reference group of children, representative of some geographical region, and presented as the frequency distribution at different ages.⁵⁶

When a linguistically and culturally appropriate measure is not available, researchers have several resources available to guide the translation and adaptation of tests. However, details are rarely available about how teams can specifically apply these guidelines to meet the needs and goals of their study.²⁴ The lack of transparency in this process has been and continues to be an obstacle to comparability between studies. For example, we are not able to compare our test MSEL results with results from other studies that do not report on their translation and adaptation process as another study may have performed direct translation only and not adaptation or may have completed an adaptation that is substantially different from ours. Similarly, the lack of information about the translation and adaptation process complicates the ability to put the results of a specific study into context by limiting a reader's understanding of potential threats from bias in the assessment process.

Another obstacle to transparency in practice is that companies owning testing instrument licenses require large fees for permission to use and translate their tests. This can be a deterrent to research and other global health teams and can prevent the sharing of information by teams who did not obtain this sometimes cost-prohibitive permission. When a project is complete, testing companies do not make the translation and adaptation available to other

projects, and therefore, teams may be duplicating the work or working without an instrument that is actually available in a translated and adapted form. Testing companies should consider the global health need, and the financial benefits of making translations and adaptations of developmental assessments more widely available to communities and research teams that could benefit from them, perhaps using a tiered pricing approach.

Future work should focus on increased transparency in the decision-making of how a test is selected and the translation and adaptation process as a step toward improving population access to an increased pool of developmental assessment measures and norms. This can help bring us closer to the ultimate goal of developing and expanding access to local child development preventive and therapeutic intervention programs.

Received August 29, 2018. Accepted for publication October 31, 2018.

Published online December 26, 2018.

Note: Supplemental materials appear at www.ajtmh.org.

Acknowledgments: We thank Walla Dempsey, Mary Smith, Kay Tomashek, and Wendy Keitel for their review of this manuscript and guidance in this project as DMID and VTEU project officers and investigators; Edgar Eduardo Barrios for the visual artwork; the families in Trifinio, Guatemala, who participated in the pilot testing of this project and the nurses of the University of Colorado clinic in Trifinio who consulted with us on local language and cultural issues.

Financial support: This work was supported by the National Institutes of Health (NIH), funding mechanism: VTEU Contract HHSN2722013000151, Task Order No. HHSN27200013-16-0057.C1D1.0058.

Conflicts of Interest: E. A. reports grants from GlaxoSmithKline and Pfizer outside the submitted work.

Authors' addresses: Amy K. Connery, Children's Hospital Colorado, Aurora, CO, and Department of Physical Medicine and Rehabilitation, University of Colorado School of Medicine, Aurora, CO, E-mail: amy.connery@childrenscolorado.org. Gretchen Berrios-Siervo, Children's Hospital Colorado, Aurora, CO, Department of Pediatrics, University of Colorado School of Medicine, Aurora, CO, and Department of Neurology, University of Colorado School of Medicine, Aurora, CO, E-mail: gretchen.berrios-siervo@childrenscolorado.org. Paola Arroyave, Desiree Bauer, Sara Hernandez, Alejandra Paniagua-Avila, Guillermo Antonio Bolaños, and Mirella Calvimontes, Center for Human Development, Fundación para la Salud Integral de los Guatemaltecos, Retalhuleu, Guatemala, E-mails: parroyave@ufm.edu, desibauerh@gmail.com, smhh93@gmail.com, alejandra.paniagua.fsigcu@gmail.com, guillermo.bolando.fsigcu@gmail.com, and mirellacalvimontes@yahoo.com. Saskia Bunge-Montes, Center for Human Development, Fundación para la Salud Integral de los Guatemaltecos, Retalhuleu, Guatemala, Center for Global Health and Department of Epidemiology, Colorado School of Public Health, Aurora, CO, E-mail: saskia.bunge@gmail.com. Hana M. El Sahly, Department of Molecular Virology and Microbiology, Baylor College of Medicine, Houston, TX, E-mail: hana.elsahly@bcm.edu. Daniel Olson and Edwin J. Asturias, Children's Hospital Colorado, Aurora, CO, Department of Pediatrics, University of Colorado School of Medicine, Aurora, CO, Center for Global Health and Department of Epidemiology, Colorado School of Public Health, Aurora, CO, E-mails: daniel.olson@childrenscolorado.org and edwin.asturias@childrenscolorado.org. Flor M. Munoz, Department of Molecular Virology and Microbiology and Department of Pediatrics, Baylor College of Medicine, Houston, TX, E-mail: florm@bcm.edu.

