

Available online at [www.sciencedirect.com](http://www.sciencedirect.com)

SciVerse ScienceDirect

journal homepage: [www.elsevier.com/locate/jval](http://www.elsevier.com/locate/jval)

## Unidimensional 12-Item Zarit Caregiver Burden Interview for the Assessment of Dementia Caregivers' Burden Obtained by Item Response Theory

Javier Ballesteros, MD, PhD<sup>1,2,\*</sup>, Borja Santos, BSc Maths<sup>2</sup>, Eduardo González-Fraile, MSc Psychol<sup>1,3</sup>, Paula Muñoz-Hermoso, BSc Psychol<sup>3,4</sup>, Ana I. Domínguez-Panchón, MSc Psychol<sup>3,5</sup>, Manuel Martín-Carrasco, MD, PhD<sup>2,3</sup>

<sup>1</sup>Department of Neuroscience – Psychiatry, Medical School, University of the Basque Country, UPV/EHU, Leioa, Spain; <sup>2</sup>CIBERSAM, Madrid, Spain; <sup>3</sup>Instituto de Investigaciones Psiquiátricas – IIP, Bilbao, Spain; <sup>4</sup>Clínica Psiquiátrica Padre Menni, Pamplona, Spain; <sup>5</sup>Hospital Aita-Menni, Arrasate-Mondragón, Spain

### ABSTRACT

**Objectives:** To develop a one-dimensional version of the 22-item Zarit Caregiver Burden Interview (ZBI) by applying item response theory approaches. **Methods:** The answers to the 22-item ZBI of 241 caregivers participating in a clinical trial were analyzed 1) with a Mokken nonparametric item response theory analysis to ascertain the dimensional structure underlying the scale and obtain a one-dimensional reduced version, 2) with the Samejima's graded response model to assess the item characteristics of the reduced version, and 3) with confirmatory factor analysis to confirm the unidimensionality of the reduced ZBI version and assess the item loadings to the burden latent variable. **Results:** Mokken analysis resulted in a major one-dimensional scale comprising 12 items directly related with burden. All items showed scalability indices over 0.30. The scalability for the overall scale was 0.44 defining a medium scale according to Mokken's criteria. An unconstrained Samejima's graded response model showed appropriate fit, and most items of the reduced 12-item ZBI

presented pertinent difficulty and discrimination parameters. The results of the 12-item ZBI confirmatory factor analysis fitted to a one-dimensional latent structure for burden (comparative fit index = 0.975; root-mean-square error of approximation = 0.067; weighted root mean square residual = 0.677). All factor loadings were above 0.40 with items 9 (strained by the relative) and 22 (overall feeling of burden) presenting the highest loadings. **Conclusions:** The reduced 12-item ZBI fits a one-dimensional latent variable of burden. Further psychometric studies, focusing on its equivalence for different populations, sensitivity to change, and minimal important difference are warranted.

**Keywords:** confirmatory factor analysis, item response theory, Mokken analysis, Samejima's graded response model, Zarit Caregiver Burden Interview.

Copyright © 2012, International Society for Pharmacoeconomics and Outcomes Research (ISPOR). Published by Elsevier Inc.

### Introduction

Originally, the Zarit Caregiver Burden Interview (ZBI) was developed in 1980 as a 29-item self-report scale aimed to assess the subjective burden experienced by an informal (not paid) caregiver, usually a relative [1]. Some years later, shorter versions of 20 and 22 items were released [2,3], and in 1991, the standard version of 22 items with two factorial subscales—Personal Strain and Role Strain—was produced [4]. While in the first version items were scored on a four-point ordinal Likert-type scale, a five-point ordinal scale (0: never; 1: rarely; 2: sometimes; 3: quite frequently; and 4: nearly always) recording the feeling the caregiver has on the corresponding statement is currently used. Despite the two factorial subscales mentioned above, the ZBI has been almost universally used as a one-dimensional measure given the high correlation between factors originally described. Later on, several attempts to find latent dimensions of the ZBI

have been successfully made [5–7]. It has been used primarily, but not exclusively, among caregivers of patients diagnosed with dementia [8–10], and caregivers of patients presenting with other pathologies or settings including palliative care, heart failure, brain injury, or schizophrenia [11–14] have also been assessed. Nowadays, the ZBI is believed to be the most commonly used measure of caregivers' burden [15].

Because of its multidimensionality, the ZBI total score discloses several underlying latent constructs with two to five factors as previously reported [7,16]. The multidimensional structure of the ZBI implies that a clinical interpretation based on its total score could not be as informative as it should be because of mixing different latent constructs in a unique observed score. Consequently, if the 22-item ZBI is used as primary outcome in trials designed to evaluate the efficacy of interventions to improve caregiver burden, and an absolute change since baseline is reported for its total score, it could be

\* Address Correspondence to: Javier Ballesteros, Department of Neuroscience – Psychiatry, Medical School, University of the Basque Country, UPV/EHU, Barrio Sarriena S/N, E-48640 Leioa, Spain.

