Cross-cultural measurement invariance of the purpose in life test - Short form (PIL-SF) in seven Latin American countries

Tomás Caycho-Rodríguez1 · Lindsey W. Vilca2 · Mauricio Cervigni3,4,5,6 · Miguel Gallegos7,8,3,6 · Pablo Martino3,4,5,6 · Manuel Calandra3,4,5 · Cesar Armando Rey Anaconá9 · Claudio López-Calle10 · Rodrigo Moreta-Herrera11 · Edgardo René Chacón-Andrade12 · Marlon Elías Lobos-Rivera12 · Perla del Carpio13 · Yazmín Quintero13 · Erika Robles14 · Macerlo Panza Lombardo15 · Olivia Gamarra Recalde16 · Andrés Buschiazzo Figares17 · Michael White18 · Carmen Burgos Videla19 · Carlos Carbajal-León20

Accepted: 7 July 2022 / Published online: 28 July 2022
© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2022

Abstract
The aim was to test the cross-cultural measurement invariance of the PIL-SF in a sample of people from seven Latin American countries. Additionally, the characteristics of the PIL-SF items were evaluated and to assess the relationship between purpose in life, as measured by the PIL-SF, and fear of COVID-19. A total of 4306 people from seven Latin American countries participated in the study. The results indicated that the PIL-SF is invariant in the seven participating countries and, therefore, there is evidence that the items reflect the purpose of life in the same way in all countries. This allows comparisons of purpose in life between countries that are free of bias, reflecting the true differences in how countries respond to items. From IRT, the discrimination parameters are adequate and indicate that the items cover a wide range of the purpose in life construct. The difficulty parameters are adequate and increase monotonically. This indicates that people would need a higher level of purpose in life to respond to the higher response categories. Thus, the PIL-SF items would be useful in determining people with a relatively high degree of purpose in life. Identifying people with different levels of purpose in life would allow them to be part of intervention programs, either to support those with low levels or to maintain and reinforce their purpose in life. The evidence of cross-country measurement invariance of the PIL-SF provides a measure to be used in cross-cultural studies about the meaning of life.

Keywords Measurement invariance · Latin America · Purpose in life, cross-cultural research/comparison · Item response theory

Introduction
In recent years, empirical studies of the meaning of life construct have increased as it has become an important concept within the theories of well-being especially driven by the development of Positive Psychology (García-Alandete, 2014; García-Alandete et al., 2017). This provides a better understanding of the potentialities and resources people have (Crea, 2016). For Frankl (2014), experiencing the meaning of life is the most important motivational characteristic for human beings. There is evidence that the meaning of life is related to a greater experience of freedom, satisfaction with life, optimism, hope and well-being (Akbari et al., 2019; Karataş et al., 2021; Stoyles et al., 2015; Yalçın & Malkoç, 2015). In addition, it has a protective role against depression, hopelessness, aimlessness, self-injurious behaviors and suicide, among other mental health problems (Disabato et al., 2017; George & Park, 2016; Lew et al., 2020; Marco et al., 2017).

Meaning in life has been defined in many ways, such as an experience of fulfillment in life, sense of purpose, authentic life, etc. (Wong, 2016). Thus, it has been suggested that meaning in life comprises both meaning and purpose, which are closely related constructs that can be used interchangeably (García-Alandete et al., 2019). However, others indicate that value and importance are synonymous with meaning; while goals and intentions are synonymous with purpose (Steger et al., 2006). From other theoretical perspectives,
the cognitive impulses that give meaning to personal experiences are important (Heintzelman & King, 2014a; Proulx & Inzlicht, 2012; Waytz et al., 2015). Similarly, importance is given to the emotional aspects of the meaning of life (Baumeister et al., 2013; King et al., 2016). Others suggest that meaning is part of a system of social relations (Lambert et al., 2013; Klein, 2017). Finally, there are studies that consider the meaning of life as a permanent search process or that try to identify the levels of meaning (McAdams, 2013; Schnell, 2009).

The difficulty to define the meaning of life also generates problems to evaluate it in a valid and reliable way. In logotherapy, there is an empirical orientation that has led to the development of instruments to quantitatively evaluate the meaning of life (Baththyany & Guttman, 2006; Schulenberg, 2004). From this approach, the Purpose in Life Test was developed with the objective of assessing meaning and purpose in life (PIL; Crumbaugh & Maholick, 1964). The PIL is probably the most widely used and researched instrument based on logotherapy (Schulenberg & Melton, 2010). According to logotherapy, people have the ability to resist the impact of external situations and their own limitations, both physical and psychological. This would allow them to develop meaning in their own lives whenever they are able to reflect on themselves (Frankl, 2014).

In its original form, the PIL is comprised of 20 items that assess a set of life experiences, such as the presence of life goals, satisfaction with daily activities, boredom and enthusiasm, among others (Shuv-Ami & Bareket-Bojmel, 2021). Numerous studies have reported that the original 20-item PIL presents evidence of validity and reliability in different populations and countries (e.g., Brunelli et al., 2012; García-Alandete et al., 2016; Haugan & Moksnes, 2013; Ortiz et al., 2012; Schulenberg & Melton, 2010; Simkin et al., 2018; Zhang et al., 2021).