REFERENCES

1. Black MM et al.; Lancet Early Childhood Development Series Steering Committee, 2017. Early childhood development coming of age: science through the life course. *Lancet* 389: 77–90.

2. John CC, Black MM, Nelson CA, 2017. Neurodevelopment: the impact of nutrition and inflammation during early to middle childhood in low-resource settings. *Pediatrics* 139 (Suppl 1): S59–S71.
3. Sudfeld CR, McCoy DC, Danaei G, Fink G, Ezzati M, Andrews KG, Fawzi WW, 2015. Linear growth and child development in low- and middle-income countries: a meta-analysis. *Pediatrics* 135: e1266–e1275.
4. Ayllón T et al., 2017. Early evidence for Zika virus circulation among *Aedes aegypti* mosquitoes, Rio de Janeiro, Brazil. *Emerg Infect Dis* 23: 1411–1412.
5. Faria NN et al., 2016. Zika virus in the Americas: early epidemiological and genetic findings. *Science* 352: 345–349.
6. Kleber de Oliveira W, Cortez-Escalante J, De Oliveira WT, do Carmo GM, Henriques CM, Coelho GE, Araújo de França GV, 2016. Increase in reported prevalence of microcephaly in infants born to women living in areas with confirmed Zika virus transmission during the first trimester of pregnancy—Brazil, 2015. *MMWR Morb Mortal Wkly Rep* 65: 242–247.
7. Baud D, Gubler DJ, Schaub B, Lanteri MC, Musso D, 2017. An update on Zika virus infection. *Lancet* 390: 2099–2109.
8. Cugola FR et al., 2016. The Brazilian Zika virus strain causes birth defects in experimental models. *Nature* 534: 267–271.
9. França GVA et al., 2016. Congenital Zika virus syndrome in Brazil: a case series of the first 1501 livebirths with complete investigation. *Lancet* 388: 891–897.
10. Gulland A, 2016. Zika virus may be linked to several birth defects, expert warns. *BMJ* 352: i1322.
11. Kapogiannis BG, Chakhtoura N, Hazra R, Spong CY, 2017. Bridging knowledge gaps to understand how Zika virus exposure and infection affect child development. *JAMA Pediatr* 171: 478–485.
12. Rasmussen SA, Jamieson DJ, Honein MA, Petersen LR, 2016. Zika virus and birth defects—reviewing the evidence for causality. *N Engl J Med* 374: 1981–1987.
13. Yopez JB, Murati FA, Pettito M, Peñaranda CF, de Yopez J, Maestre G, Arevalo JF; Johns Hopkins Zika Center, 2017. Ophthalmic manifestations of congenital Zika syndrome in Colombia and Venezuela. *JAMA Ophthalmol* 135: 440–445.
14. Beaton DE, Bombardier C, Guillemin F, Ferraz MB, 2000. Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine (Phila Pa 1976)* 25: 3186–3191.
15. Erkut S, 2010. Developing multiple language versions of instruments for intercultural research. *Child Dev Perspect* 4: 19–24.
16. Gregoire J, 2018. ITC guidelines for translating and adapting tests. *Int J Testing* 18: 101–134.
17. Geisinger KF, 1994. Cross-cultural normative assessment: translation and adaptation issues influencing the normative interpretation of assessment instruments. *Psychol Assess* 6: 304–312.
18. Hambleton RK, 2004. Issues, designs, and technical guidelines for adapting tests into multiple languages and cultures. RK Hambleton, PF Merenda, CD Spielberger, eds. *Adapting Educational and Psychological Tests for Cross-Cultural Assessment*. New York, NY: Psychology Press, 3–38.
19. Hambleton RK, Patsula L, 1999. Increasing the validity of adapted tests: myths to be avoided and guidelines for improving test adaptation practices. *J Appl Test Technol* 53: 1689–1699.
20. AERA, APA, NCME, 2014. *Standards for Educational and Psychological Testing*. Washington, DC: American Educational Research Association.
21. van de Vijver F, Tanzer NK, 2004. Bias and equivalence in cross-cultural assessment: an overview. *Rev Eur Psychol Appl* 54: 119–135.
22. World Health Organization, 2009. *Process of Translation and Adaptation of Instruments*. Available at: http://www.who.int/substance_abuse/research_tools/translation/en/. doi:entity/substance_abuse/research_tools/translation/en/index.html. Accessed August 2018.
23. Fernald LCH, Prado E, Kariger P, Raikes A, 2017. *A Toolkit for Measuring Early Childhood Development in Low and Middle-income Countries*. Available at: <https://ideas.repec.org/b/wbk/wbpubs/29000.html>. Accessed June 2017.