E-mail: [javier.ballesteros@ehu.es](mailto:javier.ballesteros@ehu.es).

1098-3015 - see front matter Copyright © 2012, International Society for Pharmacoeconomics and Outcomes Research (ISPOR). Published by Elsevier Inc.

<http://dx.doi.org/10.1016/j.jval.2012.07.005>

unclear whether some dimensions have changed differentially over the intervention process, and this is the information that could be relevant to properly address the trial aims.

Factor analysis studies of the ZBI have been mostly done to verify or explore the underlying structure of the original ZBI. Nonetheless, others were intended to develop briefer or screening versions of the ZBI, and several reduced versions of the canonical 22-item ZBI with a variable number of items have been reported in the literature [17–19]. None of those versions, however, were developed or assessed by using psychometric approaches rooted on item response theory (IRT). IRT is less concerned than classical test theory with reliability of total scores. On the contrary, it is more concerned with the analysis of the responses elicited to individual items of the evaluated scale. Benefits of IRT include comprehensive analysis and reduction of measurement error, meaningful scaling of latent variables, objective calibration and equating, evaluation of test and item bias, greater accuracy in the assessment of change due to therapeutic interventions, and evaluation of model and person fit [20]. IRT models use item endorsement frequencies as outcomes to estimate parameters that characterize the properties of an item and are increasingly used to improve the accuracy of classical psychometric tools [21,22] or to develop shortened versions [23]. Our aim in this study was to obtain a reduced and one-dimensional version of the 22-item ZBI by using both nonparametric and parametric IRT analyses of the baseline measurements of caregivers of dementia patients recruited for a randomized clinical trial designed to assess the efficacy of a psychoeducational program on the caregivers' burden (EDUCA-2 trial; ISRCTN14411440).

## Methods

### Study Design and Population

This study includes a validation sample of 241 caregivers of patients with dementia recruited within a multicenter randomized clinical trial (20 centers across Spain and Portugal). To be included, a caregiver should be informally caring (not paid for) for a patient with dementia (*Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, Text Revision*, criteria). The patient should be treated as an outpatient in memory clinics or psychogeriatric day centers at the research sites, to present impairment of at least two instrumental activities, or one activity of daily life. The caregiver (adult male or female) should care for the patient at least 4 hours daily. The recruitment for the trial began in September 2010 and lasted till October 2010.

### Data Collection Procedures and Measures

The EDUCA-2 trial included three evaluations: at baseline (visit 1), after finishing the trial intervention (visit 2 at 4 months since inception), and finally at 4 months since finishing the trial intervention (visit 3 at 8 months since inception). This article reports results of the 22-item ZBI as obtained at trial inception (prerandomization period). The classical psychometric properties of the Spanish and Portuguese versions of the ZBI used in this study have been published elsewhere [24–26] and showed appropriate internal consistency (Cronbach's  $\alpha$  values of 0.92 and 0.88, respectively) and test-retest reliability (intraclass correlation coefficient [ICC] of 0.93 for the Portuguese ZBI).

### Statistical Analysis

The analyses to obtain a reduced and one-dimensional ZBI version were conducted according to a three-step approach.

First, the 22-item responses were analyzed by a nonparametric IRT analysis (Mokken analysis) to elucidate the latent constructs and likely subscales underlying the association matrix of observed responses [27,28]. A secondary aim was to check in advance the psychometric assumptions associated with the parametric IRT model chosen for the second analytical step: the Samejima's graded response model (GRM) [29]. Two Mokken models were fitted to the data—the monotone homogeneity model and the double monotonicity model. The former aims to test whether a scale total score is a valid tool for ordering and classifying subjects according to the degree of the construct exhibited. The latter is more restrictive because it also aims to identify whether an order exists among the items to rate the corresponding construct that is independent of the selected sample. All items linked to underlying Mokken scales were retained if the scales had at least three items attached to them. The Mokken constructed scales were interpreted according to customary rules of thumb: to be considered as relevant, all items should have a scalability coefficient ( $H_i$ )  $\geq 0.30$ , and also the total scale should have a scalability ( $H$ ) of  $\geq 0.30$ . Mokken [27] suggested the following thresholds to interpret scalability coefficients for a measurement scale: weak scale for  $0.3 \leq H < 0.4$ , medium scale for  $0.4 \leq H < 0.5$ , and strong scale for  $H \geq 0.5$ .