However, there are certain concerns regarding the PIL, such as the criticism related to its dimensionality. Different studies question whether the PIL has a unidimensional structure (e.g., Marsh et al., 2003; Simkin et al., 2018), two dimensions (e.g., García-Alandete et al., 2013; Hayashi & Esmerelles, 2017; Morgan & Farsides, 2009), three dimensions (e.g., Armas et al., 2018; Gottfried, 2016; Magaña Valdadeas et al., 2004) or even up to six factors (Reker & Cousins, 1979). It has even been proposed that both two-factor and five-factor solutions work adequately (Shek, 1988). On the other hand, it has been indicated that the PIL format is cumbersome and bulky at the time of evaluation, even more so if other tests are administered in addition (Harlow et al., 1987). This is largely due to the fact that each of the PIL items has different response scales; for example, item 5 (“Every day is:” (“Cada día es:”)) has response options ranging from “exactly the same” (“exactamente lo mismo”) to “always new and different” (“siempre nuevo y diferente”); while, in item 8 (“In terms of reaching my life goals, I:” (“En términos de alcanzar mis metas de vida, yo:”)), the response options are “haven’t made any progress” (“no he hecho ningún progreso”) to “have achieved all of them completely” (“los he logrado todos por completo”) (Schulenberg et al., 2011). This can lead to confusion and difficulty in understanding the test in a unitary way (Francis et al., 2019).

There is also concern about the extreme nature of some items, which refer to suicidal thoughts or preparation for death, and the influence it may receive from certain momentary emotions, such as boredom and enthusiasm (Heintzelman & King, 2014b). It has been suggested that the content of some items may be mismeasuring other constructs, such as depression (Schulenberg et al., 2011). In addition, some items, such as those referring to life goals, or reason for existence, may be too abstract for some age groups, such as adolescents (Law, 2012). In this sense, for example, a factor structure has been proposed without the presence of three items that were not representative and did not load on any factor (Jonsén et al., 2010). This has led to the suggestion of shorter versions of the PIL by eliminating some items. Thus, there are, for example, versions of 17 items (Jonsén et al., 2010) and 19 items divided into three factors (Halama, 2009), 10 items divided into two factors (García-Alandete et al., 2013; García-Alandete, 2014; Hayashi & Esmerelles, 2017) and even a 4-item version (Schulenberg et al., 2011).

Specifically, the 4-item version of the PIL, called Purpose in Life Test - Short Form (PIL-SF; Schulenberg et al., 2011) is made up of items 3, 4, 8 and 20 of the original version. For the development of this version, items 3, 8 and 20, which make up the factor called “purposeful life” as it appeared in some studies were joined with item 4 that has specific content on meaning in life. This improved the reliability of the factor from .75 to .81 (Schulenberg et al., 2011). In addition, the 4 items have formed the same factor in different studies, which assess the presence and realization of life goals/purposes (Molcar & Stuempfig, 1988; Steger, 2006). The original psychometric study of the PIL-SF (Schulenberg et al., 2011) showed that the 4-item unidimensional model has an excellent fit and produces reliable scores. In addition, PIL-SF scores showed expected and significant correlations with other measures of life purpose and meaning, search for noetic goals or goals to find meaning in life (from logotherapy, the term noetic can be interpreted as meaning, Crumbaugh, 1977), life satisfaction, and boredom proneness. Likewise, the PIL-SF significantly predicted psychological distress. The results were similar whether the PIL-SF was administered independently or in conjunction with the original 20-item PIL. Subsequent studies have also reported an adequate fit of the unidimensional model and good reliability in Spain (García-Alandete et al., 2017) and Denmark (Pacak-Vedel et al., 2021) using confirmatory factor analysis models. From Rasch analysis, support has also been
Another study, which included countries in the Americas, indicated that the search for meaning in life may differ between cultures. Thus, the search for meaning was negatively related to happiness in American college students, but the two variables were not related in the Japanese sample. Therefore, that cultural differences may have less impact on how presence and meaning-seeking relate to well-being. This leads to the need for further studies that seek to clarify the way in which presence of, and the search for, meaning in life are experienced in different cultures. Recently, there is a growing interest in conducting studies in different countries (Boer et al., 2018; Fischer & Karl, 2019). However, this type of research poses different challenges, especially with the measurement instruments used and, specifically, with the lack of research assessing measurement invariance across countries (Matos et al., 2021). Nowadays, it is increasingly necessary to develop and
adapt measurement instruments that can be used in different countries. In view of this, although the PIL-SF has been used in different countries, it is not clear whether it shows measurement invariance in Spanish-speaking Latin American countries. Therefore, the main objective of this study was to test the invariance of the PIL-SF measurement in a sample of people from 7 Latin American countries. Additionally, the characteristics of the PIL-SF items were evaluated using an Item Response Theory (IRT) model. It is expected that the unidimensional structure of the PIL-SF will also be present in this group of Latin American countries and will present an adequate reliability estimate, as previously reported (García-Alandete et al., 2017; Pacak-Vedel et al., 2021). Likewise, although no measurement invariance studies of the PIL-SF have been carried out among different countries, it would be expected that the measure is invariant, considering that the unidimensional structure has shown good fit in previous psychometric studies. Finally, it is expected that the PIL-SF items will present good discrimination and difficulty parameters as has been previously observed (Peter et al., 2016).