24. Acquadro C, Conway K, Hareendran A, Aaronson N, 2008. Literature review of methods to translate health-related quality of life questionnaires for use in multinational clinical trials. *Value Health* 11: 509–521.
25. Olson D, Lamb MM, Lopez MR, Paniagua-Avila MA, Zacarias A, Samayoa-Reyes G, Cordon-Rosales C, Asturias EJ, 2017. A rapid epidemiological tool to measure the burden of norovirus infection and disease in resource-limited settings. *Open Forum Infect Dis* 4: ofx049.
26. Asturias EJ et al., 2015. The center for human development in Guatemala: an innovative model for global population health. *Adv Pediatr* 63: 357–387.
27. Cassepp-Borges V, Balbinotti MAA, Teodoro MLM, 2010. Tradução e validação de conteúdo: uma proposta para a adaptação de instrumentos. L Pasquali, ed. *Instrumentação Psicológica: Fundamentos e Práticas*. Brazil: Artmed Editora, 506–520.
28. Gudmundsson E, 2009. Guidelines for translating and adapting psychological instruments. *Nord Psychol* 61: 29–45.
29. Krach SK, McCreery MP, Guerard J, 2017. Cultural-linguistic test adaptations: guidelines for selection, alteration, use, and review. *Sch Psychol Int* 38: 3–21.
30. Semrud-Clikeman M, Romero RAA, Prado EL, Shapiro EG, Bangirana P, John CC, 2016. Selecting measures for the neurodevelopmental assessment of children in low- and middle-income countries 23: 761–802.
31. Bayley N, 2006. *Bayley Scales of Infant and Toddler Development*, 3rd edition. Chicago, IL: Pearson.
32. World Health Organization, 2012. *Developmental Difficulties in Early Childhood: Prevention, Early Identification, Assessment and Intervention in Low-and Middle-income Countries: A Review*. Available at: http://www.who.int/about/licensing/copyright_form/en/index.html.
33. Mullen EM, 1995. *Mullen Scales of Early Learning*. Bloomington, MN: Pearson.
34. Snow CE, Van Hemel SB; Committee on Developmental Outcomes and Assessments for Young Children, 2008. *Early Childhood Assessment Why What and How*. Washington, DC: National Academies Press.
35. Van De Vijver FJR, Poortinga YH, 2004. Conceptual and methodological issues in adapting tests. RK Hambleton, PF Merenda, CD Spielberger, eds. *Adapting Educational and Psychological Tests for Cross-Cultural Assessment*. New York, NY: Psychology Press, 39–63.
36. Sabanathan S, Wills B, Gladstone M, 2015. *Child development assessment tools in low income and middle income countries: how can we use them more appropriately?* *Archives of Disease in Childhood* 100: 482–488.
37. Mangin KS, Horwood LJ, Woodward LJ, 2017. Cognitive development trajectories of very preterm and typically developing children. *Child Dev* 88: 282–298.
38. Newborg J, 2004. *Batelle Developmental Inventory*, 2nd edition. Ithaca, IL: Riverside.
39. Bornman J, Sevcik RA, Ronski MA, Pae HK, 2010. Successfully translating language and culture when adapting assessment measures. *J Policy Pract Intellect Disabil* 7: 111–118.
40. Koura KG et al., 2013. Usefulness of child development assessments for low-resource settings in francophone Africa. *J Dev Behav Pediatr* 34: 486–493.
41. Hambleton RK, Swaminathan H, 1985. *Item Response Theory. Principles and Applications*. The Netherlands: Springer.
42. Milfont TL, Fischer R, 2010. Testing measurement invariance across groups: applications in cross-cultural research. *Int J Psychol Res* 3: 111–121.
43. Reise SP, Widaman KF, Pugh RH, 1993. Confirmatory factor analysis and item response theory: two approaches for exploring measurement invariance. *Psychol Bull* 114: 552–566.
44. Sireci SG, 2004. Using bilinguals to evaluate the comparability of different language versions of a test. RK Hambleton, PF Merenda, CD Spielberger, eds. *Adapting Educational and Psychological Tests for Cross-Cultural Assessment*. New York, NY: Psychology Press, 117–138.
45. McCarthy AM, Wehby GL, Barron S, Aylward GP, Castilla EE, Javois LC, Goco N, Murray JC, 2012. Application of neurodevelopmental screening to a sample of South American infants: the bayley infant neurodevelopmental screener (BINS). *Infant Behav Dev* 35: 280–294.
46. Pena ED, 2007. Lost in translation: methodological considerations in cross-cultural research. *Child Dev* 78: 1255–1264.