In the second step, we used the Samejima's GRM as the parametric IRT to obtain estimates of the relationship among the latent construct and the item characteristics. Specifically, we estimated the item response characteristic curve parameters (ICC) and item information. If items behave adequately, the ICCs should present an ordered shape discriminating among the category thresholds. Even if overlapping, each category within an item should present a distinct probability of being selected more than any other category for a specific difficulty. We adjusted two GRMs, one assuming equal discrimination among items (restricted model) and other relaxing such assumption (unrestricted model). Because the restricted GRM is nested within the unrestricted model, we selected the most parsimonious model according to the likelihood ratio test.

The third and final step was to assess the unidimensionality of the reduced scale so far obtained by confirmatory factor analysis (CFA). CFA was carried out by using robust weighted least squares on the sample variance-covariance matrix of polychoric correlations among the reduced ZBI items. Goodness of fit for the CFA was evaluated by using the comparative fit index (CFI), the root-mean-square error of approximation (RMSEA), and the weighted root mean square residual. A value of CFI  $> 0.95$  was considered as acceptable model fit, an RMSEA value of  $< 0.08$  was considered to reflect an adequate fit to the model, and a value of  $< 0.05$  was considered as good fit. A weighted root mean square residual value of less than 1 is customarily interpreted as a good value, but its behavior as a goodness-of-fit index is not as well studied as are the CFI and RMSEA indexes. Finally, the reliability of the final reduced scale was evaluated according to both Mokken and Cronbach estimates. The statistical packages R v2.13.1 (nonparametric and parametric IRT with libraries "mokken" and "ltm," respectively) [30,31] and Mplus v5 (CFA) were used to carry on the analyses.

## Results

Table 1 shows the main characteristics of the dyads of caregivers and patients with dementia. As seen, patients were equally distributed in the sample regarding cognitive severity. All caregivers shared home with the patient, were informal (not paid for caring for the patient), and patient's relatives (spouses 50.4%, sons/daughters 44.3%, brothers/sisters and nephews 5.3%). They were mostly females, with a caring time well beyond 8 h/d.

**Table 1 – Main characteristics of caregivers and dementia patients (N = 241).**

Variables	Caregivers	Patients
Gender, n (%)		
Males	55 (22.8)	93 (38.6)
Females	186 (77.2)	148 (61.4)
Age (y), mean $\pm$ SD	62.2 $\pm$ 13.9	77.9 $\pm$ 8.1
Daily caring hours, mean $\pm$ SD	12.6 $\pm$ 6.4	—
22-Item ZBI total score, mean $\pm$ SD	55.8 $\pm$ 14.6	—
MMSE score, mean $\pm$ SD*	—	14.6 $\pm$ 10.3
Severity of dementia, n (%)*		
Severe (MMSE score $\leq$ 9)	—	74 (33.9)
Moderate (MMSE score 10–20)	—	67 (30.7)
Mild (MMSE score $\geq$ 21)	—	77 (35.3)
MMSE, Mini Mental State Examination; ZBI, Zarit Caregiver Burden Inventory.		
* Based on 218 patients.		

In spite of this, most caregivers (67%) had no auxiliary help provided by a paid supported worker.

#### Unidimensionality of the Underlying Scales: Mokken Analysis

According to Mokken analyses, it was possible to single out three unidimensional scales tapping three latent constructs for the 22-item ZBI (Table 2). The major scale included 12 items directly

related with burden, whereas the other two scales presented only two items each and were linked to guilty and embarrassment. Six items could not be ascribed to any underlying latent construct (Table 2). These results lead to the 12-item ZBI selected for further analyses. All items showed scalability indices ( $H_i$ ) over 0.30, and the scalability for the overall scale ( $H$ ) was 0.44, a medium scale by Mokken's criteria (Table 3). The reliability of this 12-item ZBI was good, 0.89 according to both Mokken and Cronbach's  $\alpha$  criteria. The 12-item ZBI scale fitted to a Mokken monotonicity model (monotone homogeneity model) because there were no violations of its assumptions. In case of the double monotonicity model, we found significant violations for all items but 10, 11, and 12; however, the number of violations was not enough to discard the double monotonicity model as presenting a good fit.

#### Item Characteristics: Samejima's Graded Response Analysis

The best fit for the Samejima's GRM was obtained by the unconstrained model, assuming a different discrimination parameter by item. The likelihood ratio test for comparing the constrained and unconstrained models was 132.04 on 11  $df$  ( $P < 0.001$ ). This model retained most of the ZBI information (80.7%; 52.72 for the 12-item ZBI vs. 65.29 for the 22-item ZBI) and showed appropriate ICC with items 2, 3, 9, 17, and 22 as the most discriminative (Table 3). Figures 1 and 2 show the ICC from the Samejima's parametric GRM. As seen, most items presented a shape and category threshold compatible with appropriate difficulty and discrimination parameters. The main exception were item 8 (do you feel your relative is dependent on you?) and to a

