



10-2017

Measuring home environments across cultures: Invariance of the HOME scale across eight international sites from the MAL-ED Study

Paul C. Jones

Temple University, Philadelphia, PA, USA.

Laura L. Pendergast

Temple University, Philadelphia, PA, USA.

Barbara A. Schaefer

The Pennsylvania State University, State College, PA, USA.

Muneera Rasheed

Aga Khan University, muneera.rasheed@aku.edu

Erling Svensen

University of Bergen, Norway

See next page for additional authors

Follow this and additional works at: https://ecommons.aku.edu/pakistan_fhs_mc_women_childhealth_paediatr



Part of the [Maternal and Child Health Commons](#), and the [Pediatrics Commons](#)

Recommended Citation

Jones, P. C., Pendergast, L. L., Schaefer, B. A., Rasheed, M., Svensen, E., Scharf, R., Shrestha, R., Maphula, A., Roshan, R., Rasmussen, Z. (2017). Measuring home environments across cultures: Invariance of the HOME scale across eight international sites from the MAL-ED Study. *Journal of School Psychology, 64*, 109-127.

Available at: https://ecommons.aku.edu/pakistan_fhs_mc_women_childhealth_paediatr/336

Authors

Paul C. Jones, Laura L. Pendergast, Barbara A. Schaefer, Muneera Rasheed, Erling Svensen, Rebecca Scharf, Rita Shrestha, Angelina Maphula, Reeba Roshan, and Zeba Rasmussen



Contents lists available at ScienceDirect

Journal of School Psychology

journal homepage: www.elsevier.com/locate/jschpsyc



Measuring home environments across cultures: Invariance of the HOME scale across eight international sites from the MAL-ED study[☆]



Paul C. Jones^{a,*}, Laura L. Pendergast^a, Barbara A. Schaefer^b, Muneera Rasheed^c, Erling Svensen^{d,e}, Rebecca Scharf^f, Rita Shrestha^g, Angelina Maphula^h, Reeba Roshanⁱ, Zeba Rasmussen^j, Jessica C. Seidman^j, Laura E. Murray-Kolb^b, The MAL-ED Network Investigators¹

^a Temple University, Philadelphia, PA, USA

^b The Pennsylvania State University, State College, PA, USA

^c Aga Khan University, Pakistan

^d University of Bergen, Norway

^e Haydom Lutheran Hospital, Tanzania

^f University of Virginia

^g Institute of Medicine, Tribhuvan University, Kathmandu, Nepal

^h University of Venda, South Africa

ⁱ Christian Medical College, Vellore, India

^j Fogarty International Center/National Institutes of Health, Bethesda, MD, USA

ARTICLE INFO

Action Editor: Sterett Mercer

Keywords:

Measurement invariance

Validity

Culture

MAL-ED

Environment

Low-income countries

ABSTRACT

The home environment provides the context for much of a child's early development. Examples of important aspects of the home environment include safety, cleanliness, and opportunities for cognitive stimulation. This study sought to examine the psychometric properties of an adapted form of the Home Observation for the Measurement of the Environment (HOME; Caldwell & Bradley, 1984, 2003) across the eight international sites of the MAL-ED project (Dhaka, Bangladesh; Vellore, India; Bhakatapur, Nepal; Naushahro Feroze, Pakistan; Fortaleza, Brazil; Loreto, Peru; Venda, South Africa; Haydom, Tanzania), to identify a factor structure that fit the data at all sites, and to derive a subset of items that could be used to examine home environmental characteristics across sites. A three-factor structure (i.e., Emotional and Verbal Responsivity; Clean and Safe Environment; Child Cleanliness) was identified, and partial measurement equivalence/invariance across sites was supported. Overall, these findings lend support for the use of portions of this abbreviated and adapted version of the HOME for use among heterogeneous, cross-cultural groups in low- and middle-income nations.

[☆] The Etiology, Risk Factors and Interactions of Enteric Infections and Malnutrition and the Consequences for Child Health and Development Project (MAL-ED) is carried out as a collaborative project supported by the Bill & Melinda Gates Foundation, the Foundation for the NIH, and the National Institutes of Health, Fogarty International Center. The authors thank the staff and participants of the MAL-ED Network for their important contributions.

* Corresponding author at: School Psychology Program, Temple University, 1301 Cecil B. Moore Ave., Philadelphia, PA 19122, USA.

E-mail address: paul.c.jones@me.com (P.C. Jones).

¹¹ See appendix

<http://dx.doi.org/10.1016/j.jsp.2017.06.001>

Received 5 April 2016; Received in revised form 18 April 2017; Accepted 14 June 2017

Available online 29 June 2017

0022-4405/ © 2017 The Authors. Published by Elsevier Ltd on behalf of Society for the Study of School Psychology. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

1. Introduction

In recent years, school psychologists have been serving increasingly diverse populations within the U. S. and internationally (Jimerson, Annan, Skokut, & Renshaw, 2009; Mendes, Abreu-Lima, Almeida, & Simeonsson, 2014). To better serve such diverse populations, in the United States and worldwide, a comprehensive understanding of the context in which children grow and develop is necessary. School psychologists need to be aware of the connection between physical health and psychological health, and how the context of early childhood may influence later development. Experiences of early childhood, such as early exposure to toxins (Jusko et al., 2008) and parenting practices (Bradley, Corwyn, Burchinal, McAdoo, & Garcia Coll, 2001), have a significant influence on cognitive development and academic outcomes (Andrade et al., 2005; Bradley & Corwyn, 2002; Leventhal & Brooks-Gunn, 2000; Pendergast & Kaplan, 2015; Shaw, 2004; Shonkoff et al., 2012). School psychology researchers, particularly those working in low-income settings, may require a psychometrically robust measurement tool to evaluate environmental characteristics. Thus, psychometrically supported tools are necessary to ensure that early childhood development is being measured accurately across contexts.

Like most individual differences, child development varies significantly based on factors such as family values, the physical environment, or even the political structure of the community or nation. Macrosystem-level factors, such as sociocultural norms, may differentially influence home environment between cultural groups. From an ecological perspective (e.g., Bronfenbrenner, 1977), evaluation of a child's home environment may be useful as it relates to early childhood health and development. For researchers to understand the complex ways in which biological and psychosocial factors influence child development from an ecological perspective, it is also necessary to understand the environment in which the child develops – which requires the measurement of characteristics of the home environment. Moreover, a number of psychometric issues, including linguistic and cultural non-congruence, can arise when using assessments across different cultural groups (Bracken & Barona, 1991; Van De Vijver & Poortinga, 1982). Thus, systematic and psychometrically robust measurement tools are required to adequately evaluate factors relevant to cognitive, psychological, and physical development, with respect to cultural differences. The current research is an effort to provide psychometric support for the use of a home environment observation tool across multiple cultural groups.

The Home Observation for Measurement of the Environment Scale (HOME; Caldwell & Bradley, 1984, 2003) is an observational tool intended to assess characteristics of the home environment that are relevant to childhood health and development, including both physical and relational dimensions. The HOME has been used in medical and epidemiological studies worldwide (e.g., Bradley, Mundfrom, Whiteside, Casey, & Barrett (1994); Bradley, Whiteside et al. (1994); Bradley, Corwyn, & Whiteside-Mansell, 1996; Walkowiak et al., 2001; Williams et al., 2003; Black et al., 2004), but studies of the psychometric properties of the HOME across cultures are limited. Measures do not always operate similarly across cultures, and it is necessary to examine whether the constructs and items function in a similar fashion across sites (e.g., measurement equivalence/invariance; Pendergast, von der Embse, Kilgus, & Eklund, 2017). To that end, the version of the HOME adapted by Black et al. (2004) was administered and the psychometric properties examined as part a larger multinational study examining several factors related to child development across each of eight sites: Dhaka, Bangladesh; Vellore, India; Bhakatapur, Nepal; Naushahro Feroze, Pakistan; Fortaleza, Brazil; Loreto, Peru; Venda, South Africa; Haydom, Tanzania.

1.1. Environment and child development

Early life experiences play a critical role in child development and can provide the foundation for future health and well-being (National Research Council and Institute of Medicine, 2000; Shonkoff, Boyce, & McEwen, 2009; Shonkoff et al., 2012). Bronfenbrenner (1977) emphasized the importance of understanding human development in the context of the environment in which an individual exists. Accordingly, school psychologists and others working in health care and education have been encouraged to adopt an ecological perspective and consider multiple characteristics of a child's environment in research and in practice (Pendergast & Kaplan, 2015; Sheridan & Gutkin, 2000). Further, it is important to consider the role that both physical and interpersonal/social-emotional factors can play in child development.

For instance, Andrade et al. (2005) reported on the differential impact of quality household stimulation on cognitive development. Further, the effects of high levels of early stress and adversity in childhood on later development have also been well documented (e.g., Shonkoff et al., 2012). Adverse neurobiological effects can have devastating effects on children who are exposed to toxins in the home environment (e.g., Grandjean & Landrigan, 2006). Even low levels of lead exposure can have devastating effects on intellectual ability (e.g., Lanphear et al., 2005). These early childhood experiences can have significant implications for a child's later ability to learn and engage in a school environment.

There are many physical, interpersonal, and experiential factors that characterize a high-quality home environment. Factors that are commonly discussed in the psychological and educational literatures include availability of learning resources (e.g., toys and books), sensitive and responsive interactions from caregivers, and the provision of a variety of experiences to a developing child (e.g., Andrade et al., 2005). However, physical characteristics of the home environment, such as safety (e.g., Delgado et al., 2002), cleanliness (e.g., Aiello & Larson, 2002), home crowding (e.g., Solari & Mare, 2012), likelihood of exposure to neurotoxins (e.g., Grandjean & Landrigan, 2006), and likelihood of exposure to illness-causing pathogens (e.g., Cardoso, Cousens, de Góes Siqueira, Alves, & D'Angelo, 2004) also play a key role in childhood development. Quality home environments can serve as protective factors against adverse outcomes (e.g., Bradley, Whiteside et al., 1994).

1.1.1. Relational environmental characteristics (parental responsiveness)

Andrade et al. (2005) examined the effect of maternal level of education on the quality of cognitive stimulation in the home

environment, finding positive associations between maternal education and home stimulation that were associated with enhanced cognitive development. Such evidence has spurred the interest in and development of ways to assess specific variables within the context of the home (e.g., mother-child interactions; Rasheed & Yousafzai, 2015). Given the significance of the home environment on early childhood development, the home environment is sometimes used as a proxy for early childhood care and education (Iltus, 2006). Increased parent responsiveness has been linked with cognitive development and later learning (Bradley et al., 2001). Further, Evans et al. (2010) explained how maternal responsiveness mediates the effect of home crowding on cognitive development.

1.1.2. Physical environmental characteristics (cleanliness, safety, healthfulness)

Physical factors in the environment have also been shown to affect child development (Ferguson, Cassells, MacAllister, & Evans, 2013). Shaw (2004) reviewed the link between physical health and the environment, stating that children in poorer quality housing are more likely to experience increased rates of infection. Research has pointed to higher rates of physical illness such as skin infections among homes without working toilet facilities (Baillie et al., 2005). Such infectious diseases can have both direct and indirect effects on child physical and psychological development (Walker et al., 2007). Additionally, poor environmental safety may lead to increased risk of exposure to neurotoxins, such as lead, which can lead to cognitive impairments (Kippler et al., 2012; Lanphear et al., 2005; Needleman et al., 1979). Home crowding has been identified by several researchers as having one of the strongest associations with scores on the HOME (Bradley & Caldwell, 1984; Johnson, Breckenridge, & McGowan, 1984), with research demonstrating that home crowding may lead to social withdrawal and reduced parental responsivity (e.g., Bradley & Caldwell, 1984; Evans, 2006). Given that both physical and interpersonal/experiential factors influence developmental outcomes, it is important for a measure of home environment to account for both parental responsivity to children and the physical characteristics of the environment.

