March 24, 2020

Evaluating the cross-cultural validity of three family quality of life sub-scales

Renee Hepperlen

Paula A Rabaey, *St. Catherine University*

Mary Hearst, *St. Catherine University*

Article ID: MNE 203205762

Processed by Minitex on: 6/4/2020 9:34:39 AM

This material comes to you from the University of Minnesota collection or another participating library of the Minitex Library Information Network.

Patrons, please contact your library for questions about this document.

Libraries, for more information, visit: <http://minitex.umn.edu>

If you have any questions about this service, please email mtx-edel@umn.edu or call 612-625-8318.

Title: Journal of applied research in intellectual disabilities : JARID.

ArticleTitle: Evaluating the cross-cultural validity of three family quality of life sub-scales.

ArticleAuthor: Hepperlen, Renee A

Date: March 2020 Pages: -

OCLC - 35007550; ISSN - 13602322;

Publisher: Clevedon : BILD Publications, ©1996-

Source: 35007550 <TN:129607><ODYSSEY:206.107.43.36/MNES> OCLC

Copyright: CCG

NOTICE CONCERNING COPYRIGHT RESTRICTIONS:

The copyright law of the United States [[Title 17, United StatesCode](#)] governs the making of photocopies or other reproductions of copyrighted materials.

Under certain conditions specified in the law, libraries and archives are authorized to furnish a photocopy or other reproduction. One of these specific conditions is that the photocopy is not to be "used for any purpose other than private study, scholarship, or research." If a user makes a request for, or later uses, a photocopy or reproduction for purposes in excess of "fair use," that user may be liable for copyright infringement.

This institution reserves the right to refuse to accept a copying order if, in its judgment, fulfillment of that order would involve violation of copyright law.

ORIGINAL ARTICLE

JARID

Journal of Applied Research in Intellectual Disabilities

bild

WILEY

Evaluating the cross-cultural validity of three family quality of life sub-scales

Renee A. Hepperlen¹ | Paula Rabaey² | Mary O. Hearst³ ¹School of Social Work, University of St. Thomas, Saint Paul, MN, USA²Occupational Therapy Department, St. Catherine University, Saint Paul, MN, USA³Public Health Department, St. Catherine University, Saint Paul, MN, USA**Correspondence**

Renee A. Hepperlen, School of Social Work, University of St. Thomas, 2115 Summit Ave., St. Paul, MN 55105, USA.

Email: hepp1989@stthomas.edu

Funding information

This study received funding from the GHR Foundation.

Abstract

Background: Families of children with disabilities often face unique challenges. Developed in a U.S. context, the Beach Center Family Quality of Life measure assesses the effectiveness of supports and services that families receive. This study examines whether items from three sub-scales of the Beach Center instrument perform similarly for two samples, one from Lusaka, Zambia, and the second from a Midwestern U.S. state.

Methods: This cross-sectional research used secondary data and completed hierarchical ordinal regression analyses on item-level performance within the sub-scales.

Results: Only one item flagged for potential item bias with remaining items performing similarly when controlling for overall sub-scale scores.

Conclusions: This study extends existing research on the cultural and linguistic appropriateness of the Beach Center measure, providing additional validity evidence about the internal structure of the scales. Findings indicate that these items are acceptable outcome measures for policy and programme evaluations in Zambia.

KEYWORDS

children with disabilities, cross-cultural validity, differential item functioning, family quality of life, low- and middle-income country, Zambia

1 | INTRODUCTION

The United Nations Council on Disabilities ratified the Convention on the Rights of Persons with Disabilities in 2008, which strongly advocates for effective programmes to support community-based services and supports for people with disabilities. When a person with a disability is a minor, the target of these supports often includes the family. The World Health Organization (WHO) and World Bank (2011) estimate that worldwide 93 million children between the ages of 0 and 15 have a moderate-to-severe disability, with higher proportions of disability reported in low- and middle-income countries (Maulik, Mascarenhas, Mathers, Dua, & Saxena, 2011; Scior, 2016). With growing interest in developing strategies to support families of children with disabilities in low- and middle-income countries, it is essential to evaluate interventions to assure that they are high-quality and beneficial (WHO & World Bank, 2011). To establish the effectiveness of these programmes, researchers should demonstrate

validity evidence of outcome measures, particularly when applying measures to a new cross-cultural context to ensure that the measures are linguistically and culturally appropriate.

In the intellectual and developmental disabilities field, a primary working principle has been the consideration of the quality of life for individuals who receive services and supports. The *quality of life* (QOL) construct has become better defined and studied since its introduction in the 1980s, with broad agreement that it is multidimensional and hierarchical (Claes, Van Hove, van Loon, Vandevelde, & Schalock, 2010; Cummins, 2005; Rapley, 2003; Schalock, 2004; Schalock et al., 2002). QOL considers a systems perspective, incorporating micro-, meso- and macro-level perspectives, with most research concentrated in the macro-level domain (Lyons, 2010). Micro-level considerations relate to work with individuals, families, or small groups (Miley, O'Melia, & DuBoise, 2016), and examples of topics include evaluation of friendships and interpersonal interactions. Meso-level applications focus on organizations and formal

groups (Miley et al., 2016) with questions capturing programme-level supports. Finally, macro-level questions evaluate the community and policies (Miley et al., 2016) with questions about access to health care and community safety. The quality of life approach affirms that people with disabilities should experience high levels of well-being (Samuel, Rillotta, & Brown, 2012). Samuel and colleagues noted that while quality of life research for persons with disabilities focused initially on the individual/person level, over the last two decades, the inquiry has shifted to include family members, due to the growing recognition that family members are essential natural supports for people with disabilities.

Building on the original quality of life literature, the *family quality of life* (FQOL) concept is in the formative stages of development. Researchers conceptualize FQOL as multidimensional, similar to the QOL concept (Poston et al., 2003; Samuel et al., 2012). However, FQOL also considers the interaction between services, the individual with a disability and the family (Isaacs, Brown, Brown, Baum, & Myer-Schough, 2007; Poston et al., 2003; Summers et al., 2005; Zuna, Summers, Turnbull, Hu, & Xu, 2010). Family systems theory and empowerment underpin FQOL, with the position that services and supports should enhance the entire family's functioning and not just consider the individual with a disability (Hu, Summers, Turnbull, & Zuna, 2011; Samuel et al., 2012; Summers et al., 2005; Wang & Brown, 2009). Zuna et al. (2010) described FQOL as "a dynamic sense of well-being of the family, collectively and subjectively defined and informed by its members in which individual and family-level needs interact" (p. 262). Investigators and clinicians view FQOL as a long-term outcome of interventions to support families of persons with disabilities (Summers et al., 2005).