**Table 2 – Endorsement frequencies and Mokken scalability (Loevinger's  $H_i$  coefficients) for the items of the Zarit Burden Inventory (N = 241).**

Item number and description	Mean $\pm$ SD	Endorsement frequencies (%)					H <sub>i</sub>
		0	1	2	3	4	
Subscale 1: Burden							
2. Not enough time for myself	1.85 $\pm$ 1.28	48 (20)	46 (19)	69 (29)	51 (21)	27 (11)	0.36
3. Stressed for caring & other responsibilities	1.86 $\pm$ 1.26	48 (20)	40 (17)	78 (32)	48 (20)	27 (11)	0.38
8. Relative's dependence on you	3.37 $\pm$ 0.99	7 (3)	7 (3)	26 (11)	51 (21)	150 (62)	0.24
9. Strained by relative	1.71 $\pm$ 1.27	61 (25)	31 (13)	90 (37)	34 (14)	25 (10)	0.41
10. Health decrease	1.58 $\pm$ 1.37	81 (34)	27 (11)	73 (30)	33 (14)	27 (11)	0.34
11. Lack of privacy	2.08 $\pm$ 1.32	42 (17)	32 (13)	73 (30)	52 (22)	42 (17)	0.36
12. Lack of social life	1.84 $\pm$ 1.45	66 (27)	32 (13)	63 (26)	34 (14)	46 (19)	0.33
16. Unable to care much longer	1.24 $\pm$ 1.22	93 (39)	47 (20)	64 (27)	24 (10)	13 (5)	0.26
17. Lost control of life	1.26 $\pm$ 1.29	99 (41)	43 (18)	54 (22)	28 (12)	17 (7)	0.38
18. Leave the care to someone else	0.94 $\pm$ 1.12	119 (49)	48 (20)	51 (21)	15 (6)	8 (3)	0.28
19. Uncertain about what to do	1.08 $\pm$ 1.12	106 (44)	41 (17)	66 (27)	24 (10)	4 (2)	0.26
22. Overall feeling of burden	1.90 $\pm$ 1.20	38 (16)	45 (19)	91 (38)	38 (16)	29 (12)	0.40
Subscale 2: Guilty							
20. Should do more for my relative	1.12 $\pm$ 1.18	99 (41)	56 (23)	54 (22)	21 (9)	11 (5)	0.18
21. Could do a better job caring	1.16 $\pm$ 1.08	87 (36)	58 (24)	72 (30)	18 (7)	6 (2)	0.13
Subscale 3: Embarrassment							
4. Embarrassment over relative's behavior	0.46 $\pm$ 0.84	173 (72)	33 (14)	27 (11)	7 (3)	1 (0)	0.19
13. Feel uncomfortable having friends over	0.62 $\pm$ 1.11	172 (71)	19 (8)	26 (11)	17 (7)	7 (3)	0.21
Items not fitting any scale							
1. Asking for more help than needed	1.45 $\pm$ 1.31	80 (33)	48 (20)	59 (24)	33 (14)	21 (9)	0.21
5. Angry	1.16 $\pm$ 1.10	91 (38)	51 (21)	74 (31)	19 (8)	6 (2)	0.25
6. Negative relationships	1.29 $\pm$ 1.31	98 (41)	39 (16)	57 (24)	29 (12)	18 (7)	0.27
7. Afraid about the future	2.60 $\pm$ 1.26	24 (10)	17 (7)	63 (26)	64 (27)	73 (30)	0.12
14. Expecting to be cared by you	2.55 $\pm$ 1.50	41 (17)	21 (9)	40 (17)	42 (17)	97 (40)	0.15
15. Lack of enough money to pay for the expenses	1.37 $\pm$ 1.38	99 (41)	34 (14)	51 (21)	34 (14)	23 (10)	0.21

Note. Endorsement frequencies: 0: never; 1: rarely; 2: sometimes; 3: quite frequently; 4: nearly always.

**Table 3 – Mokken's scalability (Loevinger's  $H_i$  coefficients), Samejima's graded response model parameters, and standardized loadings from confirmatory factor analysis for the Zarit Burden Inventory (N = 241).**