Testing measurement invariance is important as it could then assess the applicability of the purpose in life construct in different countries (van de Vijver, 2013). This is valuable as people may have different interpretations of certain words or the entire scale due to variations in cultural characteristics (Veenhoven, 1996). In this sense, the instrument may operate differently and the underlying construct may have different theoretical structures in various countries, leading to biased estimates (Dimitrov, 2010). If the underlying factor structure of an instrument developed in one country and language is the same in different countries with different languages, this would ensure that the instrument performs equally across groups and can be useful for bias-free comparisons (Byrne & Watkins, 2003). For this to occur, it must be shown that the underlying construct is measured equally in each country, i.e., demonstrate that the factor structure is invariant when assessed simultaneously across different countries (Borsboom, 2006; Byrne & Watkins, 2003; He & van de Vijver, 2012; Milfont & Fisher, 2010). The absence of invariance would indicate that individuals or groups respond differently to the items of an instrument. This would lead to the inability to make a reasonable comparison between group means (Dimitrov, 2010; van de Schoot et al., 2012). Providing evidence of measurement invariance of the PIL-SF could improve the accuracy of comparative estimation of the meaning people find in life at the Latin American level and better track the progress of interventions in multinational contexts. As mentioned above, to our knowledge, the PIL-SF has not been applied to Latin American samples. Therefore, this would be the first study using the PIL-SF in this region and the first also to examine its measurement invariance in different Latin American countries.

Finally, IRT-based analyses would make it possible to evaluate item parameters, estimate measurement errors, and describe the relationship between the latent trait and the probability that a person would choose a particular response (Thomas, 2011, 2019). In addition, the IRT model provides a graphical representation of item functioning, which gives the opportunity for a deeper analysis, both at the item and person level, using the same metric. While the PIL-SF has previously been analyzed using IRT models in Germany, France and Italy as well as in the United States (Peter et al., 2016), this is the first time that this analysis has been performed with the Spanish version.

Method

Participants

A total of 4306 people from the general population of seven Latin American countries (Argentina, Colombia, Ecuador, El Salvador, Mexico, Paraguay and Uruguay) selected by non-probabilistic purposive sampling participated in the study. The inclusion criteria were: 1) to be of legal age (in all participating countries the age of adulthood is 18 years or older), and 2) to have given informed consent to participate. The number of participants ranged from 244 in Paraguay to 1360 in Argentina. The countries participating in this study were part of the Epidemiological Project on the mental health of the population in a pandemic situation, designed by the Neuroscience Research Center of the School of Psychology at the National University of Rosario in Argentina, together with the School of Health Sciences of the Universidad Privada del Norte in Peru. The countries were selected to try to provide a variety of contextual settings within Latin America. Although an attempt was made to include as many countries as possible, most countries from South America and only one from Central America and the Caribbean participated.

Table 1 shows that Uruguay has the participants with the highest mean age ($M = 41.8; SD = 16.6$ years); while Ecuador has the lowest mean age ($M = 24.6; SD = 7.8$ years). Likewise, the majority of participants, in all countries, were women ($> 64\%$). More than 60% of the participants have completed university studies and have a professional career ($> 50\%$), except in Ecuador (38.5%) and El Salvador (27.8%), where there is a higher proportion of people with unskilled work (45.1% and 37.5%, respectively). Table 1 reports more detailed information on the sociodemographic characteristics of the participants in each country.
Instruments

Sociodemographic Questionnaire An ad hoc questionnaire was created to collect information on some sociodemographic variables such as nationality, age, sex, educational level and occupation.

The Purpose in Life Test-Short Form (PIL-SF; Schulenberg et al., 2011) The PIL-SF is a short version of four items, derived from the original 20-item test, which measures the degree to which people perceive that their lives had meaning. Items 3, 4, 8, and 20 from the Spanish version by Simkin et al. (2018) were used in this study. Each of the PIL-SF items has seven different Likert-type response options. The total score for the PIL-SF ranges from 4 to 28 and is obtained from the sum of the scores for each item. Higher scores indicate meaning and purpose in life.