1.2. Assessment of home environment

1.2.1. Methods for assessing environmental characteristics

Several tools may be used for gathering information about and assessing the home environment, including observations, interviews, and rating scales. Observational methods (such as the HOME; Caldwell & Bradley, 1984, 2003) rely on trained observers to assess and evaluate the home environment, the behavior of the child, and the interactions between a child and an adult. Interview methods use information collected through parents'/caregivers' responses to a series of questions asked by an interviewer about the environmental characteristics of the home and the behavior of the child. Leventhal, Selner-O'Hagen, Brooks-Gunn, Bingenheimer, and Earls (2004) developed an interview tool that was based on the HOME to assess home environments in Chicago neighborhoods, also requiring a trained professional in the administration of the interview. Rating scales are completed by informants who know the child well, such as parents or caregivers, and do not require much time or expertise for scoring and interpretation. For instance, following the development of the HOME, Frankenburg and Coons (1986) developed a similar measurement tool to evaluate the quality of one's home environment. However, rather than relying on the direct observation of the environment by a trained professional, the questionnaire was completed by the parents and/or caregivers. Inherent limitations are relevant with both interviews and self-report rating scales in the form of response bias. For example, social desirability may influence parent responses (Paulhus, 2002). Additional concerns included extreme response styles and acquiescent responding (Jackson & Messick, 1958; Krosnick, 1999). Thus, direct observation of an environment has the potential to improve objectivity through the use of trained raters. Regardless of the type of assessment tool used (rating scale, interview, or observation), the scores produced by the measure must be reliable and valid among members of the population with which they are used (AERA, APA, & NCME, 2014).

1.2.2. Home observation for measurement of the environment

The Home Observation for Measurement of the Environment (HOME; Caldwell & Bradley, 1984, 2003), is the most widely recognized and commonly used instrument for evaluating the quality of the home environment and has been frequently revised and adapted. Originally developed from the work of Elardo, Bradley, and Caldwell (1975), the HOME is an observational tool that allows trained observers to quantify observations related to mother-child interactions, family living patterns and habits, orderliness of the physical home environment, and the potential of the observed environment to positively influence development. It was developed to evaluate aspects of the environment that are believed to be important for development, such as parental responsivity and opportunities to learn (Elardo & Bradley, 1981).

The HOME has been used for a number of purposes, including evaluating interventions (Brotman et al., 2005) and assessing the effect of parenting on a number of outcome variables, including: academic achievement (Bradley & Caldwell, 1984), executive functioning (Blair, Raver, and Berry (2014)), language development (Elardo, Bradley, & Caldwell, 1977), and prenatal/perinatal exposure to neurotoxins (Walkowiak et al., 2001). Brooks-Gunn, Klebanov, and Liaw (1995) demonstrated the use of the HOME to measure risk-factors and outcomes associated with families living in poverty. The relationship between the HOME and cognitive development has also been well documented over time (Elardo et al. (1975); Bradley & Caldwell, 1976; Bradley, Caldwell, & Rock, 1988; Luster & Dubow, 1992). Parental attachment has also been evaluated utilizing the HOME (NICHD Early Child Care Research Network, 2001). Within the United States, attempts have been made to evaluate the measurement equivalence/invariance of the HOME (Bradley et al., 1996; Whiteside-Mansell, Bradley, Tresch Owen, Randolph, & Cauce, 2003). Bingenheimer, Raudenbush, Leventhal, and Brooks-Gunn (2005) evaluated the HOME for its equivalence across racial and ethnic groups. Consistent with the review completed by Bradley et al. (1996), while measurement equivalence was demonstrated between Black and White families on the HOME item functioning, the scale did not function similarly among Hispanic families. Thus, while the psychometric properties of

the HOME with sample populations consistent with the original groups studied using the HOME have demonstrated equivalence (e.g., Saudino & Plomin, 1997; Shaw & Vondra, 1995; Prodromidis, Lamb, Sternberg, Hwang, & Broberg, 1995), research with other racial and ethnic groups has been mixed (Totsika & Sylva, 2004).

1.2.3. Cross-cultural adaptations of the HOME

The HOME has been shown to predict a number of child developmental outcomes within the United States (Totsika & Sylva, 2004). In order to use the inventory outside of the US, however, adaptations have been necessary (Black et al., 2004; Bradley et al., 1996). A review of the cross-cultural uses of the HOME scale by Bradley et al. (1996) revealed important differences among developing nations, relative to history, sociocultural norms, and politics, highlighting the need for the equivalence of the HOME to be demonstrated across cultures before broad generalizations are made. Rather, the authors highlight the limitations of the HOME when administered without adaptations, as key home factors likely differ significantly between cultures. Generally, items associated with cognitively stimulating environments appeared to demonstrate the highest degree of equivalence across cultures. However, equivalence of the HOME factors must be weighed against the relevance in different cultures. Several researchers have highlighted the need for and proposed guidelines for the development of cross-cultural assessment tools (Hambleton, 2001; Peña, 2007; Van Widenfelt, Treffers, De Beurs, Siebelink, & Koudijs, 2005). Geisinger (1994) outlines several steps in adapting the use of normative assessments for cross-cultural use, particularly the importance of ensuring that the same psychological construct is being measured across groups.

More recent uses of the HOME have also included participants from Thailand (Williams et al., 2003) and Bangladesh (Black et al., 2004). As noted by Bradley et al. (1996), researchers have examined the factor structure of the HOME in several different countries, and the findings have differed widely ranging from identification of a three-factor structure in Costa Rica (Stevens & Bakeman, 1985) to a seven-factor one (with many factors having very poor reliability, e.g., < 0.50) in a study of Hispanic-Americans (Bradley, Mundfrom et al., 1994; Bradley, Whiteside et al., 1994). Notably, both the study locations and the factor analytic techniques differed across research studies, so it is unclear whether the discrepant findings are the result of cultural differences or methodological ones. In summary, due to inconsistency in methods and findings across cultural groups, the factor structure of the original HOME is unknown. Moreover, the factor structure of the Bangladeshi adaptation of the HOME that was used in this study, to our knowledge has never been examined. The current paper adds to the existing knowledge base by examining the measurement equivalence/invariance of the HOME across international sites.

1.2.4. Measurement equivalence/invariance

Measurement equivalence/invariance refers to the extent to which an assessment tool measures the intended construct in a similar way across groups (Drasgow, 1984; Pendergast et al., 2017). Moreover, Pendergast and colleagues identified a number of instances when ME/I is important to understanding how tests and measures are used across different populations, such as evaluating test bias or differences in item functioning across cultural groups. To our knowledge, no study has ever explicitly tested the ME/I of HOME scores across international groups. According to Vandenberg and Lance (2000), “The establishment of measurement equivalence/invariance across groups is a necessary and logical prerequisite to conducting substantive cross-group comparisons (e.g., tests of group mean differences, invariance of structural parameter estimates)” (p. 1). In other words, researchers need to provide evidence that scores on a measurement tool are equivalent/invariant across groups before mean scores across groups can be compared. Measurement equivalence/invariance is important when examining the psychometric properties of a particular measure across groups, such as gender (Marsh, 1985, 1987), age (Marsh & Hocevar, 1985) or cultural groups (Riordan & Vandenberg, 1994).

1.3. Present study

The purpose of the current study was to identify subsets of HOME items that yielded scores with evidence of validity and measurement equivalence/invariance across the eight international MAL-ED sites. Notably, measurement equivalence/invariance (ME/I) is considered to be a necessary prerequisite for comparing mean scores across groups (Vandenberg & Lance, 2000; Pendergast et al., 2017). Moreover, ME/I is one piece of evidence that allows for increased confidence in testing primary hypotheses related to home environment across groups.

2. Methods

2.1. Overview of MAL-ED study

The MAL-ED study (The MAL-ED Network Investigators, 2014) is a multi-national and multi-disciplinary, observational, prospective, cohort study conducted at eight international sites. The primary objective of MAL-ED is to examine the extent to which nutrition, enteric infections, and gut inflammation and permeability influence growth and cognitive and psychological development of children in resource-poor environments. Many factors influencing child development were measured, including economic resources, maternal education, reasoning and depression as well as aspects of the home environment.

2.2. Participants

The MAL-ED sample is comprised of mothers and infants from research sites located in eight different nations: Dhaka, Bangladesh

(BGD; $n = 237$); Vellore, India (INV; $n = 235$); Bhaktapur, Nepal (NEB; $n = 235$); Naushahro Feroze, Pakistan (PKN; $n = 246$); Fortaleza, Brazil (BRF; $n = 198$); Loreto, Peru (PEL; $n = 257$); Venda, South Africa (SAV; $n = 231$); Haydom, Tanzania (TZH; $n = 240$). Community-based sampling was used in this study whereby all families who met criteria (e.g., singleton infant, weight above 1500 g) within a selected geographic region were invited to participate. More detailed descriptions of the demographic characteristics of each site are provided in [Appendix A](#). A more thorough description of the sites, recruitment procedures, adaptations and translation are described elsewhere ([The MAL-ED Network Investigators, 2014](#)). Of the 2145 infants enrolled within 17 days of birth, 1598 infants (74%) were followed to at least 24 months of age, when the HOME was administered. The targeted window for administration of the HOME was 24 months of age (± 15 days).

2.3. Measure

2.3.1. Home Observation for the Measurement of the Environment (HOME)

The HOME has four versions: Infant/Toddler (0–2 years), Early Childhood (3–6 years), Middle Childhood (6–10), and Early Adolescence (10–15 years). The version of the HOME scale used in the MAL-ED cohort study (provided in [Appendix B](#)) was originally adapted by [Black et al. \(2004\)](#) for use among a population of infants in Bangladesh. These modifications, made in conjunction with anthropologists, ensured that the HOME items were culturally appropriate to Bangladeshi communities. Black et al.'s adapted version used in MAL-ED is comprised of 48 items on six subscales: parent responsiveness, avoidance of restriction and punishment, organization of the environment, appropriate play materials, parental involvement, and variety in daily stimulation. The purpose of the adaptation was to ensure that items were culturally appropriate to the community in which it was being utilized. Further, additional items were added that addressed the physical layout of the environment. However, the psychometric validity of this adapted version of the HOME has not yet been examined outside of Bangladesh. Establishing psychometric support for the use of this measure in further analyses associated with the MAL-ED study is necessary, and the current research sought to address that gap.

2.4. Procedure

2.4.1. Overview

For a detailed review of the methods and procedures for the MAL-ED study see [Murray-Kolb et al. \(2014\)](#) and [The MAL-ED Network Investigators \(2014\)](#). A review of procedures specific to the HOME Inventory is provided here. Prior to beginning our research, the HOME Inventory was translated into each target language and back translated by an individual who had never seen the original form. The translations were compared by individuals with relevant measurement and cultural expertise and adaptations were made as needed. Subsequently, the HOME Inventory was administered at 24 months of age to assess the characteristics of the home environment. To ensure uniformity of administration and observation, regional on-site trainings were conducted at the eight field sites. Each training session at each of the eight field sites lasted 5–6 days with 1 day devoted to training in the adapted HOME inventory. Training included both didactic components and practice administrations.

2.4.2. Personnel selection

The HOME was administered by psychological research assistants (PRAs) at each site. The criteria agreed upon for the selection of a PRA included a level of education that is comparable to a Master's degree in psychology or child development and fluency in speaking the local language. Six of the MAL-ED sites found PRAs with the required criteria. Two of the sites had difficulty identifying individuals with these qualifications and, therefore, the PRAs in these two sites were selected based on their education level (minimum education of 14 years in Pakistan and 11 years in Tanzania), knowledge of child development, and previous experience working on research studies. PRAs who administered the HOME underwent training which included supervised practice administrations.

2.4.3. HOME administration

The HOME assessment was scheduled at a time when the primary caregiver and child were at home and the child was awake. The PRAs visited and worked with the families regularly and, in most instances, were known to the participants. Prior to beginning the HOME, the PRAs engaged in conversation and rapport building with the families. Subsequently, they informed the families that they would need to observe the child and the home setting for approximately 1 h. The HOME form was completed over the course of the hour, and any unanswered questions were posed to the child's primary caregiver after the observation was complete. Each item had guidelines for scoring and all items on the forms received binary scores (+) or (–). PRAs scored the items based on their observations during the visit and, when appropriate asked the caregiver to provide information. All responses were recorded on paper-and-pencil forms.

2.4.4. Quality assurance

As the assessment of the HOME was carried out in eight, culturally diverse sites, many quality assurance procedures were implemented to ensure consistency in assessment and scoring methods across sites. These procedures included training, weekly conference calls with key team members, random reviews of materials and forms, and assessments of inter-rater reliability conducted by senior psychologists. Weekly calls ensured that a discussion forum was available so that all sites had the opportunity to discuss relevant issues and barriers, had common understanding of expectations, and that any change was communicated to each site. Ten percent of forms were checked by the senior psychologist at each site by random selection. Inter-rater reliability, as indicated in

Murray-Kolb et al. (2014) of the HOME was addressed through simultaneous scoring by the PRA and the senior psychologist. Supervisors and PRAs were required to have 85% correspondence in ratings before PRAs could administer the scale independently (and in order to continue independent administration). All variations between the PRA and senior psychologist were discussed between the two raters.