Item no.	$H_i$	$a$	$b_0$	$b_1$	$b_2$	$b_3$	Information	CFA loadings
2	0.47	2.02	− 1.15	− 0.42	0.55	1.63	5.41	0.75
3	0.48	2.03	− 1.15	− 0.51	0.58	1.63	5.39	0.77
8	0.31	0.70	− 5.32	− 4.26	− 2.50	− 0.78	1.39	0.42
9	0.52	2.66	− 0.82	− 0.39	0.81	1.56	7.31	0.84
10	0.46	1.89	− 0.60	− 0.22	0.87	1.68	4.39	0.74
11	0.49	1.90	− 1.28	− 0.68	0.36	1.27	4.69	0.72
12	0.43	1.54	− 0.89	− 0.35	0.60	1.28	3.17	0.64
16	0.34	1.01	− 0.59	0.35	2.00	3.31	2.25	0.52
17	0.50	2.10	− 0.33	0.24	1.15	2.01	5.22	0.78
18	0.39	1.17	− 0.05	0.83	2.33	3.42	2.66	0.60
19	0.31	0.89	− 0.34	0.57	2.62	5.08	2.15	0.48
22	0.54	2.86	− 1.22	− 0.63	0.66	1.43	8.69	0.85

a, discrimination parameter; b's, threshold parameters; CFA, confirmatory factor analysis.

lesser extent items 18 (do you wish you could leave the care of your relative to someone else?) and 19 (do you feel uncertain about what to do about your relative?).

#### Assessing Unidimensionality: CFA

The 12-item ZBI CFA results fitted to a unidimensional latent structure (CFI = 0.975; RMSEA = 0.067; weighted root mean

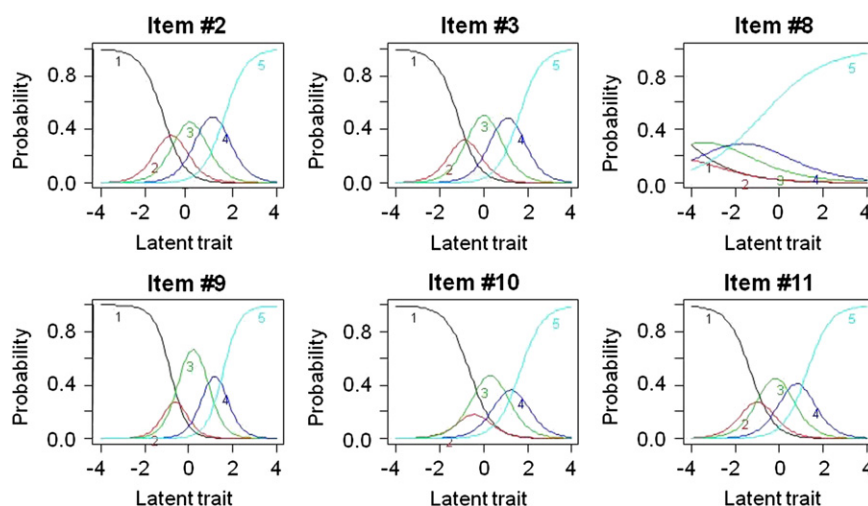


Fig. 1 – Item response category characteristic curves for 12-item short ZBI.

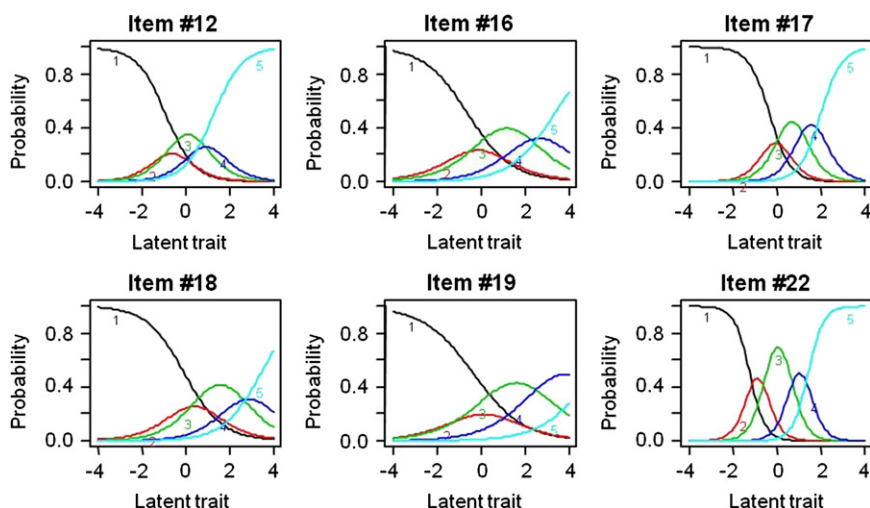


Fig. 2 – Item response category characteristic curves for 12-item short ZBI.

square residual = 0.677). Inspection of standardized residuals and modification indexes indicated no localized points of ill fit in the final solution apart from including a correlated measurement error between items 11 and 12 ( $r = 0.51$ ). All freely estimated unstandardized parameters were statistically significant ( $P < 0.001$ ). Table 3 displays the completely standardized parameter estimates (factor loadings), showing loadings for all of them above 0.40 and showing items 9 (strained by the relative) and 22 (overall feeling of burden) as the items more related with burden.