Researchers recognize that programmatic elements can influence FQOL outcomes. These factors include disability-related support and services (Bhopti, Brown, & Lentin, 2016). In providing these services to families of *children with disabilities* (CWD), families experience higher FQOL satisfaction (Bhopti et al., 2016; Kyzar, Turnbull, Summers, & Gómez, 2012). As such, to establish the impact of family-focused supports and services, programme evaluators should include measures of FQOL to establish intervention effectiveness (Bhopti et al., 2016; Summers et al., 2007).

Unlike the individual QOL field with over 1,200 measures assessing QOL (Hughes, Hwang, Kim, Eisenman, & Killian, 1995), there are only two quantitative disability-related FQOL measures identified in published studies (Bhopti et al., 2016; Wang & Brown, 2009). An international team of researchers developed the Family Quality of Life Survey-2006 (FQoLS-2006) (Brown et al., 2006). Taking a life-course perspective, the FQoLS-2006 allows for use with older individuals with disabilities and their families (Bhopti et al., 2016). The FQoLS-2006 is a comprehensive assessment and takes about 45 min to administer (Bhopti et al., 2016). The Beach Center Family Quality of Life measure (Hoffman, Marquis, Poston, Summers, & Turnbull, 2006) is the second published FQOL instrument, created in a U.S. context (Samuel et al., 2012). The Beach Center inventory was developed for use with younger families and takes a minimum of 20 min to complete (Bhopti et al., 2016). Since this pilot study

considers families who have young children, and due to time constraints with data collection, this research team elected to use the Beach Center survey.

Researchers developed the Beach Center measure through several steps to define and measure FQOL. Poston et al. (2003) completed focus groups and individual interviews with caregivers of children with disabilities, caregivers of children without disabilities, service providers, administrators and researchers who work with families and individuals with disabilities. They identified ten FQOL domains. Park et al. (2003) then created a pilot survey which they field-tested with parents of children with disabilities, evaluating the relationships of items to factor structures through a series of exploratory factor analyses. From this, they reduced the initial 10-factor structure to a 5-factor model. Continuing the psychometric validation, Hoffman et al. (2006) revised the measure through a two-part process. In the first analysis, this team completed a confirmatory factor analysis establishing a 5-factor structure and evaluated internal consistency reliability for each of the five sub-scales. They then removed items with either low factor loadings or items that did not contribute to a coefficient alpha of at least 0.8. This process resulted in a 25-item measure that Hoffman and colleagues used for phase two of the study. In the second stage, they established that each of the five sub-scales was unidimensional, demonstrated convergent validity between several sub-scales of their measure and other existing measures, and developed support for test-retest reliability of the revised measure. The five sub-scales include Family Interaction, Parenting, Emotional Well-being, Physical Well-being and Disability-Related Support (Beach Center on Disability, 2015).

The Beach Center FQOL instrument was initially developed extensively based on U.S. samples, leading to questions about how effectively the measure transfers to other country contexts. Multiple researchers recommend additional validation when adopting measures to new cultural contexts. Zumbo (2009) pointed out that measures are not "context-free" (p. 76) and items refer to social aspects that may have different connotations in other cultures, thereby introducing item bias in measures. Marotta, Tramonte, and Willms (2015) maintained that item bias could contribute in several ways, including contextual differences and communication factors. With contextual differences, the meaning of items is dissimilar for cultural groups and may introduce another construct into the measure (Van de Vijver, 2000). With linguistic factors, translations could introduce several different elements. These include convoluted wording, making it more difficult to understand than the intent of the original text (Marotta et al., 2015), not fully covering the content, or reflecting a different threshold of a construct (i.e. anxiety versus worry) (Petersen et al., 2003).

To consider these concerns, multiple researchers have examined the Beach Center's survey to determine if it is applicable to countries other than the United States. Balcells-Balcells, Giné, Guàrdia-Olmos, and Summers (2011) evaluated the factor structure of a Spanish translation of the Beach FQOL survey, finding that the factor structure in the Spanish version was the same as the English form. Parpa et al. (2016) conducted a psychometric evaluation of a Greek-translated

version of the Beach Center FQOL. They reported a consistent factor structure as the original measure as well as favourable evidence related to internal consistency reliability, construct validity, and test-retest reliability. These two international psychometric evaluations took place in the context of high-income countries and did not seek to evaluate direct differences between U.S. and international samples.

Einfeld et al. (2012) advised completing validation studies of survey instruments to determine if measures developed in the context of high-income countries related to low- and middle-income country contexts. Researchers continued to evaluate the use of the Beach Center FQOL measure in low- and middle-income country contexts. Investigators have applied this measure to samples in China, a middle-income country, and the Democratic Republic of Congo, a low-income country. Studying in the Democratic Republic of Congo, Aldersey, Francis, Haines, and Chiu (2017) conducted focus group sessions with parents of CWD to determine how families described FQOL and whether items on the Beach Center questionnaire were linguistically and culturally consistent with this setting. They found that only one question, related to dental care, was not helpful in the Democratic Republic of Congo, as many of the participants themselves had never had a dental appointment. Participants identified that the remaining items meaningfully captured the FQOL construct. Likewise, Hu et al. (2011) established that this dental care question did not apply to a Chinese cohort and again found that the other items represented FQOL well. In sum, both Aldersey et al. (2017) and Hu et al. (2011) used qualitative methods to evaluate whether families of CWD define FQOL with the same concepts as the Beach Center FQOL measure in these two low- and middle-income country contexts. They concluded that the FQOL survey represented the experiences of these families. To date, no researchers have employed quantitative methods to examine the Beach Center FQOL in low- and middle-income country contexts.

While qualitative methods allow researchers one approach to assess item performance, quantitative methods can also provide information about item functioning. Bjorner, Kreiner, Ware, Damsgaard, and Bech (1998) noted that statistical methods lead to more sensitive appraisals of item functioning in cross-cultural evaluations. Nápoles-Springer and Stewart (2002) recommended the use of differential item functioning, to quantitatively establish measurement equivalence and identify potential items that may contain bias among cultural groups.

Differential item functioning (DIF) grew out of educational testing beginning in the 1960s as a way to assess if items were unfair for one group after controlling for the student's overall knowledge about a topic. Various researchers have also used this method to establish item-level equivalence to evaluate the cross-cultural application to diverse health outcomes. Examples include measures of cognitive functioning (He & Wolfe, 2010; Ramirez, Teresi, Holmes, Gurland, & Lantigua, 2006); depression (Teresi, Ramirez, Lai, & Silvers, 2008; Zhong, Gelaye, Fann, Sanchez, & Williams, 2014); health-related quality of life (Bjorner et al., 1998; Liang et al., 2009; Scott et al., 2006, 2009; Teresi et al., 2008); and

general health (Teresi et al., 2008). However, Scott et al. (2006) noted that DIF analyses of cross-cultural applications of quality of life measures are scarce. When introducing a measure to a new cultural context, items may not function as intended for the new group, thereby introducing systematic error or construct irrelevance into the data (Frey, 2018).

