Behavioral Psychology / Psicología Conductual, Vol. 30, Nº 3, 2022, pp. 787-808
https://doi.org/10.51668/bp.8322312n

ASSESSING IRRATIONAL BELIEFS RELIABLY AND QUICKLY: REFINING AND SHORTENING THE SPANISH VERSION OF THE ATTITUDES AND BELIEFS SCALE

Patricia Gual-Montolio[1], Verónica Martínez-Borba[1,2], Macarena Paredes-Mealla[1,] Jorge Osma[2,3] and Carlos Suso-Ribera[1,4]
[1]Jaume I University; [2] Health Research Institute of Aragon; [3]University of Zaragoza; [4]CIBER Physiopathology of Obesity and Nutrition (CIBERObn), Carlos III Health Institute (Spain)

Abstract

The Attitudes and Beliefs Scale (ABS) is a widely used measure of irrational beliefs (IBs), but has important psychometric problems. Our objective is to improve the psychometric quality of a Spanish version of the scale. Classical test theory, item response theory, and confirmatory factor analyses were combined to obtain a shorter version of the scale using 2 samples: one from the general population (n= 565) and another with chronic pain (n= 514). Pearson correlations were performed with IBs, personality and health measures to investigate sources of construct validity. After eliminating half of the items (12), the factorial fit of the scale became very good (RMSEA< .08; CFI and TLI> .95). IBs were associated with more neuroticism (.21≤ r≤ .61, p≤ .001) and poorer mental health (-.17≤ r≤ -.56, p≤ .001), as well as a less extraversion and conscientiousness (-.14≤ r≤ -.41, p≤ .01). These results were replicated in both samples, but IBs were only associated with poorer physical health in the general population sample. The shortened Spanish version of the ABS is a valid and reliable instrument that can be rapidly administered in clinical settings.

KEY WORDS: *attitudes and beliefs scale, rational emotional behavioral therapy, irrational beliefs, item response theory, classical test theory*.

Resumen

La "Escala de actitudes y creencias" (EAC) es una medida de creencias irracionales (CI) muy utilizada, pero con problemas psicométricos. Nuestro objetivo fue mejorar la calidad psicométrica de la versión española de la EAC. Se combinó la teoría clásica de los tests, teoría de respuesta al ítem y análisis factorial confirmatorio para obtener una versión corta de la escala utilizando dos muestras, una de la población general (n= 565) y otra con dolor crónico (n= 514). Se realizaron correlaciones de Pearson con CIs, personalidad y medidas de salud para investigar las fuentes de validez de constructo. Tras eliminar la mitad de los

---

We would like to thank all the participants and professionals who voluntarily participated in this study.

Correspondence: Carlos Suso-Ribera, Av. Vicent Sos Baynat, s/n, 12071 Castellón de la Plana, Castellón (Spain). E-mail: susor@uji.es

ítems (12), el ajuste factorial de la escala fue bueno (RMSEA< 0,08; CFI y TLI> 0,95). Las CIs se asociaron con más neuroticismo (0,21≤ r≤0,61; p≤ 0,001), peor salud mental (-0,17≤ r≤ -0,56; p≤ 0,001), menor extraversión y responsabilidad (-0,14≤ r≤ -0,41; p≤ 0,01). Estos resultados se replicaron en ambas muestras, pero las CIs sólo se asociaron con una peor salud física en la población general. La versión española abreviada de la EAC es válida, fiable y puede administrarse rápidamente en entornos clínicos.
PALABRAS CLAVE: *escala de actitudes y creencias, terapia racional emotiva conductual, creencias irracionales, teoría de respuesta al ítem, teoría clásica de tests*.

Introduction

According to rational emotional behavioural therapy (REBT), irrational beliefs (IBs) are rigid, extreme, and illogical attitudes that people strongly hold in relation to a given situation, even when there is evidence that contradicts such beliefs (Ellis, 1995; Pérez-Acosta et al., 2008). IBs are considered to be cognitive vulnerability factors associated with the development of psychological distress in the presence of negative situations (David et al., 2010). Specifically, when personal goals, wishes or values (people's Goals, G) are frustrated or hindered by external factors or certain events/situations (Activating events, A), personal beliefs about oneself, the others, and the world are activated (Beliefs, B). Depending on the rational/irrational nature of such beliefs, emotional and behavioural consequences (Consequences, C) become healthy or unhealthy. For example, when individuals hold rational beliefs (e.g., "I wish the situation was easier, but this is only a wish and not something I can demand; I can tolerate when things do not occur as planned"), such beliefs would lead to healthy consequences, such as sadness (as opposed to depression) and the assertive expression of one's feelings to significant others. On the contrary, IBs (e.g., "Things should be easier and I can't stand when situations do not occur as I demand") would lead to unhealthy reactions, such as depressed feelings and attempts to put an end to such feelings using maladaptive coping strategies (Turner, 2016).

IBs have received a great deal of attention in clinical research (Vîslă et al., 2016; Rovira et al., 2020) and have been repeatedly associated with a wide range of psychological problems, including depressive, anxiety, and stress symptoms (Belloch et al., 2010; Chan & Sun, 2021; Duru & Balkıs, 2021); but also, with poor physical functioning and poor adaptation to medical conditions (e.g., chronic pain; Miró, Queral & Nolla, 2014; Morris et al., 2017; Suso-Ribera, Camacho-Guerrero, et al., 2019). Not surprisingly, IBs have been at the core of cognitive-behavioural interventions for decades (David et al., 2010). Encouragingly, psychological interventions using cognitive-behavioural orientations, including REBT, are effective in modifying IBs (Turner, 2016) and changes in IBs are, at least partly, responsible for the benefits of psychotherapy on outcomes (e.g., depression; Szentagotai et al., 2008). While the clinical utility of challenging irrational thoughts is well grounded, researchers have raised concerns about the reliability of the existent measures of IBs (David et al., 2010; Hyland et al., 2017), which would ultimately question the credibility of the clinical findings that result from such measures.

According to REBT, IBs can occur in the form of four irrational inferences or processes, namely (a) demandingness (absolute and inflexible requirements in the form of "musts", "should", "have to" and "ought"); (b) awfulizing/catastrophizing (believing that not having one's desires met would be the worst thing that could occur); (c) low frustration tolerance (LFT) (believing that it is unbearable to have one's desires or demands hampered); and (d) self-downing (global self-criticism about oneself, other persons and/or life; Lega & Ellis, 2001). A number of self-report measures for IBs exist, including the Attitudes and Beliefs Scale (ABS; Burgess, 1986; DiGiuseppe et al., 1988), the Attitudes and Beliefs Inventory (Burgess, 1990), or the Survey of Personal Beliefs (Demaria et al., 1989), to name some examples. Of these, the ABS is the most popular, arguably due to its fidelity with the REBT theory (David et al., 2010). Research to date, however, has generally failed to support the use of the previous measures of IBs, including the ABS, due to poor model fit (i.e., inability to replicate the four-factor structure) and construct validity (Artiran & DiGiuseppe, 2020; David et al., 2010; Hyland et al., 2017). In the Spanish context, for example, the ABS was adapted but not formally validated more than two decades ago (Caballo et al., 1996) and subsequent studies also identified psychometric problems with the scale, including low internal consistency and poor discriminant validity (Ruiz-Rodríguez et al., 2020; Suso-Ribera et al., 2016).