2.5. Data analysis

2.5.1. Overview

To accomplish our objective of deriving a measure of home environmental characteristics that is likely to be useful across diverse groups in many settings, a five-pronged approach was used to examine the validity of the scores from each scale across the eight groups. Prior to analyses, the sample was divided into two randomly selected subsamples – one for exploratory factor analysis (EFA; $n = 940$) and one for confirmatory factor analysis (CFA; $n = 939$). In the first step, EFA was conducted to better understand the nature of the factor structure in this sample. In the second step, the internal consistency of scores was examined with the EFA subsample (all eight sites). For the third step, CFAs were conducted based on the CFA subsample ($n = 939$) with data from all sites. Next, a MIMIC modeling approach was employed (see Jones, 2006; Woods, Oltmanns, & Turkheimer, 2009; and Pendergast et al., 2014 for reviews and illustrations of the MIMIC modeling approach to evaluating ME/I). Finally, convergent validity of HOME scores and crowding was examined. Details of the procedures used in each analysis are provided below.

2.5.2. Exploratory factor analysis and internal consistency

Common factor analysis (principal axis factoring, promax rotation) was selected because it is the preferred technique for identification of latent factor structures (Briggs & MacCallum, 2003; Fabrigar, Wegener, MacCallum, & Strahan, 1999). Exploratory factor analyses were conducted in SPSS (version 22) based on tetrachoric matrices (Lorenzo-Seva & Ferrando, 2012). Three methods were employed to determine the number of factors to retain: parallel analysis (PA; Horn, 1965; Watkins, 2006), minimum average partials (MAP; Velicer, 1976), and a visual scree test (Cattell, 1966; Henson & Roberts, 2006). Pattern coefficients ≥ 0.30 on only one factor were considered salient (Stevens, 2002). Factor structures with (a) \geq three items with salient pattern coefficients; (b) evidence of reliability (e.g., internal consistency ≥ 0.70 or support based on item response theory analyses); and (c) a conceptually or theoretically meaningful pattern were considered to be adequate and were submitted for confirmatory factor analysis.

2.5.3. Baseline confirmatory factor analysis

CFAs were conducted in MPlus 6.12 using mean and variance adjusted weighted least squares estimation (WLSMV). Goodness of fit was evaluated based on multiple criteria (Tanaka, 1993): (a) root mean square error of approximation (RMSEA) values < 0.075 , (b) standardized root mean square residual (SRMR) values < 0.08 , and (c) comparative fit index (CFI) values > 0.90 (Hu & Bentler, 1995; Markland, 2007). Although more stringent criteria exist (e.g., CFI > 0.95 ; Hu & Bentler, 1999), rigid adherence to these more stringent criteria could lead to inappropriate rejection of well-fitting models (e.g., Marsh, Hau, & Wen, 2004). Therefore, relatively liberal criteria (e.g., CFI > 0.90) were utilized, and multiple fit indices were examined (Markland, 2007).

2.5.4. MIMIC modeling analyses

When conducting cross-cultural research, it is important for researchers to evaluate whether their measures operate similarly across groups. Two important principles related to measurement in cross-cultural research are measurement equivalence/invariance (ME/I; see Pendergast et al., 2017 for a review) and differential item functioning (DIF; see Zumbo, 2007, for a detailed discussion of ME/I and DIF). The two most commonly used techniques for examining ME/I and DIF are multi-group confirmatory analysis and item response theory analysis. However, both of those approaches are typically used in analyses examining a small number of groups and may not be appropriate for analyses across many groups (e.g., Asparouhov & Muthén, 2014; Pendergast et al., 2017).

Multiple indicators multiple causes (MIMIC; Jöreskog & Goldberger, 1975) modeling is similar to CFA but allows for examination of DIF through the inclusion of covariates to test the influence of extraneous variables (e.g., nationality, gender, race, etc.) on the factor structure. This allows the researcher to evaluate the equivalence of scores from a measure across groups (Muthén, 1989; Jones, 2006). Advantages to MIMIC modeling, as noted by Woods et al. (2009), include ease in modeling multidimensional constructs (such as home environment), availability of high quality software, capacity for modeling categorical data (as is used in the HOME), and ability to examine DIF across multiple groups simultaneously. To examine differential item functioning (DIF) across the eight international sites, a MIMIC modeling approach was used in this study.

The MIMIC modeling analyses began by following the procedures recommended by Shih and Wang (2009). First, a subset of one or more DIF-free items were identified and specified as anchor items in subsequent analyses (Shih & Wang, 2009; Woods et al., 2009). Subsequently, the stepwise method for DIF analyses described by Jones (2006) was used, and, accordingly, multiple interim MIMIC models were estimated. Next, the remaining (non-anchored) items were tested for DIF using an iterative process whereby the items were tested one at a time, and modification indices were inspected after each analysis. If the modification indices indicated that freeing a path between an item and a cultural (site) variable would meaningfully influence the χ^2 value, then the item was considered to have DIF. Finally, effect sizes were calculated using MLR estimation to derive proportional odds ratio (OR) estimates and interpreted according to the Cole, Kawachi, Maller, and Berkman (2000) guidelines.

2.5.5. Convergent validity

Convergent validity analyses included a measure of home crowding. Home crowding was defined in the larger MAL-ED study as

the number of persons living in the home divided by the number of rooms in the home. Correlations between each factor obtained from the factor analyses of the HOME were obtained. Pearson correlation coefficients were then computed for each obtained factor and crowding. Moderate relationships, whereby increased crowding predicting poorer home environmental conditions were predicted based on prior literature (Bradley & Putnick, 2012; Conley, 2001; Matheny, Wachs, Ludwig, & Phillips, 1995).

3. Results

3.1. Preliminary analyses

The version of the HOME administered in this study included 48 items. Before beginning primary analyses, all five items from the original Avoidance of Restriction and Punishment subscale were removed from the dataset. This subscale included items that inquired about caregiver behaviors such as spanking or slapping the child, shouting at the child, and expressing overt annoyance toward the child. Such behaviors were quite rare in our sample (e.g., endorsed by 0–2% of respondents at many sites) – creating numerous empty cells and thereby violating the assumptions of categorical factor analysis.

3.2. Exploratory factor analyses

The remaining 43 items from the HOME were submitted for EFA. Bartlett's test of sphericity was statistically significant ($p < 0.001$). The Kaiser-Meyer-Olkin statistic was 0.81 lending support for the suitability of the correlation matrix for factor analysis (Kaiser, 1974). The recommended number of factors varied widely. Parallel analysis suggested 6 factors, visual inspection of the scree plot indicated 7 factors, and minimum average partials suggested 16 factors. All of these factor solutions were attempted but rejected because (a) one or more factors had an inadequate number of items (fewer than 3) with salient pattern coefficients (> 0.30), or (b) one or more factors with insufficient reliability (< 0.70). Subsequently, four- and five-factor solutions were attempted and rejected for the same reasons. The three-factor solution met the aforementioned criteria and was retained. Items that did not load on any factor were removed from analyses, and the analyses were conducted without them.

The final three-factor solution met all a priori criteria. It was comprised of factors that appeared to reflect the following dimensions: Emotional and Verbal Responsivity; Clean and Safe Environment, and Child Cleanliness. Pattern coefficients and alpha values for all factors are reported in Table 1. The Emotional and Verbal Responsivity factor was comprised of 13 items that appeared to reflect the extent to which the primary caregiver (usually the mother) responded to the child's emotional states and attempts at verbalization. Examples of items on this factor include, Caregiver spontaneously praises child's qualities or behavior and Caregiver spontaneously vocalizes to the child." Five items were included on the Clean and Safe Environment factor. This factor appeared to reflect the extent to which features of the environment were conducive to the safety and health of the child, reflecting more permanent features of the child's environment. Examples of items on this factor include "The house is relatively light" and "The stove is

Table 1

Pattern matrix for exploratory factor analysis (principal axis factoring) with promax (Oblique) Rotation of the Home Observation for the Measurement of the Environment (HOME).

Items (abbreviated)	Emotional and verbal responsivity	Child cleanliness	Environmental safety and healthfulness
Spontaneous praise for child 2 ×	.693	.035	–.059
Respond to vocalization by kid	.603	.032	–.023
Initiates verbal exchange with observer	.588	–.146	.063
Spontaneous vocalize to kid	.580	–.076	.016
Positive response	.568	–.041	–.070
Tells name of object	.566	–.135	.015
Affectionate with child	.509	.151	.129
Talks to child often	.495	.053	–.141
Smiles at child	.464	.055	.000
Encourages development	.430	.225	–.020
Believe can modify behavior	.429	.040	–.087
Expresses ideas freely	.386	–.043	.142
Positive feeling conveyed	.372	.016	.058
Child clean	–.028	.905	–.079
Clothes relatively clean	.008	.831	.078
Hair neat/washed recently	.063	.807	.004
Stove in safe area	.050	–.041	.811
Play area relatively safe free of hazards	.061	–.036	.782
House relatively light	–.089	–.012	.517
House neat orderly	–.103	.341	.366
Taken to health clinic regularly	–.031	.158	.317
Alpha coefficient	.815	.724	.811
		.704^a	.890^a

Note. Factor loadings $> .30$ are in boldface. Alpha estimates were based on EFA subsample.

^a Alpha value without "Neat and orderly" item (cross loaded item).

located in a relatively safe area.” The third factor, Child Cleanliness, included four items that reflected the cleanliness of the child’s hair, clothes, etc. Examples of items on this factor include “The child is relatively clean, with no offensive odor” and “The child’s hair is relatively clean.” Notably, one item, “House is neat and orderly,” cross-loaded on both the Clean and Safe Environment and the Child Cleanliness factor. In factor analysis, cross-loaded items are often problematic. Cross-loaded items indicate shared contributions to multiple factors, raising questions about the independence of extracted factors with regard to the common indicator. Therefore, this item was tested as loading on both factors, separately, in CFA, and model fit was compared.

3.3. Confirmatory factor analyses

The three-factor model of HOME scores was examined in CFA based on an independent subsample ($n = 939$). The model was determined to be a good fit to the data based on several criteria (e.g., CFI = 0.94; RMSEA = 0.055 (CI = 0.051–0.060)). All factors were allowed to correlate (inter-factor correlations ranged from -0.04 to 0.68). No post-hoc modifications were used. However, two models (specified a priori) were run and compared: one with the item “House is neat and orderly” loading on the Clean and Safe Environment factor and one with that item loading on the Child Cleanliness factor. The model where the item loaded on the Child Cleanliness factor had superior fit (alternative model CFI = 0.89; RMSEA = 0.08) and was retained.

3.4. MIMIC modeling analyses

3.4.1. Overview of DIF analyses

MIMIC analyses were conducted based on the entire sample. After identification of anchor items, DIF analyses proceeded in a stepwise fashion as described by Jones (2006) involving estimation of multiple interim MIMIC models (see Pendergast et al., 2014, for a more detailed illustration of this process). Subsequently, a final model was tested allowing DIF where indicated (based on findings from the interim models). A listing of items and sites for which DIF was allowed can be found in Table 2. The final model, which allowed for DIF, demonstrated good fit to the data (e.g., CFI = 0.96; RMSEA = 0.05). The final MIMIC model was re-estimated using the maximum likelihood estimator with robust standard errors (MLR) to examine the effect sizes associated with DIF by calculating proportional odds ratio (OR) estimates. As suggested by Cole et al. (2000), OR values > 2.0 or < 0.5 were considered to be indicative of meaningful DIF. Parameter estimates and OR estimates are reported in Table 2.

3.4.2. Emotional and verbal responsivity DIF

For factor I, Emotional and Verbal Responsivity (EVR), there were statistically significant paths between the EVR factor and five

Table 2
Parameter estimates and differential item functioning of HOME items across seven sites.