Procedure

The data was collected between the months of June and September 2020 at the height of the COVID-19 pandemic. The data collection process was the same in each of the participating countries. Argentina initiated and coordinated the Epidemiological Project on the mental health of the population in a pandemic situation. A general coordinator together with a team of researchers carried out the work in each participating country. An online survey was used, elaborated on the Google Forms platform, with had three distinct parts. The first consisted of the objective of the study and informed consent. If the participants gave their consent to participate in the study, they could access the second (sociodemographic questionnaire) and third (PIL-SF questions, FCV-19S and other measures that were not taken into account in this study) parts of the questionnaire. The survey was distributed by email and social networks, seeking to reach the largest number of people. The data collected were confidential and participants were free to stop answering the questions at any time. The project was reviewed and approved by the Ethics Committee of the Universidad Privada del Norte in Peru (registration number: 20213002).

Data Analysis

Descriptive Analysis

Mean (M) and standard deviation (SD) were calculated, as well as skewness (g1) and kurtosis (g2). Values within the range of ±2, for g1, and ±7 for g2 are considered adequate (Finney & DiStefano, 2013).

Confirmatory Factor Analysis

First, the Confirmatory Factor Analysis (CFA) was performed with the Diagonally Weighted Least Squares with Mean and Variance Corrected (WLSMV) estimator. This estimator was chosen since the PIL-SF items had an ordinal nature (Brown, 2015). The chi-square test (χ²; absolute fit test), Root Mean Square Error of Approximation (RMSEA, estimate of the approximation error of the proposed model), Standardized Root Mean Squared Residual (SRMR, estimate of the average size of the residuals between the sample

---

Table 1  Sociodemographic characteristics of the participants from the seven countries

| Sociodemographic data | Argentina (n = 1360) | Colombia (n = 317) | Ecuador (n = 772) | El Salvador (n = 309) | Mexico (n = 904) | Paraguay (n = 244) | Uruguay (n = 400) |
|-----------------------|---------------------|-------------------|------------------|----------------------|-----------------|-------------------|------------------|
| Age (M ± SD)          | 36.4 ± 15.3         | 32.9 ± 12         | 24.6 ± 7.8       | 28.7 ± 8.8           | 34.6 ± 11.6     | 36.9 ± 11.5       | 41.8 ± 12.6      |
| Gender, n (%)         |                    |                   |                  |                      |                 |                   |                   |
| Male                  | 284 (20.9%)         | 81 (25.6%)        | 273 (35.4%)      | 91 (29.4%)           | 267 (29.5%)     | 637 (70.5%)       | 705 (17.5%)      |
| Female                | 1076 (79.1%)        | 236 (74.4%)       | 499 (64.6%)      | 218 (70.6%)          | 237 (20.5%)     | 196 (80.3%)       | 705 (17.5%)      |
| Educational level, n (%) |              |                   |                  |                      |                 |                   |                   |
| Self-taught reading and writing | 1 (0.1%) | 0 (0%)            | 3 (4%)           | 1 (3%)               | 1 (1%)          | 0 (0%)            | 0 (0%)           |
| Basic (< 6 years)     | 0 (0%)              | 0 (0%)            | 0 (0%)           | 0 (0%)               | 1 (1%)          | 3 (1.2%)          | 0 (0%)           |
| Primary (≥ 6 years)   | 3 (0.2%)            | 0 (0%)            | 3 (4%)           | 3 (1%)               | 1 (1%)          | 0 (0%)            | 4 (1%)           |
| Secondary (≥ 9 years) | 294 (21.6%)         | 47 (14.8%)        | 237 (30.7%)      | 50 (16.2%)           | 54 (6%)         | 21 (8.6%)         | 56 (14%)         |
| Higher (diploma/bachelor’s degree) | 1062 (78.1%) | 270 (85.2%)       | 529 (68.5%)      | 255 (82.5%)          | 847 (93.7%)     | 220 (90.2%)       | 340 (85%)        |
| Occupation, n (%)     |                    |                   |                  |                      |                 |                   |                   |
| Unqualified           | 271 (19.9%)         | 59 (18.6%)        | 348 (45.1%)      | 116 (37.5%)          | 123 (13.6%)     | 31 (21.7%)        | 23 (5.8%)        |
| Manual Qualified      | 105 (7.7%)          | 10 (3.2%)         | 71 (9.2%)        | 33 (10.7%)           | 51 (5.6%)       | 11 (4.5%)         | 17 (4.3%)        |
| Qualified non-manual  | 180 (13.2%)         | 33 (10.4%)        | 37 (4.8%)        | 66 (21.4%)           | 68 (7.5%)       | 15 (6.1%)         | 64 (16%)         |
| Professional          | 735 (54%)           | 195 (61.5%)       | 297 (38.5%)      | 86 (27.8%)           | 607 (67.1%)     | 163 (66.8%)       | 264 (66%)        |
| Management            | 69 (5.1%)           | 20 (6.3%)         | 19 (2.5%)        | 8 (2.6%)             | 55 (6.1%)       | 24 (9.8%)         | 32 (8%)          |
and the hypothesized covariance matrices). Comparative Fit Index (CFI) and Tucker-Lewis Index (TLI) were used to evaluate model fit. Both the CFI and TLI are incremental fit indices that compare the fit of a hypothetical model with the fit of a reference model. Values above .95 for the CFI and TLI would indicate a good fit; whereas, values above .90 indicate an acceptable fit. (Schumacker & Lomax, 2015). For the RMSEA and SRMR, values below .05 are indicative of a good fit; while values between .05 and .08 indicate an acceptable fit (Kline, 2015).