Without an adequate measure of irrational thinking that is consistent with the REBT theory, advances in the scientific knowledge of IBs and clinical conclusions derived from their assessment will both be questionable. Attempts to improve the psychometric properties of the ABS have been made by shortening the scale (i.e., eliminating problematic items) and the results have been promising (Hyland et al., 2017; Owings et al., 2013). Such reductions also make the ABS more practical in routine care by reducing its administration time. The present study aims to further contribute to the literature into the assessment of IBs. In particular, we intent to obtain a short, theory-consistent, and psychometrically-sound measure of the ABS to be used in the Spanish population. In doing so, we will follow the recommended steps when reducing questionnaires (Goetz et al., 2013) and will combine classical test theory (CTT) and item response theory (IRT). For reliability and generalizability purposes, we will conduct the analyses separately in two samples (a sample from the general population and a sample of persons with chronic pain). Based on previous similar attempts to reduce the length of the ABS, we expect to obtain a short and psychometrically-sound version of the Spanish ABS. Specifically, we anticipate that the four-factor solution of the short ABS based on the REBT theory (i.e., demanding, awfulizing, LFT, and self-downing) will have an adequate fit. We also hypothesize that the sources of validity evidence will reveal a good construct validity of the ABS, in the form of positive and moderate associations between IBs and neuroticism (Samar et al., 2013) and mental distress (Vîslă et al., 2016) and positive but milder correlations with physical disability (Suso-Ribera et al., 2016).

Method

Participants

The study included two samples that were similar in size (Table 1). In study 1, participants were 565 individuals from the general population aged between 18 and 93 years (M= 43.91, SD= 19.03, 61.6% women). ıRegarding job status, they were either active workers (50.5%), university students (19.6%), retired (17.9%), or unemployed (11.9%). The majority of the participants were married (53%) at the time of assessment. The remaining participants were in a relationship (15.2%), single (14.0%), widowed (14.5%), or separated/divorced (3.4%). Regarding their educational level, 27.7% had completed secondary studies, 25.5% were studying at the university, 24.5% had completed primary studies, and 22.3% had technical studies.

Table 1
Sociodemographic characteristics of the study samples

Variables	General sample (n_1= 565)	Chronic pain (n_2= 514)
Mean age, years (SD)	43.91 (19.03)	58.99 (15.05)
Gender (%)		
Man	38.4	36.1
Women	61.6	63.9
Educational level (%)		
Primary	24.5	13.6
High school	27.7	41.7
Technical studies	22.3	18.4
University	25.5	26.3
Job status		
Working	50.5	29.7
Unemployed/retired/sick leave	29.9	69.9
Student	19.6	0.4
Marital status		
Married/in a relationship	68.2	60.2
Single/separated/divorced/widowed	31.8	39.8

In study 2, the sample included 514 participants (63.9% women) with chronic pain who attended the Pain Clinic of the Vall d'Hebron Hospital (Barcelona, Spain). The participants' age ranged from 18 to 89 years (M= 58.99, SD= 15.05). Most participants were not working at the time of assessment. Specifically, 54.6% of patients were retired and 15.5% were unemployed. Only 29.7% of them were active workers, and a small proportion of them were students (0.4%). The majority of the participants were married (60.1%). The remaining participants were either single (12.8%), widowed (15.0%), or separated/divorced (12.0%). Regarding literacy level, part of the sample had only completed primary education studies (13.6%). The remaining participants had completed secondary studies (41.7%), university education (26.3%), or technical studies (18.4%). The main pain diagnoses included musculoskeletal pain in the low back and neck.

Instruments

a) *Attitudes and Beliefs Scale* (ABS; Burgess, 1986; DiGiuseppe et al., 1988), Spanish adaptation by Caballo et al. (1996). The ABS has 24 IBs grouped into four processes (demandingness, awfulizing, LFT, and self-downing). In the ABS, participants are requested to rate their agreement in a series of statements using a 5-point Likert-type scale ranging from 0 "strongly disagree" to 4 "strongly agree". Higher scores reflect more irrational thinking. The internal consistency of the reduced version of the EAC obtained in this study was good for both samples in the four processes ($.67 \leq \alpha \leq .88$).

b) NEO-Five Factor Inventory (NEO-FFI; Costa & McCrae, 1992), Spanish adaptation by Solé i Fontova (2006). The NEO-FFI evaluates the Five Factor Model of personality (i.e., neuroticism, extraversion, openness, agreeableness, and conscientiousness) with 60 items. Participants respond to a 5-point Likert scale ranging from strongly disagree (0) to strongly agree (4). Each personality dimension can obtain values between 0 and 60 points. Higher scores indicate greater neuroticism, extraversion, openness, agreeableness, or conscientiousness, according to the corresponding dimension. The internal consistency of the personality dimensions in the present study was good for both samples ($.68 \leq \alpha \leq .86$).

c) *Short Form-36 Health Survey* (SF-36; Ware & Sherbourne, 1992), Spanish adaptation by Alonso et al. (1998). Physical and mental health were evaluated with the SF-36. This questionnaire includes eight dimensions of health (physical functioning, role physical, bodily pain, general health, vitality, social functioning, role emotional, and mental health) which are grouped into two composite scores of physical (PCS) and mental health (MCS). The health dimensions are evaluated using Likert-type scales, but in each case the number of response points and the text of each response option vary. The composites have a 0-100 range. Higher scores represent better health. As indicated in Table 4, the Cronbach's alphas of the PCS and the MCS in the present study were good for both samples ($.75 \leq \alpha \leq .96$).

Procedure

Participants in study 1 were recruited at the University of Barcelona, through poster advertisements, and using social media. When the participants could not complete the assessment protocol at the university, they contacted us by mail or phone and we sent them a pre-paid envelope with the complete protocol so they could complete it at home. A snow-ball sampling approach was also used and all the participants could ask for pre-paid envelopes with the complete protocol so that their relatives and friends could participate.

On the other hand, participants in study 2 completed the questionnaires onsite at the clinic the day of their first appointment.

In both samples, the assessment protocol included an explanation of the study, its procedures and risks, the contact information of the lead researcher (C.S.R.), an informed consent form, and the questionnaires. The assessment took approximately 20 minutes to complete and there was no economic compensation to participate. The ethics committee of the Vall d'Hebron Hospital and the University of Barcelona approved the study and its procedures.

Eligibility criteria for studies 1 and 2 included (1) being 18 years old or older; and (2) being able to read in Spanish. For study 2, an additional criterion was that the participants had experienced chronic pain for at least 6 months, and that they were attending the Pain Clinic of the Vall d'Hebron Hospital (Barcelona, Spain).

Data analysis

SCALE REDUCTION AND ITEM SELECTION

To shorten a questionnaire, it is necessary to reduce the number of items while preserving or enhancing its psychometric properties. To do so, we will follow the recommendations made in past research (Goetz et al., 2013). First, we will report the validity of the original scale (ABS). Next, when reducing the scale, we will take the conceptual model into account (i.e., the four-factor structure with demandingness, awfulizing, LFT, and self-downing) and try to preserve content validity (i.e., items that reflect well the irrational construct). Finally, we will attempt to improve or at least preserve the original psychometric properties, document the reasons for item selection, and validate the findings in two independent samples. Because multi-item scales clearly exceed scales composed by single items (Diamantopoulos et al., 2012) and a minimum of three items for each construct or factor are needed to demonstrate satisfactory internal consistency (Froman, 2001; Raubenheimer, 2004), we will attempt to obtain a 12-item version of the ABS (4 factors x 3 items).

Classical test theory (CTT) and item response theory (IRT) are two of the most common approaches for item selection and reduction (Jin et al., 2018). According to CTT, each respondent has a true total score, T (latent variable), and each item is representative of the T score. On the other hand, IRT assumes the unidimensionality of the latent trait of a measure and the interindependence of all items. While CTT assesses difficulty and discrimination at the item level and the reliability at the whole measure level, IRT estimates the performance (i.e., discrimination, location, and information) of each item by using a set of logistic regression models. IRT overcomes the deficiencies of CTT, but requires larger sample sizes to fit the model (Holmes & Bolin, 2017; Toland, 2014). In addition to CTT and IRT, we will conduct a confirmatory factor analysis (CFA) to validate the fit of the short version of the ABS. Two samples will be used for reliability and generalizability purposes. Because the factor structure of the ABS is well-grounded on REBT theory, a CFA is preferred over an exploratory factor analysis (EFA) based on previous recommendations on the refinement of questionnaire (Goetz et al., 2013). Thus, in this study we will conduct a CFA of the original and reduced Spanish version of the ABS to investigate and compare their fit.