## Discussion

The analyses here reported point to a feasible short 12-item ZBI obtained from a multicentric clinical trial designed to assess the efficacy of a psychoeducational intervention with caregivers of patients with dementia. The unidimensionality of the 12-item ZBI suggested by Mokken analysis was confirmed by CFA. We interpret this structure as corresponding to the latent construct of subjective burden experienced by the caregiver.

Table 4 presents a comparative view of the structure of the several ZBI short forms published until now [32–35]. As seen, our suggested scale is quite close to that proposed by Bédard et al. [32], with both scales sharing eight items over 12. It is interesting to note the high variability of items ascribed to the different short scales. In fact only three items were not presented in any of the reported short forms of the parental ZBI (item 1: asking for more help than needed, item 7: afraid about the future, and item 15: financial problems).

Some authors [19] have proposed an extreme 1-item reduction of the 22-item ZBI, suggesting that to obtain an overall summary of burden no more than item 22 is needed. It is true that the ICC for this item behaves quite well (Fig. 2) and also retained a good percentage of the total information and a strong factorial loading on the latent construct (Table 3). However, it is not less true that item 22 acts as an anchor item to summarize the whole ZBI scale once the caregiver has answered all the former items. As such, the measurement of item 22 alone to represent the overall subjective burden experienced by a caregiver presents an intrinsic bias because its scoring depends on the learning process the caregiver has had when scoring the previous items. In other words, item 22's score is as good as the summary of all previous scoring, and thus its analysis in isolation from the rest of the information does not seem to be advisable.

As a limitation of the study, we acknowledge that several items in our proposal for a unidimensional ZBI short scale do not present a good ICC. But because their elimination did not improve the fit as assessed by CFA, we opted for their maintenance. It is highly likely that by modifying the content or the wording of such items its ICC shape would be more appropriate; however, they matched the underlying construct according to the Mokken analysis and as said they did not worsen the CFA fits to a unidimensional construct. Another limitation is that our study is based on ZBI responses obtained from a convenience sample of caregivers recruited within a clinical trial and thus do not correspond to a representative sample of caregivers as has been done by others [36].

In conclusion, the 12-item ZBI derived from IRT has shown good ICC properties. It fits a one-dimensional latent variable

**Table 4 – Items included in several short forms of the Zarit Caregiver Burden Interview.**

Item no. and description	Bédard et al. [32]	Bédard et al. [32]	Gort et al. [33]	Gort et al. [34]	Arai et al. [35]	Higginson et al. [19]	This article
1. Asking for more help than needed							
2. Not enough time for myself	X	X	X	X		X	X
3. Stressed for caring & other responsibilities	X	X	X			X	X
4. Embarrassment over relative's behavior					X		
5. Angry		X		X	X		
6. Negative relationships		X	X		X	X	
7. Afraid about the future							
8. Relative's dependence on you							X
9. Strained by relative	X	X	X	X	X	X	X
10. Health decrease		X	X			X	X
11. Lack of privacy		X					X
12. Lack of social life		X			X		X
13. Feel uncomfortable having friends over					X		
14. Expecting to be cared by you				X			
15. Lack of money to pay for the expenses							
16. Unable to care much longer							X
17. Lost control of life		X	X			X	X
18. Leave the care to someone else					X		X
19. Uncertain about what to do	X	X			X		X
20. Should do more for my relative		X					
21. Could do a better job caring		X					
22. Overall feeling of burden			X				X



related with burden that could be relevant to properly interpret its total score change according to interventions or over time. Further validations with other languages, most notably English, and additional psychometric studies, mainly focusing on its psychometric equivalence on representative samples of caregivers, on its sensitivity to change over time, and on the minimal important difference according to several scenarios are warranted.