Reliability

Reliability was estimated by calculating Cronbach’s alpha (Cronbach, 1951) and omega for categorical items (Green & Yang, 2009). In both cases, values greater than .70 indicate adequate reliability (Viladrich et al., 2017).

Measurement Invariance

Measurement invariance assessment was performed using the approximate measurement invariance (AMI) approach. This method considers that the factorial loads and the intercepts are not identical between the groups and, therefore, the presence of small differences between parameters is acceptable (Byrne & van de Vijver, 2017; Fischer & Karl, 2019; Lomazzi, 2018). Within this approach, multigroup factor analysis alignment was chosen to test for invariance (Asparouhov & Muthén, 2014). First, the fit of a configurational model that did not present restrictions between groups was evaluated. Second, the configurational model was optimized with a component loss function to minimize the invariance between the means of each factor and the variances of the groups (Asparouhov & Muthén, 2014). Invariance tolerance criteria were established for factor loadings ($\lambda = .40$) and intercepts ($\nu = .20$) (Robitzsch, 2020). The alignment power for the parameters was .25 (Fischer & Karl, 2019). The evaluation of the equivalence of the parameters was carried out based on the interpretation of the R2 index, where values close to 1 express greater invariance; while values close to 0 express less invariance (Asparouhov & Muthén, 2014). The presence of up to 25% of non-invariant parameters ($\lambda$ and $\nu$) is adequate to consider the lack of invariance of a scale (Asparouhov & Muthén, 2014).

Analysis Based on Item Response Theory

Analyzes based on Item Response Theory (IRT) were performed with the 2-Parameter Graded Response Model (2-PLM) for ordinal items with three or more response alternatives (GRM, Samejima, 1997) (Hambleton et al., 2010). Discrimination ($a$) and difficulty ($b$) parameters were estimated. The parameter $a$ indicates the variation of the answers of the items according to the level of the trait. The parameter $b$ indicates the amount of trait necessary for the item to be answered in a specific way. Because the PIL-SF has 7 response categories, there are 6 difficulty estimates, one for each threshold. Item and Test Information Curves (IIC and TIC respectively) were calculated.

Statistical Package Used

Statistical analysis was performed in the RStudio environment for R. For the CFA, the “lavaan” package was used (Rosseel, 2012); while the “sirt” (Robitzsch, 2020) and “ltm” (Rizopoulos, 2006) packages were used for the Alignment and GRM methods, respectively.

Results

Descriptive Analysis

Table 2 shows that the mean scores of the PIL-SF items indicate a tendency toward high levels of purpose in life in all country samples. All mean scores were greater than 5, on a scale of 1 to 7. This is especially observed in items 1 (Enthusiasm) and 2 (Excitement in life), which have the highest mean scores in most countries. Skewness and kurtosis values are adequate in most countries ($g_1 < ±2; g_2 < ±7$). This gives evidence of the univariate normality of the data. On the other hand, the items present moderate and high correlations in all countries. Thus, evidence of convergent internal validity is provided. Furthermore, since the correlations are not greater than .80, the items are not redundant measures of the purpose in life construct (internal discriminant validity).

Validity Based on Internal Structure

Table 3 shows that the unidimensional PIL-SF model presents adequate fit indices in all countries, especially in Ecuador (RMSEA = .000 [.000-.056]; CFI = 1.00; TLI = 1.00), Mexico (RMSEA = .000 [.000-.059]; CFI = 1.00; TLI = 1.00) and Paraguay (RMSEA = .000 [.000-.083]; CFI = 1.00; TLI = 1.00). In addition, all items have high factor weights in all countries.

Scale Reliability

As mentioned, the alpha and omega coefficients for categorical items were used to assess the reliability of the PIL-SF. The scale shows adequate reliability indices in all countries ($\alpha = .83-.88; \omega = .84-.87$) (Table 3).
Table 2  Descriptive analysis of the items by nationality of the participants