Item reduction based on CTT

Based on CTT, we first used descriptive statistics (i.e., means and standard deviation) to eliminate items with unclear, ambiguous, or extreme scores because they provide little useful information or are too complicated to interpret (Streiner et al., 2015). To define the cut-off for the elimination of items, we followed the recommendations from past research and selected the lowest score option plus 20% of the score range and the highest score option minus 20% of the score (Jin et al., 2018). In the ABS, scores have a 4-point range from 1 (lowest value) to 5 (highest value). Therefore, average item scores lower than 1.8 or higher than 4.2 were considered to be extreme. Again in line with past research (Jin et al., 2018), standard deviations smaller than 1/6 of the score range (i.e., 1/6*4= 0.67) were also considered to be inadequate due to little variability in the responses. Items that met these exclusion criteria or had missing rates higher than 10% were candidates for elimination.

Additionally, we examined the internal consistency reliability (item-total correlation and Cronbach's alpha) of the four scales to examine the items at a domain level. Larger corrected item-total correlations are preferred. A correlation below 0.3 indicates poor association with the scale and suggests that the item should be eliminated.

Item reduction based on IRT

IRT is a more modern way to evaluate the relationship between items (i.e., the observed manifestations) and latent trait variables (latent construct; Finch & Bolin, 2017; Toland, 2014). Parametric IRT analysis are based on four assumptions, namely unidimensionality, local independence, functional form, and normal distribution of the latent variable (Toland, 2014).

Several IRT models exist depending on item responses (i.e., dichotomous or polytomous). Because the ABS uses a polytomous Likert-type scale, we will implement the parametric unidimensional model applicable to Likert-type item responses, known as the graded response model (GRM). The GRM is an extension of the two-parameter logistic (2PL) model developed for dichotomous item response (Toland, 2014). It incorporates individual discrimination parameter values (estimates) for each item and expresses the comparison of item response choices in terms of the likelihood of an individual selecting a particular response option or one higher, versus responding with a lower option (Holmes & Bolin, 2017).

In IRT, the parameter estimate is specific to the respondents (person parameter) and estimates the latent trait being measured by the scale (e.g., irrationality). Items with poor discrimination (those that do not effectively differentiate respondents with higher vs. lower levels of the trait) are automatically given low estimate weightings (Owings et al., 2013). The item parameters include the location or difficulty on the latent trait scale (*b*), the ability to differentiate among individuals with different levels of the construct (item discrimination, *a*), and, in some cases, the probability that an individual will support the item due to chance (*c*). Polytomous IRT models are characterized by multiple item location parameters, known as thresholds, which correspond to intersections of different response options on the latent trait scale (level of irrationality needed to have equal probability of choosing to respond above a given threshold).

The GRM estimates a unique slope parameter for each item across the ordinal response categories along with multiple between-category thresholds (e.g., b1 to b3) for items having more than two categories. Because each item on the ABS scale has five ordered response categories, there are 5-1= 4 threshold parameters and one unique slope parameter to be estimated for each item.

To conduct IRT without violating IRT assumptions (e.g., unidimensionality), we conduct an analysis separately for each of the four factors. The GRM forms each item with its own discrimination parameter and a set of parameters that identify the boundaries between the ordered options using a logistic regression approach. The "information" (reliability) that each item contributes to the factor will be assessed with the item information functions (IIFs) on the fitted GRMs. Better items are those that provide larger amounts of information. Items with a discrimination estimation below 1.0 and item information below 0.5 will be candidates for elimination (Owings et al., 2013).

FINAL EVALUATION OF THE FACTOR STRUCTURE FIT AND THE SOURCES OF CONSTRUCT VALIDITY EVIDENCE

As a final step, we evaluated and compared the fit of the two versions of the questionnaire (full length and shortened) by means of a CFA. Model fit was evaluated by means of the comparative fit index (CFI), the Tucker-Lewis index (TLI), and the root mean square error of approximation (RMSEA). CFI and TLI indicate excellent fit when values are over .95. RMSEA scores smaller than .08 indicate good fit (Schreiber et al., 2006). Additionally, we explored sources of validity evidence of the short version. In particular, we conducted Pearson correlations with the measures of personality and health status.

Results

The results of the item analysis based on CTT and IRT in the two study samples (general population and persons with chronic pain) are shown in Table 2. In the Table 2, all the infractions have been italicized for clarity reasons. The demand and the self-downing scales were the most problematic, generally due to extreme responses (ceiling effect for demand and floor effect for self-downing) or low discrimination indices (<1.0). After an inspection to the results in both samples, we eliminated items 13, 25, and 33 from the demanding scale; items 14, 18, and 22 from the awfulizing scale; items 3, 7, and 43 from the LFT scale; and, items 4, 40, and 48 from the self-downing scale. This was done after a consensus between all the study authors.

Table 2

Item reduction based on Classical Test Theory and Item Response Theory in the two study samples

	Sample 1. General population (n= 565)									Sample 2. Chronic pain (n= 514)								
	Item reduction based on CTT				Item reduction based on IRT					Item reduction based on CTT				Item reduction based on IRT				
Item	Step 1. Item level analysis		Step 2. Internal consistency		Step 3. Item response					Step 1. Item level analysis		Step 2. Internal consistency		Step 3. Item response				
	M[a]	SD[b]	r_{i-t}[c]	α (if removed)	a[d]	b1	b2	b3	b4	M[e]	SD[f]	r_{i-t}[g]	α (if removed)	a[h]	b1	b2	b3	b4
Demand				.74									.67					
13 [d,h]	1.95	1.32	.53	.70	*0.96*	-0.45	-0.62	0.04	1.63	1.95	1.33	.43	.62	*0.47*	0.57	-1.19	0.28	0.92
17 [e]	3.20	0.87	.52	.70	1.20	-2.00	-2.81	-2.05	0.28	*3.25*	0.96	.41	.62	1.06	-1.54	-2.49	-1.57	-0.33
21 [a,d,e,h]	*3.37*	0.85	.43	.72	*0.76*	-1.02	-1.99	-2.02	-0.39	*3.51*	0.78	.30	.65	*0.73*	-0.96	-2.42	-1.78	-0.91
25 [h]	2.01	1.24	.55	.69	1.11	-1.06	-0.62	-0.03	2.15	2.14	1.29	.49	.59	*0.60*	-0.11	-1.09	0.23	0.78
29	2.90	1.01	.53	.69	1.14	-1.67	-2.18	-1.34	0.82	3.06	1.02	.47	.59	1.22	-1.72	-2.25	-1.67	0.30
33 [a,d,e,h]	*3.48*	0.76	.36	.74	*0.64*	-2.85	-1.44	-1.90	-0.70	*3.55*	0.74	.31	.65	*0.82*	-0.66	-3.25	-1.90	-0.98
Awfulizing				.88									.86					
14	1.72	1.36	.67	.86	1.14	-0.54	0.29	-0.12	1.97	2.07	1.41	.67	.83	1.12	-0.59	-0.22	-0.45	1.30
18 [h]	2.04	1.31	.65	.87	1.07	-0.87	-0.39	-0.10	1.56	2.39	1.38	.62	.84	*0.97*	-0.88	-0.42	-0.65	0.66
22	1.32	1.25	.65	.87	1.17	-0.21	0.44	1.08	2.08	1.53	1.31	.61	.84	1.01	-0.02	-0.23	0.87	1.61
26	2.01	1.27	.72	.86	1.85	-1.83	-0.31	-0.21	2.94	2.26	1.37	.67	.83	1.30	-0.88	-0.50	-0.88	1.54
30	1.92	1.30	.70	.86	1.66	-1.28	-0.32	0.02	2.63	2.19	1.35	.69	.83	1.46	-1.04	-0.73	-0.45	1.68
34	1.54	1.26	.74	.85	2.21	-0.87	-0.05	1.23	4.21	1.58	1.31	.66	.84	1.27	-0.22	-0.23	0.66	2.39
LFT				.84									.81					
3 [d,h]	2.60	1.13	.51	.84	*0.76*	-1.52	-0.60	-1.13	1.16	2.79	1.22	.43	.81	*0.57*	-0.78	-0.33	-1.38	0.32
7 [d,h]	1.62	1.27	.60	.82	*0.95*	-0.57	0.17	0.33	2.05	1.73	1.34	.56	.78	*0.87*	-0.17	-0.24	0.24	1.53
11	2.00	1.28	.70	.80	1.49	-1.64	-0.09	-0.25	2.40	2.22	1.32	.65	.76	1.26	-1.23	-0.51	-0.52	1.52
39	1.97	1.36	.64	.82	1.13	-0.86	-0.17	-0.07	1.43	2.40	1.36	.60	.77	1.02	-1.08	-0.46	-0.62	0.66
43 [h]	1.06	1.16	.59	.83	1.10	0.19	0.76	1.25	2.63	1.31	1.28	.54	.78	*0.84*	0.26	-0.01	1.16	1.30
47	1.65	1.33	.70	.80	1.64	-0.82	0.26	0.30	2.89	1.95	1.44	.62	.77	1.08	-0.57	0.12	-0.34	1.14
Self-downing				.77									.76					
4 [a,d,h]	*0.78*	1.11	.52	.73	*0.88*	1.08	0.90	0.83	2.85	0.89	1.21	.44	.75	*0.56*	1.05	0.42	0.98	0.89
8 [h]	1.35	1.37	.60	.71	1.13	0.16	0.89	0.05	2.21	1.64	1.50	.56	.72	*0.91*	0.11	0.57	-0.28	1.06
12 [h]	1.11	1.28	.58	.71	1.11	0.30	1.09	0.70	2.04	1.40	1.43	.58	.71	*0.98*	0.21	0.84	0.01	1.37
40 [a,d,h]	*0.59*	1.12	.34	.77	*0.43*	1.59	0.95	1.08	-0.34	*0.71*	1.25	.37	.76	*0.43*	1.52	0.97	0.71	-0.36
44 [a]	*0.70*	1.12	.56	.72	1.13	1.22	1.56	1.18	2.01	0.85	1.27	.61	.71	1.11	1.23	1.13	0.88	1.71
48 [a,d,h]	*0.68*	1.10	.47	.74	*0.75*	1.47	1.35	0.61	1.58	1.18	1.34	.49	.74	*0.64*	0.44	0.87	0.17	1.09