## Acknowledgments

We thank the centers and collaborative authors of the EDUCA-2 trial: Centro de Día Psicogeriatrico Txurdinagabbarri, Bilbao (José Ezkerra, Nora León García, Leire Gómez Galván); Centro de Día Psicogeriatrico Joxe Migel Barandiaran, Durango (Nagore Mar-daras Barrenechea, Eva González Gerrikaetxeberria, Eduardo González Fraile); Centro de Día Psicogeriatrico Aita Menni, Arrasate, Centro de Día Psicogeriatrico Aita Menni, Eskoriatza, Centro de día Municipal de Oñate (Joana García Marcos, Elena Berecibar Ciorraga, M<sup>a</sup> José Novas Sánchez); Centro de día Psicogeriatrico Aita Menni, Deusto, Bilbao (Leire Gómez Galván, Joana García Marcos, Eduardo González Fraile); Centro Hospitalario Padre Benito Menni, Valladolid (Christian Eduardo Rueda-Calusen, Carlos Francisco Salgado Pascual); Complejo Hospitalario San Luis, Palencia (Ignacio Herreros Guilarte, Javier Aguado Rodríguez); Centro Hospitalario Padre Menni, Santander (Carmen García Orallo, Emma Fernández Álvarez); Clínica Psiquiátrica Padre Menni, Pamplona (Míriam García Sanabria, Paula Muñoz Hermoso); Sagrat Cor, Serveis de Salut Mental, Martorell (Manuel Sánchez Pérez, Francesca Amores Guevara, Carla Subirós Castresana-Isla, Ana García Plata, David Serrano Santó, Diana Romero Domínguez); Benito Menni, Complex Assistencial en Salut Mental, Sant Boi (Consuelo Almenar Monfort, M<sup>a</sup> Luisa Sánchez Guerra, Consuelo Casado Montañés, Carmen García Membrive); Hospital Mare de Déu de La Mercé, Barcelona (Cristobal Diez Aja, César Obiol Madrid, M<sup>a</sup> del Mar Ruiz Vergara); Igorco, Servicios Sociosanitarios, Bilbao (Amaia Cantera Aranburu, Carla Pérez de Muniain Lumbreras, Ainara Castaños Mejuto); Complejo Asistencial Hermanas Hospitalarias del Sagrado Corazón, Málaga (M<sup>a</sup> Dolores Nieto de Haro, David Porta Pelayo, Belén Guzmán); Clínica Psiquiátrica de San José, Lisboa, Portugal (Manuel Gonçalves Pereira, Rita Margarida Santos Costa, Jaime Manuel Novo Grácio, Catarina Madeira Janeiro); Centro Psicogeriatrico San Francisco Javier, Pamplona (Germán Jusué Erro, Concepción Esparza Lizasoain, Luisa Taus Ibarben, David Brugos Miranda); Casa de Saúde da Idanha, Belas, Portugal (Carla Alexandra da Silva Pombo, Ana Isabel Rogrigues de Moraes); Complejo Asistencial Benito Menni, Ciempozuelos, Madrid (Elisa Sánchez Santamaría, Antonio José Briz Iglesias, Susana Postigo Navarro); Fundación Aspalidiko, Portugalete, Bilbao (María García Escudero, Tamara Arenal Cantalapiedra, Ainhoa Sainz Ibarra).

Source of financial support: The EDUCA-2 trial was funded by FIS (PI08/90812). J.B., B.S., and E.G.F. were supported by UPV/EHU (GIU07/07 and GIU10/24).