Researchers have developed multiple methods to detect the two types of DIF: uniform DIF and non-uniform. In uniform DIF, one group consistently outperforms another group when compared with those at a similar level of the trait. With non-uniform DIF, one group outperforms another group at the lower levels of the trait; then, the relationship reverses with the previously higher performing group moving to a lower level performance at the higher ends of the trait. Three general methods to detect DIF exist. The Mantel-Haenszel chi-squared test is a non-parametric method that works well with small samples, but only detects uniform DIF (Holland & Thayer, 1988). Logistic regression models allow for the detection of uniform and non-uniform DIF, while also being intuitive and easy to use (Scott et al., 2009; Swaminathan & Rogers, 1990; Zumbo, 1999). Finally, Item Response Theory models permit more accurate estimates of item parameters but require large sample sizes (Lord, 1980; Raju, 1990; Thissen, Steinberg, & Wainer, 1993).

In building additional validity evidence for the Beach Center FQOL measure, the study considers two groups, one from the large urban centre of Lusaka, Zambia, a low-income country, and a U.S. sample. Structural differences exist in the care and support of CWD in these countries, such as differences in health care and education. In a study that ranked the health performance of countries, the United States ranked 37th while Zambia was 182nd, indicating differences in health outcomes between these countries (Tandon, Murray, Lauer, & Evans, 2000). In Zambia, the WHO (2020) identified a critical need for trained healthcare professionals, resulting in gaps in health information and healthcare services for CWD (Central Statistical Office, 2018). The National Center for Educational Statistics (2019) reported that in 2017–2018, 14% of all public school students ages 3–21 received special education services in the United States, whereas in Zambia, 3.15% of all students received special education services in primary or secondary education settings (Ministry of General Education, 2017), leading the Zambian Central Statistical Office (2018) to conclude that CWD are unable to access educational services.

1.1 | Study aims

The present research study represents an important extension to the existing literature on the cross-cultural validity of this tool in low-income countries. The purpose of this study is to evaluate the linguistic and cultural application of three sub-scales of the Beach Center FQOL measure using differential item functioning analysis. This method allows researchers to consider if different response patterns exist for two groups—respondents in Zambia and a U.S. comparison group—for items when conditioned on similar levels of a trait.

2 | METHODS

2.1 | Instrument

This cross-sectional, secondary data analysis study used three sub-scales of the Beach Center Family Quality of Life measure, Emotional Well-being, Disability-Related Support and Physical/Material Well-being. Hoffman et al. (2006) noted that these three sub-scales are germane to programme effectiveness as they capture information about supports and services. Since this research team intends to use this measure for programme evaluation, the Family Interaction and Parenting sub-scales were not considered. As previously noted, Aldersey and colleagues and Hu et al. (2011) previously identified the Physical Well-being sub-scale item, *my family gets dental care when needed*, as problematic in a low- and middle-income contexts. This study excluded this item from analysis resulting in all domains with four items. To provide a satisfaction rating for each item, respondents used a 5-point rating from (1) *very dissatisfied* to (5) *very satisfied*. To arrive at a sub-scale total score for each of the domains, the research team summed the item-level responses for each item to arrive at a total sub-scale score. Examples of questions in the Emotional Well-being measure include, *my family members have friends or others who provide support*. The Physical Well-being sub-scale includes questions related to accessing transportation, medical care, having enough money to pay for expenses, and feeling safe in one's neighbourhood. The Disability-Related Support measure considers assistance in school and home as well as guidance to make friends and relationships with service providers.

Prior research demonstrated that the psychometric properties of this measure confirm good content validity (Hoffman et al., 2006; Park et al., 2003; Poston et al., 2003). Researchers also substantiated solid internal consistency reliability (Hoffman et al., 2006; Parpa et al., 2016) and established the factor structure (Hoffman et al., 2006). Last, Hoffman et al. (2006) revealed concurrent validity with the APGAR and the Family Resources Scale (Hoffman et al., 2006).

2.2 | Procedures

This study used data from a U.S. Midwestern sample ($N = 1,071$) and a Zambian sample ($N = 373$) of families who have young children with disabilities. The U.S. sample drew from families whose children qualified for education services and supports in Iowa and received pre-kindergarten through 6th-grade educational services. The majority of the children were in kindergarten to 3rd grade (65.4%). Participants acquired and returned their survey during the winter of 2019. Table 1 provides demographic information about the Iowa sample.

The Zambian sample consisted of caregivers of children with disabilities in two sub-districts of Lusaka, Zambia. Participants completed a programme evaluation questionnaire and test-piloted these

questions. These caregivers were involved in a community-based programme designed to provide disability-focused supports and services to these families. These supports and services seek to enhance access to community-based rehabilitation services and promote social supports and empower these families. The average age of the children with a disability for the Zambian sample was 11.3 years. Zambian participants were from two compounds in Lusaka, the capital city. These compounds have limited infrastructure and dense populations with residents who have limited financial assets. Table 2 includes demographic information for the Zambian sample, while Table 3 includes reliability coefficients for each of the two groups for the three sub-scales. The Iowa demographic table includes information about variability in race, ethnicity and geographic characteristics, while the Zambian sample was uniformly black, non-Hispanic, and all lived in Lusaka, a large urban area.

Trained Zambian staff collected the data orally. These staff members had received college degrees with exposure to programme monitoring and evaluation. The third author also provided additional training to the Zambian staff about the Beach Center measure. Even though English is the national language, most Zambians, particularly those who live in these low-income sub-districts, prefer to speak their local language. In Lusaka, multiple languages prevail, such as Nyanja and Bemba, as well as other languages. Due to the lower literacy rates among participants and the use of multiple languages, the research team decided to allow data collectors to translate items at the time of data collection rather than using a scripted translation. Data collectors received training from programme staff and were supervised until staff were confident that items were translated correctly and using a consistent process.

Participants were asked to complete the survey in the spring of 2019. The surveys were done orally, translated to the language of the participant at the time of the interview, and collected and managed using REDCap (Research Electronic Data Capture) electronic data capture mobile application hosted at a midwestern university (Harris et al., 2009, 2019). REDCap is a secure, web-based software platform developed to support data capture for research studies. It provides (a) an intuitive interface for validated data capture; (b) the capacity to track data manipulation and export; (c) automated export procedures for straightforward data downloads to statistical packages; and (d) methods for data integration and interoperability with external sources.