Items	BGD	BRF	INV	NPB	PEL	PKN	SAV	TZH
	PE (OR)	PE(OR)	PE(OR)	PE(OR)	PE(OR)	PE(OR)	PE(OR)	PE(OR)
Caregiver vocalizes to child	-.39(.86)	.17(4.51)	NT	-.13(.16)	-	-	-	NT
Caregiver responds to child	.20(1.27)	.11(3.70)	NT	NT	-	-.09(.30)	-	NT
Caregiver teaches names of objects ^a	NT	-	NT	-	-	-	-	NT
Caregiver initiates conversation	.24(.06)	.16(3.91)	NT	.34(13.42)	-	-	-	NT
Caregiver expresses ideas freely	-	-	NT	-	-	-.33(.09)	-	NT
Caregiver praises child	.44(9.92)	-	NT	-	.55 (3.43)	-	-	NT
Caregiver discusses child positively	.68(.70)	-	NT	-	-	-.15(.23)	-	NT
Caregiver is affectionate	.78(384.01)	.12(2.41)	NT	.28(6.62)	-.66 (.39)	-.43(0.05)	-	NT
Caregiver responds positively to child	0.11(.53)	.15(4.05)	NT	0.25(5.40)	.27 (1.90)	.15(2.15)	.19(5.74)	NT
Caregiver smiles at child	NT	.21(6.10)	NT	-.20(.11)	.44(2.92)	-	.18(4.72)	NT
Caregiver talks to child while working	.70(116.26)	.17(3.30)	NT	-.20(.18)	-	-	-	NT
Caregiver encourages development	.62(55.81)	.13(2.44)	NT	.21(3.66)	-	-.22(22)	.13(2.82)	NT
CG believes behavior is modifiable	.70(118.81)	-	NT	.22(3.93)	-	-.13(.44)	.20(4.46)	NT
Factor I total items with meaningful DIF	6	8	NT	8	3	7	4	NT
Child regularly visits health clinic	.49(67.61)	.19(4.23)	-	-.10(.58)	-	-	-.33(.13)	-.39(.36)
Play area is relatively safe ^a	-.16(.18)	-	-	-	-	-	-	-
Stove is in a safe area	-	-	-	-	-	-.18(.28)	-	-
House is relatively light	-	-	-	-	-	-	-	-.20(.23)
Factor II total items with meaningful DIF	2	1	0	0	0	1	1	2
House is neat and orderly	NT	-	-	-.21(.19)	-	-.21(.18)	.19(5.81)	-.89(.24)
Child is relatively clean	NT	-	-	-	-	-.12(.15)	-	-
Child’s hair is relatively clean ^a	NT	-	.14(4.75)	-	-	-	-	-
Child’s clothes are clean	NT	-	-	-.07(.19)	.41	-	-.07(24)	-.10(0.09)
Factor III total items with meaningful DIF	NT	0	1	2	1	2	2	2

Note. - indicates that there was no DIF and the parameters were not estimated. PE = Standardized Parameter Estimate; Odds ratios are reported in parentheses, and items deemed to have meaningful DIF (ORs > 2.0 or $< .50$) are bolded. NT refers to items that were not tested for DIF due to empty cells (an absence of participants who endorsed or failed to endorse the item.) Study sites were not comprised of samples that were demographically representative of the countries in which they were located and should not be interpreted as such.

^a Denotes items used as anchor items in analyses.

sites. Four sites scored significantly lower in regard to overall EVR scores: (BGD: $PE^2 = -0.39, p < 0.001$; BRF: $PE = -0.34, p < 0.001$; PEL: $PE = -0.17, p = 0.001$; and PKN: $PE = -0.10, p = 0.016$). One site scored significantly higher on the overall EVR score (NPB: $PE = 0.28, p < 0.001$), and there was no statistically significant difference for the South Africa site. Notably, invariance could not be tested for two sites, INV and TZH, due to extremely limited variance by which observers endorsed items for every participant or nearly every participant which resulted in numerous empty cells. At the item level, after controlling for differences on the overall factor, a high number of items on each factor displayed statistically significant and meaningful DIF as shown in Table 2.

3.4.3. Clean and safe environment DIF

On factor II, Clean and Safe Environment (CSE), statistically significant paths occurred between the CSE factor and six sites. Four sites scored significantly higher in regard to the overall CSE scores (INV: $PE = 0.52, p < 0.001$; NPB: $PE = 0.12, p = 0.001$; PEL: $PE = 0.33, p < 0.001$; SAV: $PE = 0.56, p < 0.001$). Two sites scored significantly lower: (PKN: $PE = -0.38, p < 0.001$; TZH: $PE = -0.54, p < 0.001$), and no statistically significant differences were evident for the BGD and BRF sites. At the item level, after controlling for differences on the overall factor, DIF was minimal. Three sites had zero items with meaningful DIF, three sites had only one item with meaningful DIF, and two sites had two DIF items (reported in Table 2). Thus, the CSE factor was largely comparable across sites.

3.4.4. Child cleanliness DIF

For factor III, Child Cleanliness (CC), statistically significant paths emerged between the CC factor and overall scores from seven sites. Four sites scored significantly higher in regard to the overall CC score (BRF: $PE = 0.15, p = 0.004$; NPB: $PE = 0.36, p < 0.001$; PEL: $PE = 0.10, p = 0.025$; and SAV: $PE = 0.32, p < 0.001$). Three sites scored significantly lower in regard to the overall CC score (INV: $PE = -0.26, p < 0.001$; PKN: $PE = -0.21, p < 0.001$; and TZH: $PE = -0.21, p < 0.001$). One site, BGD, could not be tested for DIF on the CC factor due to extremely limited variability in responses to items. At the item level, after controlling for the overall factor scores, evidence of a moderate degree of DIF was present. One site had zero items with meaningful DIF, two sites had one item, and four sites had two items (reported in Table 2). Thus, the CC factor was somewhat comparable across sites.

3.5. Convergent validity

Pearson's correlations were analyzed based on unit-weighted factor scores, first between each of the three factors obtained through EFA (see Table 3). Emotional and verbal responsivity was significantly correlated with environmental safety ($r = 0.270, p < 0.001$), although, not surprisingly, not with cleanliness ($r = 0.057, p = 0.051$). Environmental safety was positively correlated with cleanliness ($r = 0.387, p < 0.001$).

The association between each of three HOME factors with environmental crowding (number of individuals living in the home divided by number of rooms used for sleeping) was also examined. Environmental crowding was not associated with emotional and verbal responsivity ($r = -0.028, p = 0.338$). However, as expected based on prior research (Bradley & Caldwell, 1984; Johnson et al., 1984), crowding had low to moderate correlations with both environmental safety ($r = -0.385, p < 0.001$) and cleanliness ($r = -0.336, p < 0.001$), both associations reaching statistical significance.

4. Discussion

Bronfenbrenner (1977), in his ecological systems theory, suggested that child development must be interpreted in the context of the child's environment. The empirical literature clearly demonstrates that early childhood environment has a meaningful and lasting impact on child development (Andrade et al., 2005; Bradley & Corwyn, 2002; Leventhal & Brooks-Gunn, 2000; Pendergast & Kaplan, 2015; Shaw, 2004; Shonkoff et al., 2012). This research examines the validity of scores from an observational measure of the home environment, the primary context in which child development occurs. Examining the role that environmental factors play in child development across cultures is challenging for researchers because environmental contexts vary widely across cultural groups. Moreover, the extent to which many measures of early childhood environment are reliable and valid across multiple cultural groups is unknown. This study examined the validity of the HOME across eight international sites that differ widely in regard to culture, language, and economic resources.

4.1. Overview of findings

The findings from this study largely support the psychometric properties of scores from the HOME as measures of Emotional and Verbal Responsivity, Clean and Safe Environment, and Child Cleanliness in the eight low- and middle-income settings examined in this study. The validity scores of some of the other original subscales (e.g., avoidance of physical punishment; variety in stimulation) were not supported with this sample. Moreover, the extent to which each subscale was supported, as well as the degree of measurement equivalence across sites, varies by subscale and site.

² PE = parameter estimate

Table 3

Summary of Pearson correlation coefficients, means, and standard deviations for scores on the EVR, CSE, CC, and crowding.

Measure	1	2	3	4
1. EVR	–			
2. CSE	.27**	–		
3. CC	.06	.39**	–	
4. Crowding	–.03	–.39**	–.34**	–
M	10.71	3.20	3.26	2.65
SD	2.65	1.12	1.21	1.67

Note. EVR = Emotional and Verbal Responsivity; CSE = Clean and Safe Environment; CC = Child Cleanliness.

** $p < 0.01$ based on combined EFA and CFA subsamples.

4.1.1. Emotional and verbal responsivity

Specifically, these findings indicate that HOME scores measure Emotional and Verbal Responsivity fairly well (in regard to structural validity) for six of our eight research sites located in six countries across three continents BGD, BRF, NPB, PKN, PEL, and SAV, while not functioning well at the Indian and Tanzanian MAL-ED sites. This suggests that some aspects of maternal emotional and verbal responsivity can be assessed in diverse groups, while in some areas the scores are less likely to be useful in research or clinical contexts. Nonetheless, support for the assessment of emotional and verbal responsivity across cultures is a critical finding from this study and can be used to inform future research, as this construct is closely related to child development. The dyadic interactions between parent/caregiver and child are important for the development of cognitive skills and language from an early age (e.g., Sameroff & Fiese, 2000). Thus, further research in this area should focus on determination of the impact of emotional and verbal responsivity on cognitive development across cultures.

4.1.2. Clean and safe environment

Clean and safe environments for children are critical to ensuring positive developmental trajectories. Increased rates of infection have been linked to the physical environment (Shaw, 2004; Bailie et al., 2005). The current findings support the use of the HOME Clean and Safe Environment subscale across the sites examined in this study and suggest that it is likely to be an effective measure for future researchers examining diverse populations in low- and middle-income nations. A high degree of measurement equivalence was demonstrated across sites – as 50–100% of items displayed no meaningful DIF in each site. In addition, positive correlations were observed between home crowding and environmental safety.

Children living in crowded conditions are at a higher risk for injury, including burns (Delgado et al., 2002) and respiratory infection (Cardoso et al., 2004). Given the relationship of crowding with both the risk of injury and the spread of infectious disease, identification of proper safety measures taken within the home (e.g., the placement of the stove within the home) and steps to prevent the spread of infectious diseases (e.g., access to healthcare), findings of convergent validity of our home safety and crowding variables are important and support the use of the scale.

4.1.3. Child cleanliness

These findings suggest that the HOME Child Cleanliness factor effectively measured Child Cleanliness in seven of the eight MAL-ED sites located on three continents (i.e., all sites except Bangladesh). Moreover, a high degree of measurement equivalence was evidenced across sites for this subscale as 50–100% of items displayed no meaningful DIF in each of the seven sites. This factor was also correlated positively with home crowding. These findings support the use of the HOME Child Cleanliness subscale across the seven indicated sites and suggest that it is a promising tool for future researchers examining diverse populations in low- and middle-income nations.

Environmental factors that contribute to the health and safety of the child are also consistently illustrated in empirical literature (e.g., Ferguson et al., 2013; Grandjean & Landrigan, 2006; Lanphear et al., 2005; Needleman et al., 1979; Walkowiak et al., 2001). Personal hygiene, nutrition and the risk of infection are also linked to cognitive and psychological development (Aiello & Larson, 2002; Walker et al., 2007). Thus, a measure of a child's cleanliness can serve as an indicator of the hygiene practices of the family that may contribute to the spread of infectious diseases. The relationship between infectious disease and cognitive development, for instance, is the subject of ongoing research with some evidence suggesting that chronic infections may contribute to adverse effects on cognitive development and academic achievement (Guerrant, DeBoer, Moore, Scharf, & Lima, 2013; Guerrant, Oriá, Moore, Oriá, & Lima, 2008; Walker et al., 2007). As a physical characteristic of the home, the relationship between home crowding and child cleanliness is also relevant, given the possibility of increased transmission of infectious disease among members of a household (e.g., Jaie, Baker, & Venugopal, 2011). Thus, attention to child cleanliness is critical for reducing levels of infectious disease transmission in crowded homes.

4.2. Implications

The current research sought to evaluate the psychometric properties of the HOME across cultural groups. Culturally relevant modifications to the HOME have been previously made in evaluating home environment variables (e.g., Black et al., 2004). To our knowledge, however, a measurement equivalence/invariance evaluation of the HOME scale has never been conducted outside the US,

despite suggestion by the original author encouraging the use of the HOME across cultural groups (Bradley et al., 1996). Thus, although the influence of home environment on various cognitive skills and developmental trajectories appears to remain constant across cultures, the specific indicators of the home environment that are relevant to individual cultural groups is more variable. As the sample population from which the MAL-ED data is drawn was not selected to be representative of entire regional populations, the indicators and factors that emerged as relevant for the particular cultural groups specific to these analyses may not generalize to entire regional populations.

4.3. Strengths, limitations, and future directions

Strengths of the current study include the use of a large international sample from diverse communities across eight low- and middle-income countries. The HOME measure, although previously used with international groups (Black et al., 2004; Bradley et al., 1996), has not been subject to an analysis of measurement equivalence/invariance cross culturally. Thus, the current study contributes to the existing research base in a unique way. In addition, the use of non-overlapping samples each for the EFA and CFA allowed for the built-in replication of findings.