## REFERENCES

- [1] Zarit SH, Reeve KE, Bach-Peterson J. Relatives of the impaired elderly: correlates of feelings of burden. *Gerontologist* 1980;20:649–55.
- [2] Zarit SH, Orr NK, Zarit JM. The Hidden Victims of Alzheimer's Disease: Families Under Stress. New York: New York University Press, 1985.
- [3] Zarit SH, Todd PA, Zarit JM. Subjective burden of husbands and wives as caregivers: a longitudinal study. *Gerontologist* 1986;26:260–6.
- [4] Whitlatch CJ, Zarit SH, von Eye A. Efficacy of interventions with caregivers: a reanalysis. *Gerontologist* 1991;31:9–14.
- [5] Knight BG, Fox LS, Chou CP. Factor structure of the burden interview. *J Clin Geropsychol* 2000;6:249–57.
- [6] Ankrj J, Andrieu S, Beaufils B, et al. Beyond the global score of the Zarit Burden Interview: useful dimensions for clinicians. *Int J Geriatr Psychiatry* 2005;20:254–60.
- [7] Lu L, Wang L, Yang X, Feng Q. Zarit Caregiver Burden Interview: development, reliability and validity of the Chinese version. *Psychiatry Clin Neurosci* 2009;63:730–4.
- [8] Ostwald SK, Hepburn KW, Caron W, et al. Reducing caregiver burden: a randomized psychoeducational intervention for caregivers of persons with dementia. *Gerontologist* 1999;39:299–309.
- [9] Gonyea JG, O'Connor MK, Boyle PA. Project CARE: a randomized controlled trial of a behavioral intervention group for Alzheimer's disease caregivers. *Gerontologist* 2006;46:827–32.
- [10] Martín-Carrasco M, Martín MF, Valero CP, et al. Effectiveness of a psychoeducational intervention program in the reduction of caregiver burden in Alzheimer's disease patients' caregivers. *Int J Geriatr Psychiatry* 2009;24:489–99.
- [11] Buchanan KM, Elias LJ. Psychological distress and family burden following spinal cord injury: concurrent traumatic brain injury cannot be overlooked. *Axone* 2001;22:16–7.
- [12] Leal M, Sales R, Ibañez E, et al. Evaluation of the effect of a psychoeducational program on the burden in informal caregivers of patients with schizophrenia. *Actas Esp Psiquiatr* 2008;36:63–9.
- [13] Chung ML, Pressler SJ, Dunbar SB, et al. Predictors of depressive symptoms in caregivers of patients with heart failure. *J Cardiovasc Nurs* 2010;25:411–9.
- [14] Higginson IJ, Costantini M, Silber E, et al. Evaluation of a new model of short-term palliative care for people severely affected with multiple sclerosis: a randomised fast-track trial to test timing of referral and how long the effect is maintained. *Postgrad Med J* 2011;87:769–75.
- [15] Bachner YG, O'Rourke N. Reliability generalization of responses by care providers to the Zarit Burden Interview. *Aging Mental Health* 2007;11:678–85.
- [16] Siegert RJ, Jackson DM, Tennant A, Turner-Stokes L. Factor analysis and Rasch analysis of the Zarit Burden Interview for acquired brain injury carer research. *J Rehabil Med* 2010;4:302–9.
- [17] O'Rourke N, Tuokko HA. The relative utility of four abridged versions of the Zarit Burden Interview. *J Mental Health Aging* 2003;9:55–64.
- [18] Longmire CVF, Knight BG. Confirmatory factor analysis of a brief version of the Zarit Burden Interview in black and white dementia caregivers. *Gerontologist* 2011;51:453–62.
- [19] Higginson IJ, Gao W, Jackson D, et al. Short-form Zarit Caregiver Burden Interviews were valid in advanced conditions. *J Clin Epidemiol* 2010;63:535–42.
- [20] Thomas ML. The value of item response theory in clinical assessment: a review. *Assessment* 2011;18:291–307.
- [21] Santor DA, Debrota D, Engelhardt N, et al. Optimizing the ability of the Hamilton Depression Rating Scale to discriminate across levels of severity and between antidepressants and placebos. *Depress Anxiety* 2008;25:774–86.
- [22] Levine SZ, Rabinowitz J, Rizopoulos D. Recommendations to improve the Positive and Negative Syndrome Scale (PANSS) based on item response theory. *Psychiatry Res* 2011;188:446–52.
- [23] Khan A, Lewis C, Lindenmayer JP. Use on non-parametric item response theory to develop a shortened version of the Positive and Negative Syndrome Scale (PANSS). *BMC Psychiatry* 2011;11:178.
- [24] Martín Carrasco M, Ballesteros Rodríguez J, Ibarra Gandiaga N, et al. Alzheimer's caregiver burden and psychological distress: a neglected association in the assessment of dementias. *Actas Esp Psiquiatr* 2002;30:201–6.
- [25] Martín-Carrasco M, Otermin P, Pérez-Camo V, et al. Psychometric properties of the Spanish version of the Zarit Caregiver Burden Scale. *Aging Mental Health* 2010;14:705–11.
- [26] Gonçalves-Pereira M, Carmo I, Alves da Silva J, et al. Caregiving experiences and knowledge about dementia in Portuguese clinical outpatient settings. *Int Psychogeriatr* 2010;22:270–80.
- [27] Mokken RJ. Nonparametric models for dichotomous responses. In: van der Linden WJ, Hambleton RK, editors. *Handbook of Modern Item Response Theory*. New York: Springer-Verlag; 2010. p. 351–67.
- [28] Molenaar IW. Nonparametric models for polytomous responses. In: van der Linden WJ, Hambleton RK, editors. *Handbook of Modern Item Response Theory*. New York: Springer-Verlag; 2010. p. 369–80.
- [29] Samejima F. Graded response model. In: van der Linden WJ, Hambleton RK, editors. *Handbook of Modern Item Response Theory*. New York: Springer-Verlag; 2010.
- [30] Van der Ark LA. Mokken scale analysis in R. *J Stat Software* 2007;20:1–19.
- [31] Rizopoulos D. ltm: an R package for latent variable modeling and item response theory analyses. *J Stat Software* 2006;17:1–25.

- 
- [32] Bédard M, Molloy DW, Squire L, et al. The Zarit Burden Interview: a new short version and screening version. *Gerontologist* 2001;41:652–7.
- [33] Gort AM, March J, Gómez X, et al. Escala de Zarit reducida en cuidados paliativos. *Med Clin (Barc)* 2005;124:651–3.
- [34] Gort AM, Mingot M, March J, et al. Utilidad de la escala de Zarit reducida en demencias. *Med Clin (Barc)* 2010;135:447–9.
- [35] Arai Y, Tamiya N, Yano E. The short version of the Japanese version of the Zarit Caregiver Burden Interview (J-ZBI-8): its reliability and validity. *Nippon Ronen Igakkai Zasshi* 2003;40:497–503.
- [36] O'Rourke N, Tuokko HA. Psychometric properties of an abridged version of the Zarit Burden Interview within a representative Canadian caregiver sample. *Gerontologist* 2003;4:121–7.