2.3 | Data analysis

This study used SPSS version 25 for the hierarchical ordinal regression analysis to complete the DIF evaluation. Two statistics led to decisions about the possible presence of DIF. These included the following: the change in chi-square to test for DIF significance, and the difference in Nagelkerke R^2 to indicate effect size due to elevated type I error rates or over-identification of items with DIF when only using a chi-square statistic (Jodoin & Gierl, 2001). The three hierarchical models are

TABLE 1 Iowa family member demographics (N = 1,071)

Characteristic	n	%
Gender		
Male	146	13.6
Female	768	71.7
Missing	157	14.7
Ethnicity		
Hispanic or Latino	32	3.0
Not Hispanic or Latino	872	81.4
Missing	157	14.7
Race		
White	865	80.8
Black or African American	9	0.8
American Indian or Alaska Native	1	0.1
Asian	12	1.1
Native Hawaiian or Pacific Islander	4	0.4
Other	14	1.3
Missing	166	15.5
Relationship to child		
Parent (biological, step, foster or adoptive)	902	84.2
Grandparent	9	0.8
Aunt or uncle	3	0.3
Sibling (brother or sister)	2	0.2
Other non-relative	2	0.2
Missing	153	14.3
Marital status		
Married/living with a partner	760	71.0
Not married (widowed, divorced, separated, never married)	153	14.3
Missing	158	14.8
Geographic location		
Urban (population >50,000)	151	14.1
Suburban (population between 10,000 and 50,000)	305	28.5
Rural (population less than 10,000)	442	41.3
Missing	173	16.2
Highest educational level		
Primary school, but not high school diploma	16	1.5
Secondary school (high school diploma)	205	19.1
Associates, bachelor's or graduate degree	648	60.5
Other	45	4.2
Missing	157	14.7

Model #1 (no DIF): $Y = b_0 + b_1 \text{FQOL sub-scale total}$

Model #2 (uniform DIF): $Y = b_0 + b_1 \text{FQOL sub-scale total} + b_2 \text{Group}$

Model #3 (non-uniform DIF): $Y = b_0 + b_1 \text{FQOL sub-scale total} + b_2 \text{Group} + b_3 \text{FQOL sub-scale total} \times \text{Group}$

To evaluate whether DIF is present, researchers recommend evaluating the difference between the chi-square with three

TABLE 2 Zambian family member demographics (N = 373)

Characteristic	n	%
Gender		
Male	30	8.0
Female	336	90.1
Relationship to child		
Parent (biological, step, foster, or adoptive)	273	73.2
Grandparent	55	14.7
Other relative (aunt, uncle, sibling, etc.)	44	15.0
Marital status		
Married/living with a partner	253	67.8
Not married (widowed, divorced, separated, never married)	120	32.2
Highest educational level		
Some primary	78	20.9
Primary	112	30.0
Some secondary	115	30.8
Secondary	36	9.7
Higher education	7	1.9
Other	25	6.7

degrees of freedom from model 3 and the chi-square statistic with one degree of freedom found in model 1 (Slocum, Gelin, & Zumbo, 2004; Zumbo, 1999). Due to concerns with multiple hypotheses testing with DIF analyses, this study adjusted α using Bonferroni-correction to arrive at α of .0125 (Crane, Gibbons, Jolley, & van Belle, 2006). Since the chi-square statistic can become inflated with larger sample sizes, Zumbo (1999) recommended also considering the difference in Nagelkerke R^2 between models 3 and 1 to determine the effect size. This study used the Jodoin and Gierl (2001) effect size criteria to determine the magnitude of DIF. Jodoin-Gierl considered a change in Nagelkerke R^2 of <.035 to be negligible while a change in $R^2 > .070$ was large, whereas differences in R^2 between the negligible and large effect were considered moderate. If the differences between the model 3 and 1 chi-square statistic yield a statistically significant difference and the difference in Nagelkerke R^2 is more than 0.035, then researchers should examine whether DIF was present. To be identified with DIF, items needed to have both a significant chi-square statistic and an $R^2 \Delta$ of more than 0.035.

To evaluate the type of DIF—either uniform or non-uniform—Zumbo (1999) endorsed considering the change in chi-square and magnitude for two models. With uniform DIF, researchers should examine the statistical difference between models 2 and 1, while differences in chi-squared between model 3 and model 2 provide information about non-uniform DIF. Additionally, researchers may use visual inspection of Loess lines from the two groups to determine the type of DIF. In uniform DIF, the Loess lines do not cross while with non-uniform DIF, the Loess lines for the two groups will cross (Slocum, Gelin, & Zumbo, 2003).

2.4 | Ethical considerations

As an analysis of secondary data, this study received ethical approval from the lead author's institution as an exempt research project, using only de-identified data. A second midwestern university Institutional Review Board approved the Zambian primary project data collection, and the Iowa secondary data included authorization to share de-identified secondary data.

3 | RESULTS

3.1 | Characteristics of samples

Participants for this study were two groups: one from Iowa and the second from two low-income sub-districts in Lusaka, Zambia. These samples had some similarities and differences. For example, both samples primarily were mothers of *children with disabilities* (CWD) who lived with an adult partner. However, whereas the Zambian sample was exclusively from the urban area of Lusaka, the Iowa sample included urban, suburban and rural areas. A second point of divergence was the educational background of participants. The Iowa respondents had higher levels of education, with 60.5% possessing

at least an associate degree, while 81.7% of the Zambian sample had not received a secondary level of education. Differences were also present for these children with all of the children in the U.S. sample accessing educational services, while only 33.6% of the Zambian children attended school.

3.2 | Differential item functioning

Table 4 provides information about the $\chi^2_{\text{Model 3-Model 1}}$ as well as $R^2 \Delta$ for each of the three sub-scales. The next section considers the item-level DIF analysis for each of the three sub-scales.

3.2.1 | Emotional Well-being

For the DIF analysis, three of the four items showed negligible DIF with a change in effect size between model 3 and model 1 at less than 0.035. One item, *my family members have time to pursue their own interests*, was not statistically significant for DIF $\chi^2_{\text{model 3-model 1}}$ (2, $N = 1,286$) = 4.54, $p = .103$.