Note: CTT= classical test theory; IRT= item response theory; LFT= low frustration tolerance; a= item slope (discrimination) parameter; b= item threshold (difficulty/location) parameter. [a,e]Mean scores< 0.8 or > 3.2; [b,f]SD< 0.67; [c,g]Item-total corrected correlation coefficient < .03; [d,h]discrimination< 1. Cut-offs are as recommended in the study by Jin et al. (2018). Missing rates were less than 0.1% for all items. All transgressions are italicized (eliminated items). Retained items are in bold.

Both the presence of extreme scores (CTT) and poor discrimination (IRT) were considered in the elimination process. The only case when this led to conflict was when choosing between item 21 ("I sometimes need to be relaxed") and item 25 ("I need to be liked by other people") from the demanding scale. Item 21 was negatively skewed in both samples, but its discrimination index was better than that of item 25 in the chronic pain sample, which was clearly below the established cut-off. In fact, if item 25 was selected, the model fit of the four-factor model of the reduced scale in the chronic pain sample worsened significantly (χ^2= 316.19, *df*=48, *p*< .001, RMSEA= .107, 95% RMSEA= [.096, .118], CFI= .946, TLI= .926) when compared to the model with item 21 (χ^2=181.96, *df*=48, *p*< .001, RMSEA= .075, 95% RMSEA= [.064, .087], CFI= .972, TLI= .962). The fit of the four-factor solution with item 21 was also very good in the general population sample (χ^2=141.95, *df*=48, *p*< .001, RMSEA= .059, 95% RMSEA= [.048, .070], CFI= .990, TLI= .986). Thus, we eliminated item 25 and retained item 21. The elimination of the remaining items from the scale was more straightforward and emphasized the inclusion of items with adequate discrimination indices (> 1.0 and the larger the better), non-extreme response patterns (mean scores < 0.8 or > 3.2), and sufficient variability in responses (*SD*> 0.67).

Table 3 shows the results of the CFA with the two samples (general population and chronic pain) and both the original scale (24 irrational items) and the reduced 12-item version. The absolute and incremental fit indices were between good and excellent in the reduced version in both samples (RMSEA< .08 and CFI and TLI> .95). This suggests an adequate fit of the data to the four-factor solution when using our reduced version of the scale. Conversely, the fit indices indicated a poor fit of the data to the four-factor solution when using the original scale with 24 IBs.

Table 3
Results of the Confirmatory Factor Analysis with the original scale and the reduced version (irrational items only) for each of the two study samples

Sample	Scale	χ^2	*p*	RMSEA	95% CI RMSEA	CFI	TLI
General population	Original	2003.75	< .001	.112	.108, .117	.916	.905
	Short version	141.95	< .001	.059	.048, .070	.990	.986
Chronic pain	Original	1808.13	< .001	.114	.109, .119	.865	.849
	Short version	181.96	< .001	.075	.064, .087	.972	.962

Note: RMSEA= Root mean square error of approximation; CFI= Comparative fit index; TLI= Tucker-Lewis index.

The results of the analysis of sources of validity evidence are reported in Table 4. We also indicate the internal consistency estimates for all study variables. The analyses are performed separately for both samples to explore whether the results were consistent across samples. The intercorrelations with the two samples separately evidenced small-to-moderate correlations between the demanding scale and the remaining derivatives (awfulizing, LFT, and self-downing; $.10 \leq r \leq .24$, all $p \leq .05$). The three derivatives (awfulizing, LFT, and self-downing) were between moderately and strongly associated ($.56 \leq r \leq .69$, all $p \leq .001$).

Taking personality, neuroticism was positively linked to irrational thinking ($.21 \leq r \leq .61$, all $p \leq .001$). These associations were stronger for the three derivatives ($.49 \leq r \leq .61$, all $p \leq .001$). Contrary to neuroticism, extraversion and conscientiousness were inversely related to irrational thinking. Again, the associations were stronger for the three derivatives ($-.14 \leq r \leq -.41$, all $p \leq .01$) compared to the demand scale ($< .01 \leq r \leq -.12$). Openness and agreeableness were generally unrelated or very mildly related to IBs.

Finally, IBs, especially the three derivatives, had small-to-moderate negative associations with mental health both in the general population and in persons with chronic pain ($-.17 \leq r \leq -.56$, all $p \leq .001$). The size of the correlation between IBs and mental health was strongest for self-downing ($-.50 \leq r \leq -.57$, all $p \leq .001$) and LFT ($-.46 \leq r \leq -.47$, all $p \leq .001$). Neuroticism had comparable (slightly stronger) negative associations with mental health ($-.56 \leq r \leq -.57$, all $p \leq .001$). Extraversion and conscientiousness were positively associated with mental health status in both samples ($.23 \leq r \leq .29$, all $p \leq .001$).

Different to the previous findings, the association between IBs and physical health status differed across samples. Specifically, IBs were unrelated to physical status in the chronic pain sample. By contrast, LFT and self-downing were associated with poorer physical health in the general population ($-.27 \leq$ r$\leq -.29$, all p$\leq .001$). Similarly, personality was not associated with physical health status in the chronic pain sample. However, neuroticism was associated with poorer physical health in the general population ($r = -.35$, $p < 0{,}001$). The remaining personality factors were more modestly, yet significantly associated with physical health status, either positively (extraversion, openness, and conscientiousness) or negatively (agreeableness).

As a final step, we compared the means in the study variables across the two samples (Table 5). Overall, the results evidenced more irrational thinking in the sample with chronic pain, but the differences were generally small ($d < 0.30$). Differences were also revealed for all the personality dimensions, especially in neuroticism ($t = -6.59$, $p < .001$, $d = 0.46$) and openness ($t = 4.93$, $p < 0{,}001$, $d = 0.34$). Respondents in the general population presented lower neuroticism and higher openness to experience than persons with chronic pain. These differences were between small and moderate in size. In terms of health status, mental well-being was also poorer in persons with chronic pain ($t = 5.57$, $p < .001$, $d = 0.41$). However, the most noticeable difference was in physical health status, where persons with chronic pain indicated that they were much more physical disabled than persons in the general population ($t = 26.16$, $p < .001$, $d = 1.82$).