No nationally representative data were collected from any country. Data were collected from one site in each country, which limited diversity of the sample, with respect to some factors, such as caste, ethnicity, SES, etc. The role of intersectionality (i.e., the confluence of multiple social influences) cannot be ruled out, thus the results of this project may not be applicable to each individual within a particular nation or community. Second, the current study used data on the home environment from only one time point (24 months). Future analysis should examine the longitudinal nature of the home environment, including its stability and cross cultural differences over time. Future studies may also wish to look at how practices within cultures differed across groups, which may have contributed to the psychometric differences observed in the current study. Finally, the current study demonstrates the psychometric strengths of the HOME when used cross-culturally. Utilizing the HOME, further studies should extend the current findings to examine differences in how the home environment impacts specific outcomes across cultures.

5. Conclusions

The goal of the current study was to: (1) examine the cross-cultural validity of the HOME measure among the international MAL-ED sites, and (2) adapt an existing tool for evaluating the home environment. This is the first evaluation of the HOME measure utilizing measurement equivalence analytic techniques. The results suggest that the adapted HOME may be useful in evaluating the home environment across heterogeneous groups. Future studies should seek to replicate and extend these findings, particularly the relationship between the home environment and other variables associated with childhood developmental outcomes. Furthermore, a cross-culturally validated adaptation of the HOME could then also be used to evaluate cultural differences in the home environment, and how those differences are related to such factors as academic outcomes.

Appendix A

Bangladesh site

The Bangladeshi site is located in the capital city of Dhaka, Bauniabadh area of Mirpur, within an urban, squatter settlement (slum) with one of the highest population densities in the world. Dhaka has a tropical climate with high temperatures (temperatures range from 71° F to 90° F, with a mean of 64.4° F, 18° C) and high humidity. There are three distinct seasons: the pre-monsoon hot season; the rainy season, and a cooler, dry winter. The average family size is 4.5 with gross monthly income around \$125. Overcrowding, poor house construction and poor sanitary disposals are common features of the community. Almost 63% families live in single roomed houses of metal and brick construction and 7.4% have mud or wooden floor. Around 99% of the households have access to supplied piped water in the dwellings or yard and 56% have latrine that flushes to septic tank. In regard to highest level of maternal education, 19% of mothers had completed no formal education, 47% had completed one to 5 years (primary school), 33% had completed six to 10 years, and 1% had completed secondary school or higher (> 10 years). The area has one UNICEF and one NGO operated pre-primary schools and around 25 Government and private high schools. The quality of pre-schools is limited due to large number of students. For more background and demographic information about the Bangladesh site, see Ahmed et al. (2014).

Brazil site

The Brazil site is located in the city of Fortaleza, the fifth most populated city in Brazil. Fortaleza is located in the state of Ceará. The climate of Ceará is semi-arid with low levels of rainfall throughout the year and an average temperature of 83 °F (29 °C). In this sample, the average number of people living in each home was 5.2. In regard to highest level of maternal education completed, 8% of mothers completed primary school (one to 5 years), 34% completed secondary school (six to 10 years), and 60% completed secondary school or higher (> 10 years). In regard to employment, most individuals in this area report being involved in local labor. However, it is also known that many inhabitants are involved in the narcotics trade; thus, the accuracy of the reports is unclear. In the local community, there are a few small private and one public child care pre-school centers. For more background and demographic information about the Brazil site, see Lima et al. (2014).

India site

The India site is located in Vellore, Tamil Nadu the southern region of India. The climate ranges from semi-humid to semiarid with three seasons: a southwest monsoon season, a northeast monsoon season, and a dry season with average temperatures ranging from 24 °C (74 °F) to 28 °C (83 °F) depending on the season. The average family size is 5.7. In the MAL-ED cohort at this site, the average education for the heads of household of MAL-ED families is 6.81 years (range 0 to 18 years). In regard to occupation, there is a high rate of unemployment at the India site (58.5%). For adults who are employed, most are involved in daily wage labor. For more background and demographic information about the India site, see [John et al. \(2014\)](#).

Nepal site

The Nepal site is located in the Bhakatapur municipality which is approximately 15 km outside of the capital city of Nepal, Kathmandu. The climate in this region is subtropical with a wet and hot rainy season (May to August) as well as a cool and dry season (October to March) with an average temperature of 95 °F, 35 °C. The nation of Nepal is diverse in regard to ethnic and caste groups. The majority of the citizens of Bhakatapur are members of the Newar ethnic group. In regard to family demographics, most families live in extended family systems. The average family size is 4.8, and most families are employed in agriculture. In regard to maternal education, 94% of the mothers at this site were literate but only a quarter of the mothers had completed more than 10 years of formal education. For more background and demographic information about the Nepal site, see [Shrestha et al. \(2014\)](#).

Pakistan site

The Pakistan site located in Naushahro Feroze which is a rural area in the Southern region of Pakistan. The summers are very hot (average temperature 113 °F, 45 °C). In regard to family demographics, most families live in extended family systems. The average number of children per family is 4 with an average family size of 6 to 7. Approximately 33% of mothers at this site area able to read, and the average maternal education is 2 years. Approximately 60–70% of fathers in this region are literate, and males tend to have greater access to educational opportunities. The primary occupation for men is farming and most women are homemakers or are involved in weaving and stitching. Approximately 60% of children are malnourished at 2 years of age. Children tend to start pre-school between the ages of 5 and 6. Schools in this region have highly limited resources. Some schools do not have electricity, lack basic materials, and have a high rate of teacher absenteeism. For more background and demographic information about the Pakistan site, see [Turab et al. \(2014\)](#).

Peru site

The Peru site is located along a string of semi-urban communities bordering the River Nanay 14 km from in the rural city of Iquitos in the Loreto province. Loreto is located in a low tropical rainforest climate. The average temperature is 78 °F, annual rainfall is 2.6 m a year with April being the typical month where the river reaches its peak and when flooding is common. In regard to maternal education at this site there in near-universal literacy, 1.1% of mothers had no formal education or did not respond, 21.1% had one to 5 years of education, 58.2% had six to 10 years of education, 14.1% completed their secondary education (> 10 years), and 5.6% completed come college or technical school. Primary occupations for men in this region including growing and selling vegetables, fishing, moto-taxi driving, and brick-making. The majority of women are homemakers, but some are employed as shopkeepers, maids, or work in local small scale commerce. At this site, the majority (94–99%) of children over age 6 attend school. Approximately 32.5% of families in the area reported enrollment in the JUNTOS program which is a national program that provides cash transfer to families with children under age 14 who maintain 85% school attendance, obtain civic identification documentation, maintain regular health check-ups, and receive vaccinations. For more background and demographic information about the Peru site, see [Yori et al. \(2014\)](#).

South Africa site

The South African site is located in the Dzimauli community within the Vhembe district in the Northern part of the Limpopo province. Overall, the Vhembe district has a warm, semi-arid climate. In regard to family demographics, mothers in the sample had an average of 2.4 live births. All participants in the sample reported their race as Black. On average, mothers in the sample had completed 9 years of formal education. Most participants from this site were engaged in subsistence farming, minor trading, or government agency work for more background and demographic information about the South African site, see [Bessong, Nyathi, Mahopo, and Netshandama \(2014\)](#).

Tanzania site

The Tanzania site is located in Haydom which is a rural area, approximately 300 km from the city of Arusha, in the North-Central region of Tanzania. The North-Central region of Tanzania has three seasons: a high season which is cool and dry; a shoulder season which is hot with some rain; and a low season with heavy rains. The average temperature is 67.1 °F (19.5 °C), but temperatures vary by season. In this sample, the average woman had 4.3 live births and 1.3 child deaths and there were an average of 6.7 people living

in each household. In regard to formal education, the mothers in this sample had completed 7 years of schooling on average; however, 15.3% of women in the sample had no formal schooling. Schools lack basic sanitary facilities and the teacher-pupil ratio is high, about 1 to 50. The area is populated by a diverse mix of agropastoralists, agriculturalist and pastoralists. For more background and demographic information about the Tanzania site, see [Mduma et al. \(2014\)](#).

HOME observation items in MAL-ED

Note. Item stems have been altered slightly to protect copyrighted material.

1. Caregiver spontaneously vocalizes to the child during the visit.
2. Caregiver verbally responds to child's vocalizations.
3. Caregiver tells the child the name of some person or object during the visit.
4. Caregiver's speech is clear.
5. Caregiver initiates verbal exchanges with the observer.
6. Caregiver expresses ideas and is conversational (i.e., gives more than brief answers).
7. Caregiver spontaneously praises child twice during visit.
8. When speaking to the child, caregiver's voice conveys positive feeling.
9. Caregiver does something affectionate with the child (hugs, pats, etc.) during the visit.
10. Caregiver shows some positive emotional response to the child.
11. Caregiver smiles at or laughs with the child.
12. Caregiver does not shout at the child.
13. Caregiver does not overtly display annoyance or hostility toward the child.
14. Caregiver does not slap or spank the child.
15. Caregiver reports no physical punishment within the past week.
16. Caregiver does not scold/criticize the child during the visit.
17. Caregiver keeps the child within visual range and looks at the child frequently.
18. Caregiver talks to the child while doing chores.
19. Caregiver intentionally encourages developmental advances.
20. Caregiver provides structure in the child's day.
21. Caregiver believes the child's behavior can be changed or modified by the parent.
22. When the primary caregiver is away, care is provided by one of three regular others.
23. Child is not cared for by another child (under age 12).
24. There is no evidence that anyone handles the child inappropriately.
25. Someone brings the child to a shop or market at least once a week.
27. The child is taken to the health clinic regularly.
28. The child has a special place to keep his toys and things
29. The child's play area is relatively safe and free from hazards.
30. The stove is located in a relatively safe area.
31. The house is relatively light.
32. The house is relatively ventilated.
33. The house is relatively clean.
34. The house is relatively neat and orderly.
35. There are appropriate play materials (e.g., toys, tins, balls, dolls, slates) in the house
36. The child has a riding toy for gross motor stimulation.
37. The caregiver provides toys or activities during the visit.
38. There are some magazines, newspapers, or books visible in the home.
39. The family has a pet.
40. The father plays with the child or is involved with caregiving every day.
41. The caregiver tells the child stories or nursery rhythms at least weekly.
42. The caregiver sings to the child daily.
43. The child eats at least one meal per day with the caregiver and the father.
44. The family has not moved more than once in the past year.
45. The family visits with relatives at least once per month.
46. The family visits with close friends at least once per month.
47. The child is relatively clean, with no offensive odor.
48. The child's hair is relatively clean.
49. The child's clothes are relatively clean.