3.2.2 | Physical Well-being

Three of the four items showed negligible DIF while one item, *my family has a way to take care of our expenses*, had a statistically significant $\chi^2_{\text{model 3-model 1}}$ (2, $N = 1,281$) = 149.94, $p < .001$ and meaningful DIF ($R^2 = \Delta.039$). Figure 1 provides a graphic representation of this item. In the graph, the area between the curves of the two groups is most important—with a more substantial area indicating a higher likelihood of DIF. The Midwestern group showed a greater probability of reporting higher satisfaction when compared with Zambians at a similar lower level of physical well-being. While at the upper levels of overall physical well-being, the Zambian sample tended to

TABLE 3 Coefficient alpha for the Iowa sample and Zambian sample for each of the sub-scales

Sub-scale	Location	<i>n</i>	α
Emotional Well-being	Iowa	917	0.817
	Zambia	369	0.602
Physical Well-being	Iowa	913	0.729
	Zambia	368	0.651
Disability-Related Support	Iowa	143	0.846
	Zambia	361	0.874

TABLE 4 Chi-squared statistic of the difference between Model 3 and Model 1, p -value, change in Nagelkerke R squared and effect size interpretation

Scale	Item	$\chi^2 \Delta$	p -value	$R^2 \Delta$	Effect size
Emotional Well-being	Relieve stress	89.13	<.001	.029	Negligible
	Friends or others who provide support	26.13	<.001	.006	Negligible
	Time to pursue their own interests	4.54	.103	.001	
	Outside help available	18.92	<.001	.005	Negligible
Physical Well-being	Transportation	67.203	<.001	.017	Negligible
	Medical care	60.538	<.001	.016	Negligible
	Expenses	149.945	<.001	.039	Moderate
	Safe	52.935	<.001	.016	Negligible
Disability-Related Support	Progress at school	21.53	<.001	.011	Negligible
	Progress at home	3.91	.14	.002	
	Support to make friends	10.03	<.001	.005	Negligible
	Relationship with service providers	19.992	<.001	.013	Negligible

indicate greater satisfaction with having enough money to cover expenses—noted with the crossing Loess lines.

3.2.3 | Disability-Related Support

One item in this measure did not have any indication of DIF with a $\chi^2_{\text{model 3-model 1}}(2, N = 504) = 3.91, p = .14$, while three other items showed negligible DIF. Figure 2 considers the item, *my child with a disability has support to make progress at home*, which was not statistically significant for DIF. The area between the two lines that represent the two groups is minimal, indicating that the two groups responded similarly to this item at all levels of the trait.

4 | DISCUSSION

This study explored the potential cross-cultural difference in the Beach Center Family Quality of Life survey between caregivers CWDs in the U.S. state of Iowa compared to two low-income and low-resource sub-districts in Lusaka, Zambia. Specifically, it sought to examine whether items of three subsections of the instrument functioned similarly between these two groups when controlling for total sub-scale scores. Only one item—*my family has a way to take care of our expenses*—did not function similarly. In this item, participants from Iowa were more likely to provide a higher level of satisfaction at low levels of Physical Well-being than the Zambian sample, while at the upper levels of Physical Well-being, the Zambian sample was more likely to report higher satisfaction. When analyses identify items with potential DIF, it means the item is capturing a factor unrelated to the family quality of life sub-scale related to Physical Well-being and introducing systematic error into the data (Frey, 2018).

Scott et al. (2006) identify that interpreting this significant and meaningful DIF is often equivocal. A myriad of reasons could contribute to DIF, such as format, difficulty of words and differences

in cultural applications (Teresi et al., 2008). With this analysis, it appears that the context of low-income and low-resource communities could have led to the identification of DIF in the item related to having enough money to purchase needed items. Zambian participants in this study were from two lower income communities and likely had few financial resources when compared with the Midwestern U.S. participants. The item related to expenses may reflect exposure to funding rather than physical well-being. As was previously explained, Aldersey et al. (2017) and Hu et al. (2011) excluded the dental item from this sub-scale, noting that many people cannot acquire dental care due to the limited resources in the Democratic Republic of Congo or parts of China. These two questions, access to dental care and expenses, both drawn from the Physical Well-being sub-scale, appear to capture access. This content relates to structural concerns of a low-income and low-resource contexts rather than aspects that caregivers can readily influence or change.

The use of a two-step process to identify items with potential bias led to fewer items identified with possible DIF. If this study only used the difference in chi-square from model 3 to model 1, all but two items would exhibit DIF. In using both statistical significance and an effect size measure, this study likely reduced type I error rates.

If directly comparing the FQOL scores between Zambian and U.S. participants, it would be necessary to account for this measurement difference identified in the item related to expenses through statistical methods (Desjardins & Bulut, 2018). However, there is still value in seeking this information from Zambian participants. For example, Zambia recently implemented a short-term Social Cash Transfer programme to support families who are providing care for a person with a disability (Arruda & Dubois, 2018). Additionally, non-governmental organizations are implementing micro-finance strategies as a way to buffer families of people with disabilities from poverty. Elimination of this item as an outcome measure would not allow the opportunity to assess these policies and programme changes adequately. These programmes might explain the non-uniform DIF identified in the question related to having money to afford expenses. It is

FIGURE 1 Item with differential item functioning

FIGURE 2 Item not displaying evidence of differential item functioning

possible that the few Zambian families who either receive the Social Cash Transfer programme or micro-finance support are those who had overall high levels of Physical Well-being and in turn reported greater financial ability to pay for expenses when compared with individuals at the same level of Physical Well-being in the Iowa sample.

For applied research, a coefficient alpha of .7 or above is considered acceptable. However, two of the Zambian sub-scales—Emotional and Physical Well-being had coefficient alphas of <.7. The Emotional Well-being scale revealed $\alpha = .602$, which indicates that almost 40% of score variance is due to random error. Deleting any item in the Emotional Well-being scale did not improve the reliability index for this measure. The Physical Well-being scale had a coefficient alpha of .651, indicating that 65% of score variance is actual score variance. Removing the DIF item related to expenses did not improve the reliability coefficient for the Physical Well-being measure. The Disability-Related Support sub-scale, however, had a reliability index above 0.8. A plausible explanation is that the complexity of the items related to the Emotional and Physical Well-being concepts and translating the items at the time of data collection may have led to an increase in random error. Due to the lower educational attainment of the Zambian sample, it was necessary to verbally ask participants the questions rather than have participants read the questions. The lack of consistent language could have led to misunderstanding the concepts. In the future, researchers could use standardized translations in the various local languages and pilot these translations on a few local participants using a think-aloud procedure to assure that the participants understand the translated items with the same meaning as the original English questions (Jääskeläinen, 2010).

This study includes several limitations. First, with the Disability-Related Support sub-scale, the Iowa sample received few supports and did not respond to these questions. Scott et al. (2009) noted that with scales longer than three questions, groups should consist of 300 participants to provide adequate power to detect DIF. The Iowa sample, however, only had 143 responses. This means that it would be challenging to detect DIF with this sample size. Second, the trained data collectors translated the items into the participants' preferred language at the time of data collection, and it is possible that this process led to lower reliability coefficients for the Zambian sample. These low reliability indices will make discerning any difference over time difficult. Third, this study only considers validity evidence for three of the five Beach Center sub-scales, excluding the Family Interaction and Parenting sub-scales. Last, the two samples were not from representative samples from either the United States or Zambia. The Iowa comparison group may have different experiences with family quality of life due to programmes or supports available exclusively through this state. In contrast, the Zambian sample draws from the large urban capital city of Lusaka and specifically two low-income sub-districts. Caregivers' experience with family quality of life may likely be different in rural areas due to the more limited health and social supports.