Table 4
Internal consistency estimates and intercorrelations between study variables (sources of validity evidence)

Variables	Cronbach's α	Pearson intercorrelations (General population above the diagonal / Chronic pain below the diagonal)										
	General/Pain	1	2	3	4	5	6	7	8	9	10	11
Irrational beliefs												
1. Demand	.63 / .57	--	.40***	.44***	.24*	.30***	< .01	.13*	-.01	-.05	.08	-.27***
2. Awfulizing	.86 / .80	.31***	--	.67***	.56***	.49***	-.14**	.08	-.13*	-.21***	.08	-.43***
3. LFT	.80 /.75	.35***	.69***	--	.66***	.61***	-.31***	-.05	-.12*	-.26***	-.27***	-.46***
4. Self-downing	.71 /.72	.10*	.55***	.57***	--	.60***	-.31***	-.08	-.09	-.27***	-.29***	-.56***
Personality												
5. Neuroticism	.86 / .84	.21***	.49***	.55***	.57***	--	-.36***	-.01	-.18***	-.40***	-.35***	-.56***
6. Extraversion	.83 /.82	-.12**	-.25***	-.28***	-.31***	-.40***	--	.28***	.09	.20***	.16**	.24***
7. Openness	.79 /.72	-.01	-.03	-.06	-.03	-.07	.27***	--	-.07	.06	.22***	.04
8. Agreeableness	.68 / .70	.06	-.09*	-.11*	-.12*	-.27***	.25***	.03	--	.15**	-.12*	.11*
9. Conscientiousness	.80 /.81	.08	-.31***	-.32***	-.41***	-.42***	.42***	.08	.26***	--	.12*	.23***
Health-related QoL												
10. PCS	.87-.96 /.75-.90	.05	.08	.09	.07	.08	.03	.09	-.09	.01	--	-.14*
11. MCS	.89-.92 /.79-.92	-.17***	-.42***	-.47***	-.50***	-.57***	.29***	.09	.12**	.29***	-.13**	--

Notes: LFT= Low frustration tolerance; QoL= Quality of life; PCS= Physical composite score; MCS= Mental composite score. Cronbach's α internal consistency coefficients are from the general sample, followed by those from the chronic pain sample. Internal consistency estimates for the PCS correspond to Physical functioning, Role physical, Bodily pain, and General health. Internal consistency estimates for the MCS correspond to Vitality, Social functioning, Role emotional, and Mental health. *$p< .05$; **$p< .01$; ***$p< .001$.

Table 5
Mean differences in study variables

Variables	Range	General sample (*n*= 565)	Chronic pain (*n*= 514)	*t* (*p*)	Cohen's *d*
		M (*SD*)	*M* (*SD*)		
Irrational beliefs					
1. Demand	0-12	9.47 (2.07)	9.81 (2.07)	-2.61 (.009)	0.16
2. Awfulizing	0-12	5.47 (3.39)	6.04 (3.39)	-2.70 (.007)	0.17
3. LFT	0-12	5.62 (3.35)	6.56 (3.38)	-4.54 (< .001)	0.28
4. Self-downing	0-12	3.15 (3.00)	3.87 (3.37)	-3.63 (< .001)	0.23
Personality					
5. Neuroticism	0-48	20.82 (8.67)	24.86 (8.93)	-6.59 (< .001)	0.46
6. Extraversion	0-48	27.86 (7.58)	26.08 (8.00)	3.29 (.001)	0.23
7. Openness	0-48	25.46 (7.37)	23.07 (6.69)	4.93 (< .001)	0.34
8. Agreeableness	0-48	30.97 (5.87)	31.84 (6.15)	-2.08 (.037)	0.15
9. Conscientious-ness	0-48	32.17 (6.56)	31.02 (7.23)	2.38 (.017)	0.17
Health-related QoL					
10. PCS	0-100	46.11 (12.05)	27.55 (7.91)	26.16 (< .001)	1.82
11. MCS	0-100	44.45 (12.49)	39.17 (13.31)	5.57 (< .001)	0.41

Note: LFT= Low frustration tolerance; QoL= Quality of life; PCS= Physical composite score; MCS= Mental composite score.

Discussion

IBs have been consistently associated with psychological (i.e., depressive, anxiety and stress symptoms) and physical (i.e., performance in chronic pain) impairment (Chan & Sun, 2021; Hyland et al., 2014; Morris et al., 2017). Encouragingly, psychological therapies, particularly, cognitive-behavioural interventions, can help challenge IBs and adopt more rational forms of thinking (Turner, 2016). Valid and reliable measures of IBs are necessary to evaluate the baseline status of patients before a treatment is proposed, as well as to monitor individual changes in IBs during treatment. The ABS is a widely-used instrument of IBs (David et al., 2010). However, there are still some concerns regarding its factorial structure, length, discriminant validity, and internal consistency (Artiran & DiGiuseppe, 2020; Hyland et al., 2017; Ruiz-Rodríguez et al., 2020). Thus, the main objective of the present study was to validate a reduced version of the Spanish ABS in two different samples. Following both CTT and IRT, we obtained psychometrically-sound and short version of the ABS with a significant reduction of items (from 24 to 12 IBs). Also importantly, by doing so we clearly improved the fit of the four-factor structure of the ABS (i.e., demandingness, awfulizing, LFT, and self-downing) that is consistent with the REBT theory. In addition, the shortened ABS obtained excellent construct validity results in two different samples.

In the last decades, there has been an increased interest in development and validation of brief scales that can be easier and faster to administer (Kemper et al., 2019). The use of these short measures can reduce the burden of assessment in applied clinical settings and also in research contexts (for both patients and professionals) in terms of time and related costs (Kemper et al., 2019). Abbreviated assessments reduce fatigue and negative reactions in empirical studies and improve participation rates in research studies (Credé et al., 2012; Edwards et al., 2004). In the context of IBs, several authors have tried to reduce the number of items in the ABS (Hyland et al., 2017; Lindner et al., 1999). The present investigation represents a new attempt to obtain a reduced and psychometrically-sound version of the ABS and is the first formal attempt to validate and reduce the scale in Spanish language. As recently recommended by the American Psychological Association (American Psychological Association, 2020), the brief ABS proposed in this work is consistent with theoretical basis of REBT. The four-factor solution of the short Spanish ABS has been tested using a CFA and obtained a very good model fit. In fact, the model fit of the short version of the ABS clearly overtook that of the original full-length one. Ultimately, this means that the short version of the ABS allows the reliable assessment of the four processes described in theoretical REBT model with only 12 irrational items, which makes the reduced ABS a more feasible tool for routine care.

Since the initial development of the ABS, different efforts have been conducted to explore the psychometric properties of the scale. Even in the first full-length version, poor discriminant validity was evidenced (Ruiz-Rodríguez et al., 2020; Suso-Ribera et al., 2016). This may indicate that items do not distinguish people with maladaptive IBs. This may be problematic because psychological test scores are usually used to give feedback to the patients about their psychological status and to make therapeutic recommendations (e.g., to continue with a set of more specific assessments to complete the diagnosis, to begin a psychological therapy, to consult a specialist, etc.; American Psychological Association, 2020). Instruments that do not differentiate between people with problematic and non-problematic IBs may result in inadequate recommendations and, therefore, inadequate management of the patients' suffering. To overcome this issue, items with low discrimination indices have been removed from the scale. In our study, the most problematic issues regarding discrimination indices emerged in the demandingness and the self-downing subscales. Specifically, there were little or no differences between participants with high and low scores at some demandingness (e.g., "I need to be liked by others", "It is essential to be liked by some people", and "It is essential to be relaxed at times") and self-downing items (e.g., "Sometimes, when people do not like me, I think I am a bad person", "If some people do not like me, that means I am a bad person", and "I will never be relaxed"). It has been suggested that recommendations derived from test administration should be fair and minimize bias (American Psychological Association, 2020), so future efforts should be conducted to develop items that do not lead to extreme responses and provide wide variability in responses. In this work, after removing problematic items, the short scale was composed of items with high discrimination indices, non-extreme responses patterns, and variability in responses.