| Country       | Items | M    | SD   | g1    | g2    | Polychoric correlation matrix |
|---------------|-------|------|------|-------|-------|------------------------------|
|               |       |      |      |       |       | 1   | 2   | 3   | 4   |
| Argentina     | 1     | 5.77 | 1.44 | −1.45 | 1.75  | 1   |     |     |     |
| (n = 1360)    | 2     | 5.79 | 1.54 | −1.44 | 1.45  | .62 | 1   |     |     |
|               | 3     | 5.38 | 1.35 | −1.06 | 1.06  | .54 | .59 | 1   |     |
|               | 4     | 5.73 | 1.27 | −1.25 | 1.71  | .74 | .76 | .63 | 1   |
| Colombia      | 1     | 6.00 | 1.21 | −1.57 | 2.65  | 1   |     |     |     |
| (n = 317)     | 2     | 5.95 | 1.39 | −1.33 | 1.03  | .75 | 1   |     |     |
|               | 3     | 5.36 | 1.32 | −.97  | .90   | .61 | .65 | 1   |     |
|               | 4     | 5.77 | 1.26 | −1.21 | 1.45  | .76 | .77 | .63 | 1   |
| Ecuador       | 1     | 5.93 | 1.24 | −1.27 | 1.27  | 1   |     |     |     |
| (n = 772)     | 2     | 5.82 | 1.39 | −1.22 | .99   | .68 | 1   |     |     |
|               | 3     | 5.33 | 1.29 | −.67  | .26   | .59 | .62 | 1   |     |
|               | 4     | 5.75 | 1.34 | −1.16 | 1.03  | .70 | .71 | .62 | 1   |
| El Salvador   | 1     | 6.33 | 1.14 | −2.16 | 5.12  | 1   |     |     |     |
| (n = 309)     | 2     | 6.21 | 1.27 | −1.75 | 2.57  | .74 | 1   |     |     |
|               | 3     | 5.57 | 1.43 | −.99  | .47   | .66 | .67 | 1   |     |
|               | 4     | 6.20 | 1.22 | −1.83 | 3.38  | .75 | .77 | .71 | 1   |
| Mexico        | 1     | 6.03 | 1.28 | −1.59 | 2.33  | 1   |     |     |     |
| (n = 904)     | 2     | 6.20 | 1.24 | −1.81 | 3.03  | .68 | 1   |     |     |
|               | 3     | 5.59 | 1.29 | −.96  | .76   | .59 | .64 | 1   |     |
|               | 4     | 6.01 | 1.21 | −1.51 | 2.28  | .71 | .74 | .67 | 1   |
| Paraguay      | 1     | 6.02 | 1.31 | −1.84 | 3.66  | 1   |     |     |     |
| (n = 244)     | 2     | 5.98 | 1.35 | −1.28 | .84   | .62 | 1   |     |     |
|               | 3     | 5.61 | 1.29 | −1.07 | 1.42  | .54 | .64 | 1   |     |
|               | 4     | 5.92 | 1.23 | −1.23 | 1.38  | .66 | .75 | .65 | 1   |
| Uruguay       | 1     | 5.86 | 1.32 | −1.61 | 2.72  | 1   |     |     |     |
| (n = 400)     | 2     | 6.10 | 1.41 | −2.04 | 4.12  | .68 | 1   |     |     |
|               | 3     | 5.61 | 1.21 | −1.29 | 2.30  | .56 | .62 | 1   |     |
|               | 4     | 5.89 | 1.18 | −1.27 | 1.69  | .76 | .77 | .69 | 1   |

M Mean, SD Standard Deviation, g1 Skewness, g2 Kurtosis

Table 3  Fit indices, factorial weights and reliability of the models by nationality of the participants

| Country       | Adjustment indexes | Factorial weight | Reliability |
|---------------|--------------------|------------------|-------------|
|               | χ²     | df   | p    | CFI    | TLI    | SRMR | RMSEA [90%CI] | 1   | 2   | 3   | 4   | ω  | α  |
| Argentina     | 17.11  | 2    | .000 | .99   | .99   | .012 | .075 [.045–.109] | .78 | .81 | .69 | .93 | .85 | .85 |
| Colombia      | 2.66   | 2    | .264 | 1.00  | .99   | .007 | .032 [.000–.121] | .85 | .88 | .72 | .88 | .87 | .87 |
| Ecuador       | .86    | 2    | .650 | 1.00  | 1.00  | .003 | .000 [.000–.056] | .82 | .84 | .73 | .85 | .85 | .86 |
| El Salvador   | 1.07   | 2    | .586 | 1.00  | 1.00  | .006 | .000 [.000–.094] | .85 | .86 | .79 | .90 | .87 | .87 |
| Mexico        | 1.37   | 2    | .504 | 1.00  | 1.00  | .005 | .000 [.000–.059] | .79 | .84 | .75 | .88 | .86 | .86 |
| Paraguay      | .50    | 2    | .779 | 1.00  | 1.00  | .005 | .000 [.000–.083] | .74 | .85 | .74 | .89 | .83 | .84 |
| Uruguay       | 2.14   | 2    | .344 | 1.00  | 1.00  | .010 | .013 [.000–.101] | .81 | .84 | .73 | .94 | .88 | .86 |

χ² Chi square, df degrees of freedom, SRMR Standardized Root Mean Square Residual, TLI Tucker-Lewis Index, CFI Comparative Fit Index, RMSEA Root Mean Square Error of Approximation, α Cronbach’s Alpha, ω McDonald’s Omega
Factor Invariance by Country

The Alignment method showed that the factorial structure of the PIL-SF was invariant for the factorial loads ($R^2 = .99$) and the intersections of the items ($R^2 = .99$), as can be seen in Table 4. The finding of a single non-invariant parameter indicates that the low percentage of non-invariant parameters (3.6%). Therefore, these findings demonstrate the metric and scalar invariance of the scale.