Further study with a larger sample that is more representative of contexts both in the United States and Zambia would be helpful. In

using a larger sample, it would then be possible to complete a statistical analysis using item response theory, which is a more sensitive method of detecting DIF (Desjardins & Bulut, 2018). Additionally, a larger Zambian sample would also allow for additional psychometric evaluations, such as assessing the factor structure of the Zambian items.

This study extends the existing literature of the Beach Center Family Quality of Life survey in two important ways. First, while other researchers have examined the cultural and linguistic differences of the Beach Center Family Quality of Life survey using qualitative methods (Aldersey et al., 2017; Hu et al., 2011), this is the first study to use a quantitative method (i.e. DIF analysis). Second, as this instrument was developed exclusively in the context of the United States, this study specifically examined potential socio-economic differences between participants in low versus high-income countries. Researchers have, in previous studies, examined the validity of using this instrument in other cultural and linguistic contexts. They have generally found that this instrument has favourable psychometric qualities in this new context. This study also supports this finding but noted the possible introduction of construct irrelevant information in one question that evaluates a family's financial ability to provide support. In the context of a low- or middle-income countries, this question may capture the financial strain that all individuals experience when living in poverty.

In conclusion, DIF analysis was an effective and efficient means of assessing the cross-cultural validity of the Beach Center Family Quality of Life survey. The three sub-scales maintain promising psychometric properties and are appropriate to use in low-income communities in Lusaka, Zambia, as an evaluation tool of programme implementation activities. Assuring validity across context is an important step for all researchers, and practitioners who want to support improved outcomes for families of children with disabilities.

ACKNOWLEDGMENTS

The authors wish to thank Dr. Kathleen Kyzar for her generous sharing of the Midwestern data set used as the comparison group in this study. The authors recognize that this research would not be possible without collaboration with our in-country partner, Catholic Medical Mission Board—Zambia. We appreciate all of your expertise and wisdom that you have shared.

CONFLICT OF INTEREST

The authors do not have any conflicts of interest associated with this research.

ORCID

Renee A. Hepperlen <https://orcid.org/0000-0002-8833-275X>

Paula Rabaey <https://orcid.org/0000-0002-1476-7793>

Mary O. Hearst <https://orcid.org/0000-0002-9068-827X>

REFERENCES

Aldersey, H. M., Francis, G. L., Haines, S. J., & Chiu, C. Y. (2017). Family quality of life in the Democratic Republic of the Congo. *Journal of*