47. Chen C et al., 2017. Adapting a developmental screening measure: exploring the effects of language and culture on a parent-completed social-emotional screening test. *Infants Young Child* 30: 111–123.
48. Macrine SL, Heji H, Sabri A, Dalton S, 2015. Cross-cultural adaptation of a developmental assessment for Arabic-speaking children with visual impairment. *Int J Sch Educ Psychol* 3: 256–266.
49. Panter-Brick C, Hadfield K, Dajani R, Eggerman M, Ager A, Ungar M, 2017. Resilience in context: a brief and culturally grounded measure for Syrian refugee and Jordanian host-community adolescents. *Child Dev* 89: 1803–1820.
50. Byrne BM, van de Vijver FJR, 2010. Testing for measurement and structural equivalence in large-scale cross-cultural studies: addressing the issue of nonequivalence. *Int J Test* 10: 107–132.
51. Fischer VJ, Morris J, Martinez J, 2014. Developmental screening tools: feasibility of use at primary healthcare level in low- and middle-income settings. *J Heal Popul Nutr* 32: 314–326.
52. De Onis M, 2006. WHO motor development study: windows of achievement for six gross motor development milestones. *Acta Paediatr Suppl* 450: 86–95.
53. McCoy DC, Peet ED, Ezzati M, Danaei G, Black MM, Sudfeld CR, Fawzi W, Fink G, 2016. Early childhood developmental status in low- and middle-income countries: national, regional, and global prevalence estimates using predictive modeling. *PLoS Med* 13: e1002034.
54. Janus M et al., 2007. *The Early Development Instrument: A Population-Based Measure for Communities*. Copublished by Offord Centre for Child Studies. ISBN 978-1-894088-84-8.
55. United Nation, 2001. *Road Map towards the Implementation of the United Nations Millennium Declaration*. Available at: www.un.org/documents/ga/docs/56/a56326.pdf.
56. Cole TJ, 2012. The development of growth references and growth charts. *Ann Hum Biol* 39: 382–394.