Another shortcoming found in the literature is the low internal consistency reported when using the ABS (Ruiz-Rodríguez et al., 2020; Suso-Ribera et al., 2016). To explore the internal consistency of our reduced version, both inter-item correlations (reliability at the subscale level) and intercorrelations between subscales (reliability at the questionnaire level) were computed. First, to explore intern-item correlations, Cronbach's alphas were calculated for each irrational belief subscale (demandingness, awfulizing, LFT, and self-downing). Our results indicated acceptable to good internal consistency estimates in both samples. The demandingness scale obtained the lowest internal consistency values. These results were also supported by the intercorrelations because the demandingness scale obtained the most modest associations (small-moderate correlations) with the remaining IBs. According to REBT, the irrational profile may be homogeneous (i.e., general irrationality) or heterogeneous in content (i.e., irrational just in some domains; David, 2003). Our results support this idea. Demandingness appears to be a heterogeneous subscale that includes a large variety of contents. Therefore, people may manifest approval ("I must be loved by others"), but not comfort needs ("I must have a pleasant, comfortable life most of the time"), which compromises the internal consistency of the subscale. Contrary to this, people with high scores in awfulizing, LFT, and self-downing seem to score high in all items regardless of the content (approval, perfectionism, and comfort), further supporting the homogeneity of the three inferences (Duru & Balkıs, 2021). Future studies should test whether this is due to the theoretical differentiation between inferences and demands or if items in the demandingness subscale are not adequately formulated and should be rewritten.

In addition to the inter-item correlations discussed in the previous lines, internal consistency results (at the questionnaire level) were also investigated. Specifically, the intercorrelations between ABS subscales showed that the three irrational inferences (awfulizing, LFT, and self-downing) were moderately-to-strongly associated. These results provide useful information and confirm that, contrary to the full length version of the ABS (Suso-Ribera et al., 2016), the shortened version has adequate internal consistency as evidenced by correlations between the subscales correlated in the expected direction of irrationality.

To test the construct validity of the ABS, bivariate associations with related constructs (i.e., personality and health) were calculated. Consistent with past research (Samar et al., 2013; Suso-Ribera, Martínez-Borba, et al., 2019), our results showed that people with more engrained IBs also present higher neuroticism and lower extraversion and conscientiousness scores. The stronger associations were found with neuroticism. Also congruent with past studies (Samar et al., 2013; Suso-Ribera, Martínez-Borba, et al., 2019), associations between IBs and openness and agreeableness were negligible in our study. It has been suggested that personality traits may be useful in clinical contexts as they can help professionals to conceptualize therapy cases and propose an adequate intervention (Osma et al., 2021; Samar et al., 2013). It is also reasonable to think that instruments that allow the assessment of patients' progress throughout the psychological interventions are needed. While acknowledging the role of neuroticism on wellbeing, especially in clinical practice it is also necessary to assess constructs that are more modifiable and more proximal to the final behavior, such as IBs. For instance, IBs have been linked to health behaviors in the context of the COVID-19 pandemic (Teovanović et al., 2021). Previous studies also revealed that IBs contribute incremental validity over personality in the prediction of life satisfaction (Spörrle et al., 2010). Future efforts should be conducted to further confirm the mediator role of IBs in psychological interventions and check whether changes in neuroticism during cognitive-behavioral therapy (Osma et al., 2021) are partly or totally explained by modifications in IBs.

IBs are at the core of cognitive-behavioral interventions (David et al., 2010). Therefore, it is crucial to have a short, theory-consistent scale that is easily applied in clinical practice. Our study findings are consistent with the results reported by previous authors (Chan & Sun, 2021; Hyland et al., 2014), where a relationship between IBs and poorer mental health (in both samples) is revealed. Interestingly, differences across samples emerged in the relation between IBs and physical health. In people with chronic pain, the relationship between IBs and physical health was not significant, whereas, in the general population, irrational thinking was associated with worse physical health. This difference may be explained by the fact that, in people with good health, where pain is likely to be less intense, IBs may have more room to influence functionality. On the contrary, when the health status is less favorable and people experience intense pain (sample of people with chronic pain), there might be less room for psychological variables (i.e., demandingness, awfulizing, LFT, and self-downing) to play a role on physical functioning above and beyond pain status (Suso-Ribera et al., 2017). This leads us to think that, if we want to improve the functional capacity of individuals, especially healthy persons (i.e., prevention), awfulizing, LFT, and self-downing, may be sensible targets. A clear example of this is the application of REBT in athletes (Tingaz, 2020), where therapists are challenging IBs to increase performance and wellbeing (Wood et al., 2018).

An important strength of the study was the inclusion of 2 samples. Both had an optimal distribution with respect to sex, which corresponds to the expected distribution [both at the population level (Instituto Nacional de Estadística, 2021) and in patients with chronic pain (Racine et al., 2015)]. We also had a broad age distribution, ranging from 18 to 93 years in the general population and from 18 to 89 years in persons with chronic pain, which is favorable for generalizability purposes. Finally, and different to many other validation studies, the samples used in the present investigation did not only comprise university students and highly literate participants (Duru & Balkıs, 2021; Hyland et al., 2017). This inclusion of different working groups (e.g., active workers, retired persons, and unemployed individuals) with a wide variety educational level is favorable for the generalizability of the findings.

An important finding was that the results presented above were replicated in two different samples. After the analyses with CTT and IRT, an encouraging result was that the set of poorly functioning items was generally comparable in the general population and in people with chronic pain. In addition, it is also positive that the model fit of the short version of the Spanish ABS improved in both samples. The American Psychological Association (American Psychological Association, 2020) recently recommended that instruments need to be validated and tested in different samples, especially when instruments are used in populations that may differ from those in which the test was original developed. Replicating our results in two different samples supports the stability and replicability of the factorial structure found in this study. According to our results, the reduced Spanish version of the ABS could be used in a wide range of populations, from individuals of the general population to more specific samples, such as persons with chronic pain.

The results of this study must be understood in the context of certain limitations. For example, due to the cross-sectional nature of the study, we cannot establish causal relationships between variables (e.g., the direction of effects between irrational thinking and mental and physical health or the existence of third explanatory variable). In addition, there is clinical information missing in our samples, such as an exhaustive diagnostic interview to investigate the existence of psychopathology. Therefore, the applicability of the findings for specific populations, such as persons with depression or anxiety, is unclear. A final limitation refers to the reliance on self-report data only. Including more objective information, such as data on physical performance using biosensors or wearable devices (Colombo et al., 2020), would have provided additional evidence about the validity of the shortened Spanish version of the ABS.

To conclude, the development of a short and psychometrically-sound measure of IBs in Spanish was necessary both for clinical and research purposes. For clinical practice, routine outcome monitoring during the whole therapeutic process (e.g., between sessions) is highly recommended, especially for not-on-track patients (i.e., patients who do not easily respond to the planned treatment; Gual-Montolio et al., 2020). In research settings, the use of short measures is also preferable to facilitate ecological momentary assessment, which consists of the evaluation of individuals in a naturalistic setting and in real-time, instead of using paper-and-pencil retrospective self-reports (Colombo et al., 2020). The use of the Spanish short version of the ABS will therefore facilitate the repeated assessment of IBs in clinical and research settings, while keeping the assessment burden in the patients at a minimum. To reduce the evaluator burden as well, researchers are increasingly incorporating mobile applications for repeated monitoring (Colombo et al., 2020). Because this intensive assessment also requires short evaluations to reduce participant burden due its recurrent nature, the present short Spanish version of the ABS might also be beneficial for professionals interested in evaluating IBs via mobile applications.