Appendix B. MAL-ED investigators and institutional affiliations

Angel Mendez Acosta¹

Rosa Rios de Burga¹
 Cesar Banda Chavez¹
 Julian Torres Flores¹
 Maribel Paredes Olotegui¹
 Silvia Rengifo Pinedo¹
 Mery Sigvas Salas¹
 Dixner Rengifo Trigos¹
 Angel Orbe Vasquez¹
 Imran Ahmed²
 Didar Alam²
 Asad Ali²
 Zulfiqar A Bhutta²
 Shahida Qureshi²
 Muneera Rasheed²
 Sajid Soofi²
 Ali Turab²
 Anita KM Zaidi²
 Ladaporn Bodhidatta³
 Carl J Mason³
 Sudhir Babji⁴
 Anuradha Bose⁴
 Ajila T George⁴
 Dinesh Hariraju⁴
 M Steffi Jennifer⁴
 Sushil John⁴
 Shiny Kaki⁴
 Gagandeep Kang⁴
 Priyadarshani Karunakaran⁴
 Beena Koshy⁴
 Robin P Lazarus⁴
 Jayaprakash Muliyil⁴
 Mohan Venkata Raghava⁴
 Sophy Raju⁴
 Anup Ramachandran⁴
 Rakhi Ramadas⁴
 Karthikeyan Ramanujam⁴
 Anuradha Rose⁴
 Reeba Roshan⁴
 Srujan L Sharma⁴
 Shanmuga Sundaram E⁴
 Rahul J Thomas⁴
 William K Pan^{5,6}
 Ramya Ambikapathi⁶
 J Daniel Carreon⁶
 Vivek Charu⁶
 Viyada Doan⁶
 Jhanelle Graham⁶
 Christel Hoest⁶
 Stacey Knobler⁶
 Dennis R Lang^{6, 7}
 Benjamin JJ McCormick⁶
 Monica McGrath⁶
 Mark A Miller⁶
 Archana Mohale⁶
 Gaurvika Nayyar⁶
 Stephanie Psaki⁶
 Zeba Rasmussen⁶
 Stephanie A Richard⁶
 Jessica C Seidman⁶
 Vivian Wang⁶
 Rebecca Blank⁷

Michael Gottlieb⁷
Karen H Tountas⁷
Caroline Amour⁸
Eliwaza Bayyo⁸
Estomih R Mduma⁸
Regisiana Mvungi⁸
Rosemary Nshama⁸
John Pascal⁸
Buliga Mujaga Swema⁸
Ladislaus Yarrot⁸
Tahmeed Ahmed⁹
AM Shamsir Ahmed⁹
Rashidul Haque⁹
Iqbal Hossain⁹
Munirul Islam⁹
Mustafa Mahfuz⁹
Dinesh Mondal⁹
Fahmida Tofail⁹
Ram Krishna Chandyo¹⁰
Prakash Sunder Shrestha¹⁰
Rita Shrestha¹⁰
Manjeswori Ulak¹⁰
Aubrey Bauck¹¹
Robert Black¹¹
Laura Caulfield¹¹
William Checkley^{11,6}
Margaret N Kosek¹¹
Gwenyth Lee¹¹
Kerry Schulze¹¹
Pablo Peñataro Yori¹¹
Laura E Murray-Kolb¹²
A Catharine Ross¹²
Barbara Schaefer^{12,6}
Suzanne Simons¹²
Laura Pendergast¹³
Cláudia B Abreu¹⁴
Hilda Costa¹⁴
Alessandra Di Moura¹⁴
José Quirino Filho^{14,6}
Alexandre Havt¹⁴
Álvaro M Leite¹⁴
Aldo AM Lima¹⁴
Noélia L Lima¹⁴
Ila F Lima¹⁴
Bruna LL Maciel¹⁴
Pedro HQS Medeiros¹⁴
Milena Moraes¹⁴
Francisco S Mota¹⁴
Reinaldo B Oriá¹⁴
Josiane Quetz¹⁴
Alberto M Soares¹⁴
Rosa MS Mota¹⁴
Crystal L Patil¹⁶
Pascal Bessong¹⁷
Cloupas Mahopo¹⁷
Angelina Maphula¹⁷
Emanuel Nyathi¹⁷
Amidou Samie¹⁷
Leah Barrett¹⁸
Rebecca Dillingham¹⁸
Jean Gratz¹⁸

Richard L Guerrant¹⁸
 Eric Houpt¹⁸
 William A Petri, Jr¹⁸
 James Platts-Mills¹⁸
 Rebecca Scharf¹⁸
 Binob Shrestha¹⁹
 Sanjaya Kumar Shrestha¹⁹
 Tor Strand^{19,15}
 Erling Svensen^{20,8}

Institutions

- ¹A.B. PRISMA, Iquitos, Peru
²Aga Khan University, Karachi, Pakistan
³Armed Forces Research Institute of Medical Sciences, Bangkok, Thailand
⁴Christian Medical College, Vellore, India
⁵Duke University, Durham, NC, USA
⁶Fogarty International Center/National Institutes of Health, Bethesda, MD, USA
⁷Foundation for the NIH, Bethesda, MD, USA
⁸Haydom Lutheran Hospital, Haydom, Tanzania
⁹icddr,b, Dhaka, Bangladesh
¹⁰Institute of Medicine, Tribhuvan University, Kathmandu, Nepal
¹¹Johns Hopkins University, Baltimore, MD, USA
¹²The Pennsylvania State University, University Park, PA, USA
¹³Temple University, Philadelphia, PA, USA
¹⁴Universidade Federal do Ceara, Fortaleza, Brazil
¹⁵University of Bergen, Norway
¹⁶University of Illinois at Chicago, IL, USA
¹⁷University of Venda, Thohoyandou, South Africa
¹⁸University of Virginia, Charlottesville, VA, USA
¹⁹Walter Reed/AFRIMS Research Unit, Kathmandu, Nepal
²⁰Haukeland University Hospital, Bergen, Norway

References

- Ahmed, T., Mahfuz, M., Islam, M. M., Mondal, D., Hossain, M. I., Ahmed, A. S., ... Petri, W. A. (2014). The MAL-ED cohort study in Mirpur, Bangladesh. *Clinical Infectious Diseases*, 59(Suppl. 4), S280–S286. <http://dx.doi.org/10.1093/cid/ciu458>.
- Aiello, A. E., & Larson, E. L. (2002). What is the evidence for a causal link between hygiene and infections? *Lancet Infectious Diseases*, 2, 103–110. [http://dx.doi.org/10.1016/S1473-3099\(02\)00184-6](http://dx.doi.org/10.1016/S1473-3099(02)00184-6).
- American Educational Research Association, American Psychological Association, and the National Council on Measurement in Education (2014). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.
- Andrade, S. A., Santos, D. N., Bastos, A. C., Pedromônico, M. R. M., de Almeida-Filho, N., & Barreto, M. L. (2005). Family environment and child's cognitive development: An epidemiological approach. *Revista de Saúde Pública*, 39, 606–611. <http://dx.doi.org/10.1590/S0034-89102005000400014>.
- Asparouhov, T., & Muthén, B. (2014). Multiple-group factor analysis alignment. *Structural Equation Modeling: A Multidisciplinary Journal*, 21, 495–508. <http://dx.doi.org/10.1080/10705511.2014.919210>.
- Baillie, R. S., Stevens, M. R., McDonald, E., Halpin, S., Brewster, D., Robinson, G., & Guthridge, S. (2005). Skin infection, housing and social circumstances in children living in remote indigenous communities: Testing conceptual and methodological approaches. *BMC Public Health*, 5, 128. <http://dx.doi.org/10.1186/1471-2458-5-128>.
- Bessong, P. O., Nyathi, E., Mahopo, T. C., & Netshandama, V. (2014). Development of the dzimauli community in Vhembe district, Limpopo province of South Africa, for the MAL-ED cohort study. *Clinical Infectious Diseases*, 59(Suppl. 4), S317–S324. <http://dx.doi.org/10.1093/cid/ciu418>.
- Bingenheimer, J. B., Raudenbush, S. W., Leventhal, T., & Brooks-Gunn, J. (2005). Measurement equivalence and differential item functioning in family psychology. *Journal of Family Psychology*, 19, 441–455. <http://dx.doi.org/10.1037/0893-3200.19.3.441>.
- Black, M. M., Baqui, A., Zaman, K., Persson, L., Arifeen, S., Le, K., ... Black, R. (2004). Iron and zinc supplementation promote motor development and exploratory behavior among Bangladeshi infants 1-3. *American Journal of Clinical Nutrition*, 80, 903–910. Retrieved from <http://ajcn.nutrition.org/content/80/4/903.full>.
- Blair, C., Raver, C. C., & Berry, D. J. (2014). Two approaches to estimating the effect of parenting on the development of executive function in early childhood. *Developmental Psychology*, 50, 554–565. <http://dx.doi.org/10.1037/a0033647>.
- Bracken, B. A., & Barona, A. (1991). State of the art procedures for translating, validating and using psychoeducational tests in cross-cultural assessment. *School Psychology International*, 12, 119–132. <http://dx.doi.org/10.1177/0143034391121010>.
- Bradley, R. H., & Caldwell, B. M. (1976). The relation of infants' home environments to mental test performance at fifty-four months: A follow-up study. *Child Development*, 47, 1172–1174. Retrieved from <http://www.jstor.org/stable/1128457>.
- Bradley, R. H., & Caldwell, B. M. (1984). 174 children: A study of the relationship between home environment and cognitive development during the first 5 years. In A. W. Gottfried (Ed.), *Home environment and early cognitive development: Longitudinal research* (pp. 5–56). Orlando, FL: Academic Press.
- Bradley, R. H., Caldwell, B. M., & Rock, S. L. (1988). Home environment and school performance: A ten-year follow-up and examination of three models of environmental action. *Child Development*, 59, 852–867. Retrieved from <http://www.jstor.org/stable/1130253>.
- Bradley, R. H., & Corwyn, R. F. (2002). Socioeconomic status and child development. *Annual Review of Psychology*, 53, 371–399. <http://dx.doi.org/10.1146/annurev.psych.53.100901.135233>.
- Bradley, R. H., Corwyn, R. F., Burchinal, M., McAdoo, H. P., & Garcia Coll, C. (2001). The home environments of children in the United States part II: Relations with behavioral development through age thirteen. *Child Development*, 72, 1868–1886. <http://dx.doi.org/10.1111/1467-8624.t01-1-00383>.