Item Response Theory Model

The discrimination parameters of the items have values greater than 1, which indicates adequate discrimination (Table 5) (Zickar et al., 2002). As for the difficulty parameters, all the threshold estimators increased monotonically. Figure 1 allows observing the IIC where it is indicated that item 4 is the most precise to measure the purpose of life. In addition, the TIC would indicate that the PIL-SF is more reliable in a scale range between −3 and .5.

Discussion

The aim of this study was to examine the cross-cultural measurement invariance of the PIL-SF in samples from seven Latin American countries. Rigorous approaches were adopted to assess whether a single latent factor model would explain the covariances between PIL-SF items across countries. First, the unidimensional model of the PIL-SF was tested and shown to fit the data very well across countries. This provides evidence that the PIL-SF measures a single underlying construct. In this sense, the sum of the PIL-SF item scores provides an overall measure of purpose in life. In addition, the factor loadings of the items were high, which provided further support for the internal structure-based validity of the PIL-SF. Thus, it appears that the PIL-SF is a unidimensional and brief instrument for assessing meaning in life in the participating countries. It has been suggested that instruments with more items may have better psychometric evidence than shorter ones (Gosling et al., 2003). However, in many situations it is not practical to use longer instruments, for example, in Internet-based studies with a longitudinal design or those in which many variables are used (Robins et al., 2001). Short instruments, such as the PIL-SF, take less time to answer, are cheaper, easier to administer, and reduce boredom (Herzberg & Brähler, 2006; Joseph et al., 2004). Also, the PIL-SF was shown to be equally reliable in all 7 countries. These results are similar to those reported in previous studies (García-Alandete et al., 2017; Pacak-Vedel et al., 2021; Schulenberg et al., 2011). The findings support the factorial structure and reliability in different countries and contribute to expanding the use of the PIL-SF, both for use in research and in clinical settings.

In addition, the invariance of the PIL-SF factorial structure in the seven countries with a new alignment optimization method. This method identifies non-invariant parameters and does not require exact invariance or different model modifications to make reliable comparisons between means (Asparouhov & Muthén, 2014). Overall, the results provided...
evidence of metric and scalar invariance of the PIL-SF. The findings suggest that the PIL-SF evaluates the same purpose in life construct in the participating Latin American countries. This is the first research study to test the measurement invariance of the PIL-SF across countries. The findings would make it possible to avoid errors when comparing the purpose of life between different countries. This allows for more solid conclusions in cross-cultural research on the meaning people find in life in Latin America. Therefore, those who wish to use the PIL-SF to assess purpose in life across Latin American countries can be assured that the differences obtained are a true reflection of the differences in the purpose in life construct and not a measurement error. In addition, there is also more evidence to evaluate the relationships between purpose in life and other theoretically and culturally relevant variables in all participating countries. Using the alignment optimization method, item intersections were reported to have a greater amount of invariance compared to factor loadings. Furthermore, the results may provide important information for the continued development of the PIL-SF. For example, item 3 was the least invariant in El Salvador and dealt with the achievement of life goals. This could be explained by the fact that personal goals vary according to different developmental stages, roles and life situations (Salmela-Aro et al., 2007; Zhang et al., 2009). Similarly, it is possible that this lack of invariance is also explained by differences in the characteristics of the sample across countries. For example, El Salvador was one of the countries with the lowest average age. This is important because personal goals change over time. Younger people tend to have more self-centered and education-related goals. Older people have goals more related to classic developmental tasks, such as starting a family (Krings et al., 2008).

In this study, Item Response Theory (IRT) was used to assess the characteristics of the PIL-SF items. IRT analysis allows for item and latent trait estimates that are independent of sample characteristics, trait-level standard errors, and item content (Hays et al., 2000). Evidence was provided about the high informativeness and adequate discrimination provided by each item of the PIL-SF, covering a wide range of the latent purpose in life construct. The findings indicate that people who respond to the highest response categories need a greater purpose in life. That is, the PIL-SF items are useful for determining people with a relatively high degree...
of purpose in life. Identifying people with higher levels of purpose in life would allow them to be part of intervention programs, being a support and model for those who have low levels. Likewise, identifying these people would allow them to be guided to maintain or reinforce their purpose in life. In addition, the information function of the PIL-SF indicated an adequate degree of accuracy when the level of purpose in life was medium to high. Specifically, item 4, which refers to the discovery of clear goals and a satisfactory life purpose, is the most accurate item of the PIL-SF to assess the latent trait life purpose. This can be explained by the fact that the study of the meaning people find in life is based on two main factors: the presence of a life purpose, which is experienced when a person understands himself/herself, understands the world around him/her and identifies his/her purpose in it, and the search for meaning (Steger et al., 2006).