- Policy and Practice in Intellectual Disabilities*, 14(1), 78–86. <https://doi.org/10.1111/jppi.12189>
- Arruda, P., & Dubois, L. (2018). A brief history of Zambia's social cash transfer programme (No. 62). *International policy center for inclusive growth* (pp. 1–9). Brasília, Brazil: International Policy Centre for Inclusive Growth. https://www.ipcundp.org/pub/eng/PRB62_A_brief_history_of_Zambia_s_social_cash_transfer_programme.pdf
- Balcells-Balcells, A., Giné, C., Guàrdia-Olmos, J., & Summers, J. A. (2011). Family quality of life: Adaptation to Spanish population of several family support questionnaires. *Journal of Intellectual Disability Research*, 55(12), 1151–1163. <https://doi.org/10.1111/j.1365-2788.2010.01350.x>
- Beach Center on Disability (2015). *Beach center family quality of life scale: Psychometric characteristics and scoring key*. Retrieved from www.beachcenter.lsi.ku.edu
- Bhothi, A., Brown, T., & Lentin, P. (2016). Family quality of life: A key outcome in early childhood intervention services—A scoping review. *Journal of Early Intervention*, 38(4), 191–211. <https://doi.org/10.1177/1053815116673182>
- Bjorner, J. B., Kreiner, S., Ware, J. E., Damsgaard, M. T., & Bech, P. (1998). Differential item functioning in the Danish translation of the SF-36. *Journal of Clinical Epidemiology*, 51(11), 1189–1202. [https://doi.org/10.1016/S0895-4356\(98\)00111-5](https://doi.org/10.1016/S0895-4356(98)00111-5)
- Brown, I., Brown, R. I., Baum, N. T., Isaacs, B. J., Myerscough, T., Neikrug, S., ... Wang, M. (2006). *Family quality of life survey: Main caregivers of people with intellectual disabilities*. Toronto, ON: Surrey Place Centre.
- Central Statistical Office (2018). *Zambia national disability survey 2015*. Retrieved from <https://www.unicef.org/zambia/media/1141/file/Zambia-disability-survey-2015.pdf>
- Claes, C., Van Hove, G., van Loon, J., Vandeveld, S., & Schalock, R. L. (2010). Quality of life measurement in the field of intellectual disabilities: Eight principles for assessing quality of life-related personal outcomes. *Social Indicators Research*, 98(1), 61–72. <https://doi.org/10.1007/s11205-009-9517-7>
- Crane, P. K., Gibbons, L. E., Jolley, L., & van Belle, G. (2006). Differential item functioning analysis with ordinal logistic regression techniques: DIFdetect and difwithpar. *Medical Care*, 44, S115–S123. <https://doi.org/10.1097/01.mlr.0000245183.28384.ed>
- Cummins, R. A. (2005). Moving from the QOL concept to a theory. *Journal of Intellectual Disability Research*, 49(10), 699–770. <https://doi.org/10.1111/j.1365-2788.2005.00738.x>
- Desjardins, C. D., & Bulut, O. (2018). *Handbook of educational measurement and psychometrics using R*. Boca Raton, FL: Taylor & Francis Group.
- Einfeld, S. L., Stancliffe, R. J., Gray, K. M., Sofronoff, K., Rice, L., Emerson, E., & Yasamy, M. T. (2012). Interventions provided by parents for children with intellectual disabilities in low- and middle-income countries. *Journal of Applied Research in Intellectual Disabilities*, 25(2), 135–142. <https://doi.org/10.1111/j.1468-3148.2011.00678.x>
- Frey, B. (2018). *The SAGE encyclopedia of educational research, measurement, and evaluation* (Vol. 1–4). Thousand Oaks, CA: SAGE Publications, Inc.
- Harris, P. A., Taylor, R., Minor, B. L., Elliott, V., Fernandez, M., O'Neal, L., ... Duda, S. N. (2019). The REDCap consortium: Building an international community of software platform partners. *Journal of Biomedical Informatics*, 95, 103208. <https://doi.org/10.1016/j.jbi.2019.103208>
- Harris, P. A., Taylor, R., Thielke, R., Payne, J., Gonzalez, N., & Conde, J. G. (2009). Research electronic data capture (REDCap)—a metadata-driven methodology and workflow process for providing translational research informatics support. *Journal of Biomedical Informatics*, 42(2), 377–381. <https://doi.org/10.1016/j.jbi.2008.08.010>
- He, W., & Wolfe, E. W. (2010). Item equivalence in English and Chinese translation of a cognitive development test for preschoolers. *International Journal of Testing*, 10(1), 80–94. <https://doi.org/10.1080/15305050903534738>
- Hoffman, L., Marquis, J., Poston, D., Summers, J. A., & Turnbull, A. (2006). Assessing family outcomes: Psychometric evaluation of the beach center family quality of life scale. *Journal of Marriage and Family*, 68(4), 1069–1083. <https://doi.org/10.1111/j.1741-3737.2006.00314.x>
- Holland, P. W., & Thayer, D. T. (1988). Differential item performance and the Mantel-Haenszel procedure. In H. Wainer, & H. I. Braun (Eds.), *Test validity* (pp. 129–145). Hillsdale, NJ: Lawrence Erlbaum Associates Publishers.
- Hu, X., Summers, J. A., Turnbull, A., & Zuna, N. (2011). The quantitative measurement of family quality of life: A review of available instruments. *Journal of Intellectual Disability Research*, 55(12), 1098–1114. <https://doi.org/10.1111/j.1365-2788.2011.01463.x>
- Hughes, C., Hwang, B., Kim, J. H., Eisenman, L. T., & Killian, D. J. (1995). Quality of life in applied research: A review and analysis of empirical measures. *American Journal on Mental Retardation*, 99, 623–641.
- Isaacs, B. J., Brown, I., Brown, R. I., Baum, N. T., Myer-Schough, T., Neikrug, S., & Wang, M. (2007). The international family quality of life project: Goals and description of a survey tool. *Journal of Policy and Practice in Intellectual Disabilities*, 4(3), 177–185. <https://doi.org/10.1111/j.1365-2788.2005.00751.x>
- Jääskeläinen, R. (2010). *Think-aloud Protocol. Handbook of Translation Studies*, 1, 371–374.
- Jodoin, M. G., & Gierl, M. J. (2001). Evaluating type I error and power rates using an effect size measure with the logistic regression procedure for DIF detection. *Applied Measurement in Education*, 14(4), 329–349. https://doi.org/10.1207/s15324818AME1404_2
- Kyzer, K. B., Turnbull, A. P., Summers, J. A., & Gómez, V. A. (2012). The relationship of family support to family outcomes: A synthesis of key findings from research on severe disability. *Research and Practice for Persons with Severe Disabilities*, 37(1), 31–44. <https://doi.org/10.2511/027494812800903247>
- Liang, W. M., Chang, C. H., Yeh, Y. C., Shy, H. Y., Chen, H. W., & Lin, M. R. (2009). Psychometric evaluation of the WHOQOL-BREF in community-dwelling older people in Taiwan using Rasch analysis. *Quality of Life Research*, 18(5), 605–618. <https://doi.org/10.1007/s11136-009-9471-5>
- Lord, F. M. (1980). *Applications of item response theory to practical testing problems*. Hillsdale NJ: Erlbaum.
- Lyons, G. (2010). Quality of life for persons with intellectual disabilities: A review of the literature. In Ralph Kober (Ed), *Enhancing the quality of life of people with intellectual disabilities* (pp. 73–126). Dordrecht, The Netherlands: Springer.
- Marotta, L., Tramonte, L., & Willms, J. D. (2015). Equivalence of testing instruments in Canada: Studying item bias in a cross-cultural assessment for preschoolers. *Canadian Journal of Education*, 38(3).
- Maulik, P. K., Mascarenhas, M. N., Mathers, C. D., Dua, T., & Saxena, S. (2011). Prevalence of intellectual disability: A meta-analysis of population-based studies. *Research in Developmental Disabilities*, 32(2), 419–436. <https://doi.org/10.1016/j.ridd.2010.12>
- Miley, K. K., O'Melia, M. O., & DuBoise, B. (2016). *Generalist social work practice: An empowering approach* (8th ed.). Boston, MA: Allyn & Bacon.
- Ministry of General Education (2017). *Zambia educational statistical bulletin 2017*. Retrieved from https://www.moge.gov.zm/download/statistics/annual_statistics/Educational-Statistical-Bulletin/EDUCATIONAL-STATISTICAL-BULLETIN-2017.PDF
- Nápoles-Springer, A. M., & Stewart, A. L. (2002). Use of health-related quality of life measures in older and ethnically diverse U.S. populations. In J. H. Skinner & D. Holmes (Eds.), *Multicultural measurement in older populations*. (pp. 173–180). New York, NY: Springer Publishing Company.
- National Center for Educational Statistics (2019). *Children and youth with disabilities*. Retrieved from https://nces.ed.gov/programs/coe/indicator_cgg.asp