References

Alonso, J., Regidor, E., Barrio, G., Prieto, L., Rodrigues, C., de la Fuente, L., Rodriguez, C., de la Fuente, L., Rodrigues, C., & de la Fuente, L. (1998). Valores poblacionales de referencia de la versión española del Cuestionario de salud SF-36 [Population-based reference values for the Spanish version of the Health Survey SF-36]. *Medicina Clínica*, *111*(11), 410-416.

American Psychological Association. (2020). *APA GUIDELINES for Psychological Assessment and Evaluation*. Washington, DC: Author.

Artiran, M., & DiGiuseppe, R. (2020). A Turkish translation of a measure of irrational and rational beliefs: Reliability, validity studies and confirmation of the four cognitive processes model. *Journal of Rational - Emotive and Cognitive - Behavior Therapy*, *38*(3), 369-398. doi: 10.1007/s10942-020-00340-9

Belloch, A., Morillo, C., Luciano, J., García-Soriano, G., Cabedo, E., & Carrió, C. (2010). Dysfunctional belief domains related to obsessive-compulsive disorder: A further examination of their dimensionality and specificity. *The Spanish Journal of Psychology*, *13*(1), 376-388. doi: 10.1017/S1138741600003930

Burgess, P. (1986). *Belief systems and emotional disturbance: Evaluation of the rational emotive model* [Unpublished doctoral dissertation]. University of Melbourne, Australia.

Burgess, P. M. (1990). Toward resolution of conceptual issues in the assessment of belief systems in rational-emotive therapy. *Journal of Cognitive Psychotherapy: An International Quarterly*, *4*(2), 171-184.

Caballo, V. E., Lega, L. I., & González, S. (1996, November). *A factor analysis of a back translated Spanish version of the Scale of Attitudes and Beliefs (SAB)*. [Paper presentation]. 30th Annual Convention of the Association for Advancement of Behavior Therapy, New York, NY, United States.

Chan, H. W. Q., & Sun, C. F. R. (2021). Irrational beliefs, depression, anxiety, and stress among university students in Hong Kong. *Journal of American College Health*, *69*(8), 827-841. doi: 10.1080/07448481.2019.1710516

Colombo, D., Suso-Ribera, C., Fernández-Álvarez, J., Cipresso, P., García-Palacios, A., Riva, G., & Botella, C. (2020). Technology-based ecological momentary assessment in clinical psychology. *Reference Module in Neuroscience and Biobehavioral Psychology*, 1-15. doi: 10.1016/b978-0-12-818697-8.00003-0

Costa, P. T., & McCrae, R. R. (1992). *Revised NEO Personality Inventory (NEO-PI-R) and NEO Five-Factor Inventory (NEO-FFI) professional manual*. Odessa, Ucraine: Psychological Assessment Resources.

Credé, M., Harms, P., Niehorster, S., & Gaye-Valentine, A. (2012). An evaluation of the consequences of using short measures of the big five personality traits. *Journal of Personality and Social Psychology*, *102*(4), 874-888. doi: 10.1037/a0027403

David, D. (2003). Rational Emotive Behavior Therapy (REBT): The view of a cognitive psychologis. In W. Dryden (Ed.), *Rational emotive behaviour therapy: Theoretical developments* (pp. 130-159). New York, NY: Brunner-Routledge.

David, D., Lynn, S. J., & Ellis, A. (2010). *Rational and irrational beliefs*. New York, NY: Oxford University Press.

Demaria, T. P., Kassinove, H., & Dill, C. A. (1989). Psychometric properties of the Survey of Personal Beliefs: A rational-emotive measure of irrational thinking. *Journal of Personality Assessment*, *53*(2), 329-341. doi: 10.1207/s15327752jpa5302_10

Diamantopoulos, A., Sarstedt, M., Fuchs, C., Wilczynski, P., & Kaiser, S. (2012). Guidelines for choosing between multi-item and single-item scales for construct measurement: A predictive validity perspective. *Journal of the Academy of Marketing Science*, *40*(3), 434-449. doi: 10.1007/s11747-011-0300-3

DiGiuseppe, R., Leaf, R., Exner, T., & Robin, M. V. (1988, September). *The development of a measure of rational/irrational thinking* [Paper presentation]. World Congress of Behavior, Edinburg, Scotland.

Duru, E., & Balkıs, M. (2021). Childhood trauma, depressive symptoms and rational /irrational beliefs: A moderated mediation model. *Current Psychology*, *1*(2010). doi: 10.1007/s12144-020-01282-2

Edwards, P., Roberts, I., Sandercock, P., & Frost, C. (2004). Follow-up by mail in clinical trials: Does questionnaire length matter? *Controlled Clinical Trials*, *25*(1), 31-52. doi: 10.1016/j.cct.2003.08.013

Ellis, A. (1995). Changing rational-emotive therapy (RET) to rational emotive behavior therapy (REBT). *Journal of Rational-Emotive & Cognitive-Behavior Therapy*, *13*(2), 85-89. doi: 10.1007/BF02354453

Froman, R. D. (2001). Elements to consider in planning the use of factor analysis elements to consider in planning the use of factor analysis. *Southern Online Journal of Nursing Research*, *2*(5), 1-22.

Goetz, C., Coste, J., Lemetayer, F., Rat, A. C., Montel, S., Recchia, S., Debouverie, M., Pouchot, J., Spitz, E., & Guillemin, F. (2013). Item reduction based on rigorous methodological guidelines is necessary to maintain validity when shortening composite measurement scales. *Journal of Clinical Epidemiology*, *66*(7), 710-718. doi: 10.1016/j.jclinepi.2012.12.015

Gual-Montolio, P., Martínez-Borba, V., Bretón-López, J. M., Osma, J., & Suso-Ribera, C. (2020). How are information and communication technologies supporting routine outcome monitoring and measurement-based care in psychotherapy? A systematic review. *International Journal of Environmental Research and Public Health*, *17*(9), 3170. doi: 10.3390/ijerph17093170

Holmes, W., & Bolin, J. (2017). Brief review of latent variable modeling in Mplus. In *Multilevel modeling using Mplus* (pp. 171-233). Boca Raton, FL: Chapman and Hall/CRC.

Hyland, P., Fox, R., Treacy, G., Maguire, P., Boduszek, D., & Vallières, F. (2017). Assessing the factorial validity of the Attitudes and Belief Scale 2-Abbreviated Version: A call for the development a gold standard method of measuring rational and irrational beliefs. *Journal of Rational-Emotive & Cognitive-Behavior Therapy*, *35*(2), 111-124. doi: 10.1007/s10942-016-0243-1

Hyland, P., Shevlin, M., Adamson, G., & Boduszek, D. (2014). Modeling the structure of the Attitudes and Belief Scale 2 using CFA and bifactor approaches: Toward the development of an abbreviated version. *Cognitive Behaviour Therapy*, *43*(1), 60-71. doi: 10.1080/16506073.2013.777467

Instituto Nacional de Estadística. (2021). *Estadística del padrón continuo 2021 [Statistics of the Spanish Census 2021]*. https://www.ine.es/dyngs/INEbase/es/operacion.htm?c=Estadistica_C&cid=1254736177012&menu=ultiDatos&idp=1254734710990

Jin, X., Liu, G. G., Gerstein, H. C., Levine, M. A. H., Steeves, K., Guan, H., Li, H., & Xie, F. (2018). Item reduction and validation of the Chinese version of Diabetes Quality-of-Life measure (DQOL). *Health and Quality of Life Outcomes*, *16*(1), 1-11. doi: 10.1186/s12955-018-0905-z

Kemper, C. J., Trapp, S., Kathmann, N., Samuel, D. B., & Ziegler, M. (2019). Short versus long scales in clinical assessment: Exploring the trade-off between resources saved and psychometric quality lost using two measures of obsessive-compulsive symptoms. *Assessment*, *26*(5), 767-782. doi: 10.1177/1073191118810057

Lega, L. I., & Ellis, A. (2001). Rational emotive behavior therapy (REBT) in the new millenium: A cross-cultural approach. *Journal of Rational-Emotive & Cognitive-Behavior Therapy*, *19*(4), 201-222. doi: 10.1023/A:1012537814117

Lindner, H., Kirkby, R., Wetheim, E., & Birch, P. (1999). A brief assessment of irrational thinking: The Shortened General Attitude and Belief Scale. *Cognitive Therapy and Research*, *23*(6), 651-663.