- Bradley, R. H., Corwyn, R. F., & Whiteside-Mansell, L. (1996). Life at home: Same time, different places — An examination of the HOME inventory in different cultures. *Early Development and Parenting*, 5, 251–269. [http://dx.doi.org/10.1002/\(SICI\)1099-0917\(199612\)5:4<251::AID-EDP137>3.0.CO;2-I](http://dx.doi.org/10.1002/(SICI)1099-0917(199612)5:4<251::AID-EDP137>3.0.CO;2-I).
- Bradley, R. H., Mundfrom, D. J., Whiteside, L., Casey, P. H., & Barrett, K. (1994a). A factor analytic study of the infant-toddler and early childhood versions of the HOME inventory administered to white, black and Hispanic American parents of children born preterm. *Child Development*, 65, 880–888. <http://dx.doi.org/10.1111/j.1467-8624.1994.tb00790.x>.
- Bradley, R. H., & Putnick, D. L. (2012). Housing quality and access to material and learning resources within the home environment in developing countries. *Child Development*, 83, 76–91. <http://dx.doi.org/10.1111/j.1467-8624.2011.01674.x>.
- Bradley, R. H., Whiteside, L., Mundfrom, D. J., Casey, P. H., Kelleher, K. J., & Pope, S. K. (1994b). Early indications of resilience and their relation to experiences in the home environments of low birthweight, premature children living in poverty. *Child Development*, 65, 346–360. <http://dx.doi.org/10.1111/1467-8624.ep9405315114>.
- Briggs, N. E., & MacCallum, R. C. (2003). Recovery of weak common factors by maximum likelihood and ordinary least squares estimation. *Multivariate Behavioral Research*, 38(1), 25–56. http://dx.doi.org/10.1207/S15327906MBR3801_2.
- Bronfenbrenner, U. (1977). Toward an experimental ecology of human development. *American Psychologist*, 32, 513–531. <http://dx.doi.org/10.1037/0003-066X.32.7.513>.
- Brooks-Gunn, J., Klebanov, P. K., & Liaw, F. (1995). The learning, physical, and emotional environment of the home in the context of poverty: The infant health and development program. *Children and Youth Services Review*, 17, 251–276. [http://dx.doi.org/10.1016/0190-7409\(95\)00011-Z](http://dx.doi.org/10.1016/0190-7409(95)00011-Z).
- Brotman, L. M., Gouley, K. K., Chesir-Teran, D., Dennis, T., Klein, R. G., & Shrout, P. (2005). Prevention for preschoolers at high risk for conduct problems: Immediate outcomes on parenting practices and child social competence. *Journal of Clinical Child and Adolescent Psychology*, 34, 724–734. http://dx.doi.org/10.1207/s15374424jccp3404_14.
- Caldwell, B. M., & Bradley, R. H. (1984). *Home Observation for Measurement of the Environment: Administration Manual*. Little Rock, AK: University of Arkansas at Little Rock.
- Caldwell, B. M., & Bradley, R. H. (2003). *Home Observation for Measurement of the Environment: Administration manual*. Tempe, AZ: Family & Human Dynamics Research Institute, Arizona University.
- Cardoso, M. R. A., Cousens, S. N., de Góes Siqueira, L. F., Alves, F. M., & D'Angelo, L. A. V. (2004). Crowding: Risk factor or protective factor for lower respiratory disease in young children? *BMC Public Health*, 4, 19–26. <http://dx.doi.org/10.1186/1471-2458-4-19>.
- Cattell, R. B. (1966). The scree test for the number of factors. *Multivariate Behavioral Research*, 1, 245–276. http://dx.doi.org/10.1207/s15327906mbr0102_10.
- Cole, S. R., Kawachi, I., Maller, S. J., & Berkman, L. F. (2000). Test of item-response bias in the CES-D scale: Experience from the New Haven EPESE Study. *Journal of Clinical Epidemiology*, 53, 285–289. [http://dx.doi.org/10.1016/S0895-4356\(99\)00151-1](http://dx.doi.org/10.1016/S0895-4356(99)00151-1).
- Conley, D. (2001, June). A room with a view or a room of one's own? Housing and social stratification. *Sociological Forum*. Vol. 16, No. 2. *Sociological Forum* (pp. 263–280). Kluwer Academic Publishers-Plenum Publishers.
- Delgado, J., Ramirez-Cardich, M. E., Gilman, R. H., Lavarello, R., Dahodwala, N., Bazan, A., ... Lescano, A. (2002). Risk factors for burns in children: Crowding, poverty, and poor maternal education. *Injury Prevention*, 8, 38–41. <http://dx.doi.org/10.1136/ip.8.1.38>.
- Dragow, F. (1984). Scrutinizing psychological tests: Measurement equivalence and equivalent relations with external variables are the central issues. *Psychological Bulletin*, 95, 134–135. <http://dx.doi.org/10.1037/0033-2909.95.1.134>.
- Elardo, R., & Bradley, R. H. (1981). The Home Observation for Measurement of the Environment (HOME) Scale: A review of research. *Developmental Review*, 1, 322–329. [http://dx.doi.org/10.1016/0273-2297\(81\)90028-9](http://dx.doi.org/10.1016/0273-2297(81)90028-9).
- Elardo, R., Bradley, R. H., & Caldwell, B. M. (1975). The relation of infants' home environments to mental test performance from six to thirty-six months: A longitudinal analysis. *Child Development*, 46, 71–76. Retrieved from <http://www.jstor.org/stable/1128835>.
- Elardo, R., Bradley, R. H., & Caldwell, B. M. (1977). A longitudinal study of the relation of infants' home environments to language development at age three. *Child Development*, 48, 595–603. <http://dx.doi.org/10.1111/j.1467-8624.1977.tb01202.x>.
- Evans, G. W. (2006). Child development and the physical environment. *Annual Review of Psychology*, 57, 423–451. <http://dx.doi.org/10.1146/annurev.psych.57.102904.190057>.
- Evans, G. W., Ricciuti, H. N., Hope, S., Schoon, I., Bradley, R. H., Corwyn, R. F., & Hazan, C. (2010). Crowding and cognitive development: The mediating role of maternal responsiveness among 36-month-old children. *Environment and Behavior*, 42, 135–148. <http://dx.doi.org/10.1177/0013916509333509>.
- Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods*, 4, 272. <http://dx.doi.org/10.1037/1082-989X.4.3.272>.
- Ferguson, K. T., Cassells, R. C., MacAllister, J. W., & Evans, G. W. (2013). The physical environment and child development: An international review. *International Journal of Psychology*, 48, 437–468. <http://dx.doi.org/10.1080/00207594.2013.804190>.
- Frankenburg, W. K., & Coons, C. E. (1986). Home screening questionnaire: Its validity in assessing home environment. *The Journal of Pediatrics*, 108, 624–626. [http://dx.doi.org/10.1016/S0022-3476\(86\)80853-8](http://dx.doi.org/10.1016/S0022-3476(86)80853-8).
- Geisinger, K. F. (1994). Cross-cultural normative assessment: Translation and adaptation issues influencing the normative interpretation of assessment instruments. *Psychological Assessment*, 6, 304–312. <http://dx.doi.org/10.1037/1040-3590.6.4.304>.
- Grandjean, P., & Landrigan, P. J. (2006). Developmental neurotoxicity of industrial chemicals. *Lancet*, 368, 2167–2178. [http://dx.doi.org/10.1016/S0140-6736\(06\)69665-7](http://dx.doi.org/10.1016/S0140-6736(06)69665-7).
- Guerrant, R. L., DeBoer, M. D., Moore, S. R., Scharf, R. J., & Lima, A. A. M. (2013). The impoverished gut—A triple burden of diarrhoea, stunting and chronic disease. *Nature Reviews Gastroenterology & Hepatology*, 10(4), 220–229. <http://dx.doi.org/10.1038/nrgastro.2012.239>.
- Guerrant, R. L., Oriá, R. B., Moore, S. R., Oriá, M. O. B., & Lima, A. A. M. (2008). Malnutrition as an enteric infectious disease with long-term effects on child development. *Nutrition Reviews*, 66, 487–505. <http://dx.doi.org/10.1111/j.1753-4887.2008.00082.x>.
- Hambleton, R. K. (2001). The next generation of the ITC test translation and adaptation guidelines. *European Journal of Psychological Assessment*, 17, 164–172. <http://dx.doi.org/10.1027//1015-5759.17.3.164>.
- Henson, R. K., & Roberts, J. K. (2006). Use of exploratory analysis in published research: Common errors and some comments on improved practice. *Educational and Psychological Measurement*, 66, 393–416. <http://dx.doi.org/10.1177/0013164405282485>.
- Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika*, 30, 179–185. <http://dx.doi.org/10.1007/BF02289447>.
- Hu, L. T., & Bentler, P. M. (1995). Evaluating model fit. In R. H. Hoyle (Ed.), *Structural equation modeling: Concepts, issues, and applications* (pp. 76–99). Thousand Oaks, CA: Sage Publications.
- Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, 6, 1–55. <http://dx.doi.org/10.1080/10705519909540118>.
- Iltus, S. (2006). *Significance of home environments as proxy indicators for early childhood care and education*. Education for all global monitoring report 2007 New York: NY.
- Jackson, D. N., & Messick, S. (1958). Content and style in personality assessment. *Psychological Bulletin*, 55, 243–252. <http://dx.doi.org/10.1037/h0045996>.
- Jaine, R., Baker, M., & Venugopal, K. (2011). Acute rheumatic fever associated with household crowding in a developed country. *The Pediatric Infectious Disease Journal*, 30, 315–319. <http://dx.doi.org/10.1097/INF.0b013e3181fbd85b>.
- Jimerson, S. R., Annan, J., Skokut, M., & Renshaw, T. L. (2009). Educational psychology in New Zealand: Results of the 2006 International School Psychology Survey. *School Psychology International*, 30(5), 443–455. <http://dx.doi.org/10.1177/0143034309341617>.
- John, S. M., Thomas, R. J., Kaki, S., Sharma, S. L., Ramanujam, K., Raghava, M. V., ... Ramchandran, A. (2014). Establishment of the MAL-ED birth cohort study site in Vellore, southern India. *Clinical Infectious Diseases*, 59(Suppl. 4), S295–S299. <http://dx.doi.org/10.1093/cid/ciu390>.
- Johnson, D. L., Breckenridge, J. N., & McGowan, R. J. (1984). Home environment and early cognitive development in Mexican-American children. In A. W. Gottfried (Ed.), *Home Environment and Early Cognitive Development: Longitudinal Research* (pp. 151–196). Orlando, FL: Academic Press.
- Jones, R. N. (2006). Identification of measurement differences between English and Spanish language versions of the Mini-Mental State Examination. Detecting differential item functioning using MIMIC modeling. *Medical Care*, 44, S124–S133. <http://dx.doi.org/10.1097/01.mlr.0000245250.501114.f0>.
- Jöreskog, K. G., & Goldberger, A. S. (1975). Estimation of a model with multiple indicators and multiple causes of a single latent variable. *Journal of the American*