- Park, J., Hoffman, L., Marquis, J., Turnbull, A. P., Poston, D., Mannan, H., ... Nelson, L. L. (2003). Toward assessing family outcomes of service delivery: Validation of a family quality of life survey. *Journal of Intellectual Disability Research*, 47(4–5), 367–384. <https://doi.org/10.1046/j.1365-2788.2003.00497.x>
- Parpa, E., Katsantonis, N., Tsilika, E., Galanos, A., Sassari, M., & Mystakidou, K. (2016). Psychometric properties of the family quality of life scale in Greek families with intellectual disabilities. *Journal of Developmental and Physical Disabilities*, 28(3), 393–405. <https://doi.org/10.1007/s10882-016-9477-1>
- Petersen, M. A., Groenvold, M., Bjorner, J. B., Aaronson, N., Conroy, T., Cull, A., ... Sullivan, M. (2003). Use of differential item functioning analysis to assess the equivalence of translations of a questionnaire. *Quality of Life Research*, 12(4), 373–385.
- Poston, D., Turnbull, A., Park, J., Mannan, H., Marquis, J., & Wang, M. (2003). Family quality of life: A qualitative inquiry. *Mental Retardation*, 41(5), 313–328. [https://doi.org/10.1352/0047-6765\(2003\)41<313:FQOLAQ>2.0.CO;2](https://doi.org/10.1352/0047-6765(2003)41<313:FQOLAQ>2.0.CO;2)
- Raju, N. S. (1990). Determining the significance of estimated signed and unsigned areas between two item response functions. *Applied Psychological Measurement*, 14, 197–207. <https://doi.org/10.1177/014662169001400208>
- Ramirez, M., Teresi, J. A., Holmes, D., Gurland, B., & Lantigua, R. (2006). Differential item functioning (DIF) and the Mini-Mental State Examination (MMSE): Overview, sample, and issues of translation. *Medical Care*, 44, S95–S106.
- Rapley, M. (2003). *Quality of life research: A critical introduction*. Thousand Oaks, CA: Sage.
- Samuel, P. S., Rillotta, F., & Brown, I. (2012). The development of family quality of life concepts and measures. *Journal of Intellectual Disability Research*, 56(1), 1–16. <https://doi.org/10.1111/j.1365-2788.2011.01486.x>
- Schalock, R. L. (2004). The concept of quality of life: What we know and do not know. *Journal of Intellectual Disability Research*, 48(3), 203–216. <https://doi.org/10.1111/j.1365-2788.2003.00558.x>
- Schalock, R. L., Brown, I., Brown, R., Cummins, R. A., Felce, D., Matikka, L., ... Parmenter, T. (2002). Conceptualization, measurement, and application of quality of life for persons with intellectual disabilities: Report of an international panel of experts. *Mental Retardation*, 40(6), 457–470. [https://doi.org/10.1352/00476765\(2002\)040<0457:CMAAOQ>2.0.CO;2](https://doi.org/10.1352/00476765(2002)040<0457:CMAAOQ>2.0.CO;2)
- Scior, K. (2016). Toward understanding intellectual disability stigma: Introduction. In K. Scior & S. Werner (Eds.), *Intellectual disability and stigma: Stepping out from the margins* (pp. 3–14). London, UK: Palgrave MacMillan.
- Scott, N. W., Fayers, P. M., Aaronson, N. K., Bottomley, A., de Graeff, A., Groenvold, M., ... EORTC Quality of Life Group (2009). A simulation study provided sample size guidance for differential item functioning (DIF) studies using short scales. *Journal of Clinical Epidemiology*, 62(3), 288–295. <https://doi.org/10.1016/j.cjlin.epi.2008.06.003>
- Scott, N. W., Fayers, P. M., Bottomley, A., Aaronson, N. K., de Graeff, A., Groenvold, M., ... Sprangers, M. A. G. (2006). Comparing translations of the EORTC QLQ-C30 using differential item functioning analyses. *Quality of Life Research*, 15(6), 1103–1115. <https://doi.org/10.1007/s11136-006-0040-x>
- Slocum, S. L., Gelin, M. N., & Zumbo, B. D. (2003). Statistical and graphical modeling to investigate differential item functioning for rating scale and Likert item formats. *Developments in the theories and applications of measurement, evaluation, and research methodology across the disciplines*, 1.
- Slocum, S. L., Gelin, M. N., & Zumbo, B. D. (2004). *Statistical and graphical modeling to investigate differential item functioning for rating scale and Likert item formats*. Vancouver, BC: Edgeworth Laboratory, University of British Columbia.
- Summers, J. A., Poston, D. J., Turnbull, A. P., Marquis, J., Hoffman, L., Mannan, H., & Wang, M. (2005). Conceptualizing and measuring family quality of life. *Journal of Intellectual Disability Research*, 49(10), 777–783. <https://doi.org/10.1111/j.1365-2788.2005.00751.x>
- Summers, J. A., Marquis, J., Mannan, H., Turnbull, A. P., Fleming, K., Poston, D. J., ..., Kupzyk, K. (2007). Relationship of perceived adequacy of services, family–professional partnerships, and family quality of life in early childhood service programmes. *International Journal of Disability, Development and Education*, 54(3), 319–338. <https://doi.org/10.1080/10349120701488848>
- Swaminathan, H., & Rogers, H. J. (1990). Detecting differential item functioning using logistic regression procedures. *Journal of Educational Measurement*, 27(4), 361–370. <https://doi.org/10.1111/j.1745-3984.1990.tb00754.x>
- Tandon, A., Murray, C. J., Lauer, J. A., & Evans, D. B. (2000). *Measuring overall health system performance for 191 countries*. Geneva, Switzerland: World Health Organization.
- Teresi, J. A., Ramirez, M., Lai, J. S., & Silver, S. (2008). Occurrences and sources of Differential Item Functioning (DIF) in patient-reported outcome measures: Description of DIF methods, and review of measures of depression, quality of life and general health. *Psychology Science Quarterly*, 50(4), 538.
- Thissen, D., Steinberg, L., & Wainer, H. (1993). Detection of differential item functioning using the parameters of item response models. In P. Holland & H. Wainer (Eds.), *Differential item functioning* (pp. 67–113). Hillsdale, NJ: Erlbaum.
- UNCRPD (2008). *United Nations Convention on the Rights of Persons with Disabilities*. New York, NY: UN General Assembly.
- Van de Vijver, F. (2000). The nature of bias. In R. H. Dana (Ed), *Handbook of cross-cultural and multicultural personality assessment* (pp. 105–124). Chicago: Routledge.
- Wang, M., & Brown, R. (2009). Family quality of life: A framework for policy and social service provisions to support families of children with disabilities. *Journal of Family Social Work*, 12(2), 144–167. <https://doi.org/10.1080/10522150902874842>
- World Health Organization (2020). *Zambia*. Retrieved from <https://www.who.int/workforcealliance/countries/zmb/en/>
- World Health Organization and World Bank (2011). *Summary world report on disability*. Geneva, Switzerland: WHO. Retrieved from http://disabilities/world_report/2011/accessible_en.pdf
- Zhong, Q., Gelaye, B., Fann, J. R., Sanchez, S. E., & Williams, M. A. (2014). Cross-cultural validity of the Spanish version of PHQ-9 among pregnant Peruvian women: A Rasch item response theory analysis. *Journal of Affective Disorders*, 158, 148–153. <https://doi.org/10.1016/j.jad.2014.02.012>
- Zumbo, B. D. (1999). *A handbook on the theory and methods of differential item functioning (DIF): Logistic regression modeling as a unitary framework for binary and Likert-type (ordinal) item scores*. Retrieved from <http://faculty.educ.ubc.ca/zumbo/DIF/handbook.pdf>
- Zumbo, B. D. (2009). *Validity as contextualized and pragmatic explanation, and its implications for validation practice*. In The Concept of Validity: Revisions, New Directions and Applications. IAP Information Age Publishing.
- Zuna, N., Summers, J. A., Turnbull, A. P., Hu, X., & Xu, S. (2010). Theorizing about family quality of life. In Ralph Kober (Ed), *Enhancing the quality of life of people with intellectual disabilities* (pp. 241–278). Dordrecht, The Netherlands: Springer.

How to cite this article: Rabaey P, Hearst MO, Hepperlen RA. Evaluating the cross-cultural validity of three family quality of life sub-scales. *J Appl Res Intellect Disabil*. 2020;00:1–10. <https://doi.org/10.1111/jar.12727>