Miró, J., Queral, R., & Nolla, M. (2014). Pain-related attitudes and functioning in elderly primary care patients. *The Spanish Journal of Psychology*, 17, E104. doi: 10.1017/sjp.2014.103

Morris, R., Tod, D., & Eubank, M. (2017). "It's the end of the world as we know it (and I feel fine)": The use of rational emotive behavior therapy (REBT) to increase function and reduce irrational beliefs of an injured athlete. *Rational Emotive Behavior Therapy in Sport and Exercise*, 220-229. doi: 10.4324/9781315541808

Osma, J., Peris-Baquero, O., Suso-Ribera, C., Sauer-Zavala, S., & Barlow, D. H. (2021). Predicting and moderating the response to the unified protocol: Do baseline personality and affective profiles matter? *Cognitive Therapy and Research,* 45(4), 817-830.

Owings, L. R., Thorpe, G. L., Mcmillan, E. S., Burrows, R. D., Sigmon, S. T., & Alley, D. C. (2013). Scaling irrational beliefs in the General Attitude and Belief Scale: An analysis using item response theory methodology. *SAGE Open*, *3*(2), 1-10. doi: 10.1177/2158244013484156

Pérez-Acosta, A. M., Rodríguez-Sánchez, A., & Vargas, M. C. (2008). Albert Ellis (1913-2007): Pionero de los modelos mediacionales de intervención [Albert Ellis (1913-2007): Pioneer of mediational models of intervention]. *Behavioral Psychology/Psicología Conductual*, *16*(2), 341-348.

Racine, M., Castarlenas, E., de la Vega, R., Tome-Pires, C., Sole, E., Miro, J., Jensen, M. P., Moulin, D. E., & Nielson, W. R. (2015). Sex differences in psychological response to pain in patients with fibromyalgia syndrome. *Clinical Journal of Pain*, *31*(5), 425-432. doi: 10.1097/ajp.0000000000000132

Raubenheimer, J. (2004). An item selection procedure to maximise scale reliability and validity. *SA Journal of Industrial Psychology*, *30*(4), 59-64. doi: 10.4102/sajip.v30i4.168

Rovira, M., Lega, L., Suso-Ribera, C. & Orue, I. (2020). Creencias irracionales en mujeres españolas sobre el rol femenino tradicional [Irrational beliefs in Spanish women about the traditional female role]. *Behavioral Psychology/Psicología Conductual, 28*(2), 265-291.

Ruiz-Rodríguez, J., Suso-Ribera, C., & Sorribes-Vall, F. (2020). Development and psychometric properties of a new measure of irrational thinking: The Scale of Irrational Contents and Styles-Basics (SICS-B). *Journal of Rational - Emotive and Cognitive - Behavior Therapy*, *38*(4), 498-521. doi: 10.1007/s10942-020-00348-1

Samar, S. M., Walton, K. E., & McDermut, W. (2013). Personality traits predict irrational beliefs. *Journal of Rational - Emotive and Cognitive - Behavior Therapy*, *31*(4), 231-242. doi: 10.1007/s10942-013-0172-1

Schreiber, J. B., Nora, A., Stage, F. K., Barlow, E. A., & King, J. (2006). Reporting structural equation modeling and confirmatory factor analysis results: A review. *The Journal of Educational Research*, *99*(6), 323-338. doi: 10.3200/JOER.99.6.323-338

Solé i Fontova, M. D. (2006). *Validació i estandarització espanyola del NEO-PI-R, NEO-FFI,NEO-FFI-R i escales de Schinka, en mostres universitèries i població general* [Validation and Spanish standardization of the NEO-PI-R, NEO-FFI, NEO-FFI-R and Schinka scales in university samples and the general population]. [Doctoral dissertation, University of Lérida, Spain]. http://www.tdx.cat/handle/10803/8294

Spörrle, M., Strobel, M., & Tumasjan, A. (2010). On the incremental validity of irrational beliefs to predict subjective well-being while controlling for personality factors. *Psicothema*, *22*(4), 543-548.

Streiner, D. L., Norman, G. R., & Cairney, J. (2015). *Health Measurement Scales* (Vol. 1). New York, NY: Oxford University Press.

Suso-Ribera, C., Camacho-Guerrero, L., Osma, J., Suso-Vergara, S., & Gallardo-Pujol, D. (2019). A reduction in pain intensity is more strongly associated with improved physical functioning in frustration tolerant individuals: A longitudinal moderation study in chronic pain patients. *Frontiers in Psychology*, *10*(APR), 1-12. doi: 10.3389/fpsyg.2019.00907

Suso-Ribera, C., García-Palacios, A., Botella, C., & Ribera-Canudas, M. V. (2017). Pain catastrophizing and its relationship with health outcomes: Does pain intensity matter? *Pain Research and Management*, *2017*, 1-8. doi: 10.1155/2017/9762864

Suso-Ribera, C., Jornet-Gibert, M., Ribera Canudas, M. V., McCracken, L. M., Maydeu-Olivares, A., & Gallardo-Pujol, D. (2016). There's more than catastrophizing in chronic pain: Low frustration tolerance and self-downing also predict mental health in chronic pain patients. *Journal of Clinical Psychology in Medical Settings*, *23*(2), 192-206. doi: 10.1007/s10880-016-9454-y

Suso-Ribera, C., Martínez-Borba, V., Martín-Brufau, R., Suso-Vergara, S., & García-Palacios, A. (2019). Individual differences and health in chronic pain: Are sex-differences relevant? *Health and Quality of Life Outcomes*, *17*(1), 128. doi: 10.1186/s12955-019-1182-1

Szentagotai, A., David, D., Lupu, V., & Cosman, D. (2008). Rational emotive behavior therapy versus cognitive therapy versus pharmacotherapy in the treatment of major depressive disorder: Mechanisms of change analysis. *Psychotherapy*, *45*(4), 523-538. doi: 10.1037/a0014332

Teovanović, P., Lukić, P., Zupan, Z., Lazić, A., Ninković, M., & Žeželj, I. (2021). Irrational beliefs differentially predict adherence to guidelines and pseudoscientific practices during the COVID-19 pandemic. *Applied Cognitive Psychology*, *35*(2), 486-496. doi: 10.1002/acp.3770

Tingaz, E. O. (2020). The mediating role of mindfulness in the relationship between the cognitive flexibility and irrational performance beliefs of university student-athletes. *Current Psychology*, *39*(4), 1208-1214. doi: 10.1007/s12144-020-00891-1

Toland, M. D. (2014). Practical guide to conducting an item response theory analysis. *Journal of Early Adolescence, 34*(1), 120-151. doi: 10.1177/0272431613511332

Turner, M. J. (2016). Rational emotive behavior therapy (REBT), irrational and rational beliefs, and the mental health of athletes. *Frontiers in Psychology*, *7*(SEP), 1-16. doi: 10.3389/fpsyg.2016.01423

Vîslă, A., Flückiger, C., Grosse Holtforth, M., & David, D. (2016). Irrational beliefs and psychological distress: A meta-analysis. *Psychotherapy and Psychosomatics*, *85*(1), 8-15. doi: 10.1159/000441231

Ware, J., & Sherbourne, C. (1992). The MOS 36-item Short-Form Health Survey (SF-36). I. Conceptual framework and item selection. *Medical Care*, *30*(6), 473-483.

Wood, A. G., Barker, | J B, Turner, | M J, & Sheffield, | D. (2018). Examining the effects of rational emotive behavior therapy on performance outcomes in elite paralympic athletes. *Scandinavian Journal of Medicine and Science in Sports*, *28*(1), 329-339. doi: 10.1111/sms.12926

RECEIVED: August 12, 2021
ACCEPTED: October 28, 2021