- Statistical Association, 70, 631–639. <http://dx.doi.org/10.1080/01621459.1975.10482485>.
- Jusko, T. A., Henderson, C. R., Lanphear, B. P., Cory-Slechta, D. A., Parsons, P. J., & Canfield, R. L. (2008). Blood lead concentrations 10 μ g/dL and child intelligence at 6 years of age. *Environmental Health Perspectives*, 116, 243–248. <http://dx.doi.org/10.1289/ehp.10424>.
- Kaiser, H. F. (1974). An index of factorial simplicity. *Psychometrika*, 39, 31–36. <http://dx.doi.org/10.1007/BF02291575>.
- Kippler, M., Tofail, F., Hamadani, J. D., Gardner, R. M., Grantham-McGregor, S. M., Bottai, M., & Vahter, M. (2012). Early-life cadmium exposure and child development in 5-year-old girls and boys: A cohort study in rural Bangladesh. *Environmental Health Perspectives*, 120, 1462–1468. <http://dx.doi.org/10.1289/ehp.1104431>.
- Krosnick, J. A. (1999). Maximizing questionnaire quality. In J. P. Robinson, P. R. Shaver, & L. S. Wrightsman (Eds.), *Measures of Political Attitudes* (pp. 37–58). San Diego, CA: Academic Press.
- Lanphear, B. P., Hornung, R., Khoury, J., Yolton, K., Baghurst, P., Bellinger, D. C., ... Roberts, R. (2005). Low-level environmental lead exposure and children's intellectual function: An international pooled analysis. *Environmental Health Perspectives*, 113, 894–899. <http://dx.doi.org/10.1289/ehp.7688>.
- Leventhal, T., & Brooks-Gunn, J. (2000). The neighborhoods they live in: The effects of neighborhood residence on child and adolescent outcomes. *Psychological Bulletin*, 126, 309–337. <http://dx.doi.org/10.1037/0033-2909.126.2.309>.
- Leventhal, T., Selner-O'Hagen, M. B., Brooks-Gunn, J., Bingenheimer, J. B., & Earls, F. J. (2004). The Homelife interview from the project on human development in Chicago neighborhoods: Assessment of parenting and home environment for 3- to 15-year-olds. *Parenting: Science and Practice*, 4, 211–241. <http://dx.doi.org/10.1080/15295192.2004.9681271>.
- Lima, A. A., Oriá, R. B., Soares, A. M., José Filho, Q., de Sousa, F., Abreu, C. B., ... Maciel, B. (2014). Geography, population, demography, socioeconomic, anthropometry, and environmental status in the MAL-ED cohort and case-control study sites in Fortaleza, Ceará, Brazil. *Clinical Infectious Diseases*, 59(Suppl. 4), S287–S294. <http://dx.doi.org/10.1093/cid/ciu438>.
- Lorenzo-Seva, U., & Ferrando, P. J. (2012). TETRA-COM: A comprehensive SPSS program for estimating the tetrachoric correlation. *Behavior Research Methods*, 44, 1191–1196. <http://dx.doi.org/10.3758/s13428-012-0200-6>.
- Luster, T., & Dubow, E. (1992). Home environment and maternal intelligence as predictors of verbal intelligence: A comparison of preschool and school-age children. *Merrill-Palmer Quarterly*, 38(2), 151–175. Retrieved from <http://www.jstor.org/stable/23087379>.
- Markland, D. (2007). The golden rule is that there are no golden rules: A commentary on Paul Barrett's recommendations for reporting model fit in structural equation modelling. *Personality and Individual Differences*, 42, 851–858. <http://dx.doi.org/10.1016/j.paid.2006.09.023>.
- Marsh, H. W. (1985). The structure of masculinity/femininity: An application of confirmatory factor analysis to higher-order factor structures and factorial invariance. *Multivariate Behavioral Research*, 20, 427–449. <http://dx.doi.org/10.1207/s15327906mbr2004.5>.
- Marsh, H. W. (1987). The factorial invariance of responses by males and females to a multidimensional self-concept instrument: Substantive and methodological issues. *Multivariate Behavioral Research*, 22, 457–480. <http://dx.doi.org/10.1207/s15327906mbr2204.5>.
- Marsh, H. W., Hau, K. T., & Wen, Z. (2004). In search of golden rules: Comment on hypothesis testing approaches to setting cutoff values for fit indexes and dangers in overgeneralizing Hu and Bentler's (1999) findings. *Structural Equation Modeling: A Multidisciplinary Journal*, 11, 320–341. <http://dx.doi.org/10.1207/s15328007sem1103.2>.
- Marsh, H. W., & Hocevar, D. (1985). Application of confirmatory factor analysis to the study of self-concept: First- and higher order factor models and their invariance across groups. *Psychological Bulletin*, 97, 562–582. <http://dx.doi.org/10.1037/0033-2909.97.3.562>.
- Matheny, A. P., Wachs, T. D., Ludwig, J. L., & Phillips, K. (1995). Bringing order out of chaos: Psychometric characteristics of the confusion, hubbub, and order scale. *Journal of Applied Developmental Psychology*, 16, 429–444. [http://dx.doi.org/10.1016/0193-3973\(95\)90028-4](http://dx.doi.org/10.1016/0193-3973(95)90028-4).
- Mduma, E. R., Gratz, J., Patil, C., Matson, K., Dakay, M., Liu, S., ... Ernst, A. (2014). The etiology, risk factors, and interactions of enteric infections and malnutrition and the consequences for child health and development study (MAL-ED): Description of the Tanzanian site. *Clinical Infectious Diseases*, 59(Suppl. 4), S325–S330. <http://dx.doi.org/10.1093/cid/ciu439>.
- Mendes, S. A., Abreu-Lima, I., Almeida, L. S., & Simeonsson, R. J. (2014). School psychology in Portugal: Practitioners' characteristics and practices. *International Journal of School & Educational Psychology*, 2(2), 115–125. <http://dx.doi.org/10.1080/21683603.2013.863171>.
- Murray-Kolb, L. E., Rasmussen, Z. A., Scharf, R. J., Rasheed, M. A., Svensen, E., Seidman, J. C., ... The MAL-ED Network Investigators (2014). The MAL-ED cohort study: Methods and lessons learned when assessing early child development and caregiving mediators in infants and young children in 8 low- and middle-income countries. *Clinical Infectious Diseases*, 59, S261–S272. <http://dx.doi.org/10.1093/cid/ciu437>.
- Muthén, B. O. (1989). Latent variable modeling in heterogeneous populations. *Psychometrika*, 54, 557–585. <http://dx.doi.org/10.1007/BF02296397>.
- National Research Council and Institute of Medicine (2000). From Neurons to Neighborhoods: The Science of Early Childhood Development. Committee on Integrating the Science of Early Childhood Development. In J. P. Shonkoff, & D. A. Phillips (Eds.), *Board on Children, Youth and Families, Commission on Behavioral and Social Sciences Education*. Washington, DC: National Academy Press.
- Needleman, H. L., Gunnoe, C., Leviton, A., Reed, R., Peresie, H., Maher, C., & Barrett, P. (1979). Deficits in psychologic and classroom performance of children with elevated dentine lead levels. *New England Journal of Medicine*, 300, 689–695. <http://dx.doi.org/10.1056/NEJM197903293001301>.
- NICHD Early Child Care Research Network (2001). Child-care and family predictors of preschool attachment and stability from infancy. *Developmental Psychology*, 37, 847–862. <http://dx.doi.org/10.1037/0012-1649.37.6.847>.
- Paulhus, D. L. (2002). Socially desirable responding: The evolution of a construct. In H. I. Braun, D. N. Jackson, & D. E. Wiley (Eds.), *The Role of Constructs in Psychological and Educational Measurement* (pp. 49–69). Mahwah, NJ: Erlbaum.
- Peña, E. D. (2007). Lost in translation: Methodological considerations in cross-cultural research. *Child Development*, 78, 1255–1264. <http://dx.doi.org/10.1111/j.1467-8624.2007.01064.x>.
- Pendergast, L. L., & Kaplan, A. (2015). Instructional context and student motivation, learning, and development: Commentary and implications for school psychologists. *School Psychology International*, 36, 638–647. <http://dx.doi.org/10.1177/0143034315613560>.
- Pendergast, L. L., Scharf, R. J., Rasmussen, Z. A., Seidman, J. C., Schaefer, B. A., Svensen, E., ... the MAL-ED Network Investigators (2014). Post-partum depression across time and place: The structural and temporal invariance of the SRQ-20 among women from multi-site, international MAL-ED study. *Journal of Affective Disorders*, 167, 178–186. <http://dx.doi.org/10.1016/j.jad.2014.05.039>.
- Pendergast, L. L., von der Embse, N., Kilgus, S., & Eklund, K. (2017). Measurement equivalence: A non-technical primer on categorical multi-group confirmatory factor analysis in school psychology. *Journal of School Psychology*, 60, 65–82. <http://dx.doi.org/10.1016/j.jsp.2016.11.002>.
- Prodromidis, M., Lamb, M. E., Sternberg, K. J., Hwang, C. P., & Broberg, A. G. (1995). Aggression and noncompliance among Swedish children in centre-based care, family day care, and home care. *International Journal of Behavioral Development*, 18(1), 43–62. <http://dx.doi.org/10.1177/016502549501800103>.
- Rasheed, M. A., & Yousafzai, A. K. (2015). The development and reliability of an observational tool for assessing mother-child interactions in field studies: experience from Pakistan. *Child: Care, Health and Development*, 41, 1161–1171. <http://dx.doi.org/10.1111/cch.12287>.
- Riordan, C. M., & Vandenberg, R. J. (1994). A central question in cross-cultural research: Do employees of different cultures interpret work-related measures in an equivalent manner? *Journal of Management*, 20, 643–671. [http://dx.doi.org/10.1016/0149-2063\(94\)90007-8](http://dx.doi.org/10.1016/0149-2063(94)90007-8).
- Sameroff, A. J., & Fiese, B. H. (2000). Transactional regulation: The developmental ecology of early intervention. In J. P. Shonkoff, & S. J. Meisels (Eds.), *Handbook of early childhood intervention* (pp. 135–159). Cambridge, UK: Cambridge University Press.
- Saudino, K. J., & Plomin, R. (1997). Cognitive and temperamental mediators of genetic contributions to the home environment during infancy. *Merrill-Palmer Quarterly*, 43, 1–23. Retrieved from <http://www.jstor.org/stable/23093725>.
- Shaw, M. (2004). Housing and public health. *Annual Review of Public Health*, 25, 397–418. <http://dx.doi.org/10.1146/annurev.publhealth.25.101802.123036>.
- Shaw, D. S., & Vondra, J. I. (1995). Infant attachment security and maternal predictors of early behavior problems: A longitudinal study of low-income families. *Journal of Abnormal Child Psychology*, 23, 335–357. <http://dx.doi.org/10.1007/BF01447561>.
- Sheridan, S. M., & Gutkin, T. B. (2000). The ecology of school psychology: Examining and changing our paradigm for the 21st century. *School Psychology Review*, 29, 485–502. Retrieved from <http://search.proquest.com/docview/219658013?accountid=14270>.
- Shih, C. L., & Wang, W. C. (2009). Differential item functioning detection using the multiple indicators, multiple causes method with a pure short anchor. *Applied Psychological Measurement*, 33, 184–199. <http://dx.doi.org/10.1177/0146621608321758>.

- Shonkoff, J. P., Boyce, W. T., & McEwen, B. S. (2009). Neuroscience, molecular biology, and the childhood roots of health disparities. *The Journal of the American Medical Association*, 301, 2252–2259. <http://dx.doi.org/10.1001/jama.2009.754>.
- Shonkoff, J. P., Garner, A. S., Siegel, B. S., Dobbins, M. I., Earls, M. F., Garner, A. S., ... Wood, D. L. (2012). The lifelong effects of early childhood adversity and toxic stress. *Pediatrics*, 129, e232–e246. <http://dx.doi.org/10.1542/peds.2011-2663>.
- Shrestha, P. S., Shrestha, S. K., Bodhidatta, L., Strand, T., Shrestha, B., Shrestha, R., ... Mason, C. J. (2014). Bhaktapur, Nepal: The MAL-ED birth cohort study in Nepal. *Clinical Infectious Diseases*, 59(Suppl. 4), S300–S303. <http://dx.doi.org/10.1093/cid/ciu459>.
- Solari, C. D., & Mare, R. D. (2012). Housing crowding effects on children's well-being. *Social Science Research*, 41, 464–476. <http://dx.doi.org/10.1016/j.ssresearch.2011.09.012>.
- Stevens, J. P. (2002). *Applied multivariate statistics for the social sciences* (4th ed). Mahwah, NJ: Lawrence Erlbaum.
- Stevens, J. H., & Bakeman, R. (1985). A factor analytic study of the HOME scale for infants. *Developmental Psychology*, 21, 1196–1203. <http://dx.doi.org/10.1037/0012-1649.21.6.1196>.
- Tanaka, J. S., & Long, J. S. (1993). Multifaceted conceptions of fit in structural equation models. In K. A. Bollen (Ed.), *Testing Structural Equation Models*. Newbury Park, CA: Sage.
- The MAL-ED Network Investigators (2014). The Malnutrition and Enteric Disease Study (MAL-ED): Understanding the consequences for child health and development [supplement]. *Clinical Infectious Diseases*, 59(Suppl. 4), S193–S330. <http://dx.doi.org/10.1093/cid/ciu804>.
- Totsika, V., & Sylva, K. (2004). The home observation for measurement of the environment revisited. *Child and Adolescent Mental Health*, 9, 25–35. <http://dx.doi.org/10.1046/j.1475-357X.2003.00073.x>.
- Turab, A., Soofi, S. B., Ahmed, I., Bhatti, Z., Zaidi, A. K., & Bhutta, Z. A. (2014). Demographic, socioeconomic, and health characteristics of the MAL-ED network study site in rural Pakistan. *Clinical Infectious Diseases*, 59(Suppl. 4), S304–S309. <http://dx.doi.org/10.1093/cid/ciu391>.
- Van De Vijver, F. J. R., & Poortinga, Y. H. (1982). Cross-cultural generalization and universality. *Journal of Cross-Cultural Psychology*, 13, 387–408. <http://dx.doi.org/10.1177/0022002182013004001>.
- Van Widenfelt, B. M., Treffers, P. D. A., De Beurs, E., Siebelink, B. M., & Koudijs, E. (2005). Translation and cross-cultural adaptation of assessment instruments used in psychological research with children and families. *Clinical Child and Family Psychology Review*, 8, 135–147. <http://dx.doi.org/10.1007/s10567-005-4752-1>.
- Vandenberg, R. J., & Lance, C. E. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organizational Research Methods*, 3, 4–70. <http://dx.doi.org/10.1177/109442810031002>.
- Velicer, W. F. (1976). Determining the number of components from the matrix of partial correlations. *Psychometrika*, 41, 321–327. <http://dx.doi.org/10.1007/BF02293557>.
- Walker, S. P., Wachs, T. D., Meeks Gardner, J., Lozoff, B., Wasserman, G. A., Pollitt, E., & Carter, J. A. (2007). Child development: Risk factors for adverse outcomes in developing countries. *The Lancet*, 369, 145–157. [http://dx.doi.org/10.1016/S0140-6736\(07\)60076-2](http://dx.doi.org/10.1016/S0140-6736(07)60076-2).
- Walkowiak, J., Wiener, J. A., Fastabend, A., Heinzow, B., Krämer, U., Schmidt, E., ... Winneke, G. (2001). Environmental exposure to polychlorinated biphenyls and quality of the home environment: Effects on psychodevelopment in early childhood. *Lancet*, 358, 1602–1607. [http://dx.doi.org/10.1016/S0140-6736\(01\)06654-5](http://dx.doi.org/10.1016/S0140-6736(01)06654-5).
- Watkins, M. W. (2006). Determining parallel analysis criteria. *Journal of Modern Applied Statistical Methods*, 5, 344–346. Retrieved from <http://digitalcommons.wayne.edu/jmasm/vol5/iss2/8>.
- Whiteside-Mansell, L., Bradley, R. H., Tresch Owen, M., Randolph, S. M., & Cauce, A. M. (2003). Parenting and children's behavior at 36 months: Equivalence between African American and European American mother - child dyads. *Parenting: Science and Practice*, 3, 197–234. http://dx.doi.org/10.1207/S15327922PAR0303_02.
- Williams, P., Piamjariyakul, U., Williams, A., Hornboonherm, P., Meena, P., Channukool, N., ... Chamusri, S. (2003). Thai mothers and children and the Home Observation for Measurement of the Environment (HOME inventory): Pilot study. *International Journal of Nursing Studies*, 40, 249–258. [http://dx.doi.org/10.1016/S0020-7489\(02\)00083-4](http://dx.doi.org/10.1016/S0020-7489(02)00083-4).
- Woods, C. M., Oltmanns, T. F., & Turkheimer, E. (2009). Illustration of MIMIC-model DIF testing with the schedule for nonadaptive and adaptive personality. *Journal of Psychopathology and Behavioral Assessment*, 31, 320–330. <http://dx.doi.org/10.1007/s10862-008-9118-9>.
- Yori, P. P., Lee, G., Olórtégui, M. P., Chávez, C. B., Flores, J. T., Vasquez, A. O., ... Caulfield, L. E. (2014). Santa Clara de Nanay: The MAL-ED cohort in Peru. *Clinical Infectious Diseases*, 59(Suppl. 4), S310–S316. <http://dx.doi.org/10.1093/cid/ciu460>.
- Zumbo, B. D. (2007). Three generations of DIF analyses: Considering where it has been, where it is now, and where it is going. *Language Assessment Quarterly*, 4, 223–233. <http://dx.doi.org/10.1080/15434300701375832>.