**Strengths and Limitations**

This study has several strengths. First, it was the first to examine the measurement invariance of the PIL-SF in different Latin American countries. Second, it has made an important contribution to increasing the scope of application of the PIL-SF in the Latin American context. This is valuable because of the evidence that the scale is related to well-being (Pacak-Vedel et al., 2021). In addition, samples from different countries have been used, which helps to make the findings more robust. Also, robust statistical analyses have been performed, which have considered the ordinal level of items, as well as good and novel practices for assessing invariance across many groups. However, it is important to consider the limitations of the study. First, samples from only seven Latin American countries were accessed, mostly from South America (Argentina, Colombia, Ecuador, Paraguay and Uruguay) and only one each from North America (Mexico) and Central America (El Salvador). This leads to the possibility of examining in the future whether the results can be generalized to other Latin American countries. Similarly, it would be ideal to test the measurement invariance of the PIL-SF in more diverse countries and languages, such as those in little-studied contexts like Africa and the Middle East. Second, the use of non-probabilistic convenience sampling has not allowed us to obtain representative samples from each country. This has therefore made it impossible to generalize the results to the general population. Third, possible differences in the meaning of life between subgroups within countries (e.g., between people of different sexes and ages) were not assessed. Previously, it has been reported that women have higher meaning of life scores than men (Hamama & Hamama-Raz, 2021). Additionally, there is greater meaning found in life as age increases (Dhanjal, 2019). Fourth, for the most part, the participants in each country had completed university studies. Forthcoming studies should examine whether the invariance of PIL-SF holds in samples with lower educational attainment in different countries. It has been suggested that university students do not necessarily represent the general population of a country, often being a privileged group (Vogel et al., 2013). This could influence the degree to which people give meaning to their lives compared to people of other educational levels. In addition, people with lower educational levels were likely to have limited access to the internet. Fifth, because the study was conducted during the COVID-19 pandemic, infection and death rates were different in each participating country. This may have affected the meaning people gave to their lives. This leads to the suggestion that, future work could compare the structure of PIL-SF within countries, but at different time intervals. Sixth, the study used a cross-sectional design, so future research should use longitudinal designs to assess the temporal stability of the PIL-SF. Finally, the data were self-reported, so responses could have been affected by recall bias or social desirability.

**Theoretical Implications**

At the theoretical level, the findings confirm the unidimensionality of the purpose in life construct measured with the PIL-SF. As mentioned before, the 4 items have been grouped in the same factor that assesses the presence and realization of life goals/purposes (Molcar & Stuempfig, 1988; Steger, 2006). The fact that the 4 items are grouped into one factor is consistent with Frankl’s proposal, who mentions that when we are aware of our goals or purposes in life and our way of living is consistent with those goals, we are more likely to perceive our life as meaningful (Schulenberg et al., 2008). This finding is also consistent with other studies that recommend evaluating life purpose as a unidimensional concept (Marsh et al., 2003; Simkin et al., 2018; Steger, 2006). Similarly, having a one-dimensional measure indicates that each PIL-SF item reflects a single latent construct (Gefen, 2003). If the PIL-SF had one or more items that measure other aspects related to purpose in life, your total score should also include information on these other aspects. This would lead to a misinterpretation of the PIL-SF (Ziegler & Hagemann, 2015). Likewise, identifying that item three is the least invariant among all the items would lead to this item being chosen as the anchor or reference item when constructing a cross-country CFA model.

**Practical Implications**

The findings have several practical implications. First, the PIL-SF can be used in epidemiological studies to quickly determine the relationship between purpose in life and other constructs in the countries that participated in this study. Second, the PIL-SF can be used as an early screening measure to identify groups
of people with low purpose in life or to identify those who are likely to benefit most from increased purpose in life. Third, having a consistent measure across a group of Latin American countries can be useful to assess whether or not an intervention has had a desired effect on the development of purpose in life. In this way, evidence-based information can be obtained to support the development of policies to improve the mental health and well-being of the population. Fourth, the findings provide the opportunity for meaningful comparisons of latent means to detect variables that may influence meaning in life. However, the measurement and comparison of purpose in life across countries should be done on the basis of approximate invariance.

**Conclusion**

In conclusion, this study provided evidence of invariance of the PIL-SF, a psychometric measure of the purpose in life construct, in seven Latin American countries. Thus, the PIL-SF items are interpreted in the same way by people regardless of the country to which they belong. However, future studies could gather more evidence on the nature of the PIL-SF items by testing the one-factor model in other cultures less and more distant from those of the countries evaluated here.

**Data Availability** The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

**Declarations**

**Competing Interests** The authors declare that they have no conflict of interest.

**Informed Consent** Informed consent was obtained from all individual participants included in the study.

**References**

Ahorsu, D. K., Lin, C. Y., Imani, V., Saffari, M., Griffiths, M. D., & Pakpour, A. H. (2020). The fear of COVID-19 scale: Development and initial validation. *International Journal of Mental Health and Addiction*, 1-9. https://doi.org/10.1007/s11469-020-00270-8

Aiena, B. J., Buchanan, E. M., Smith, C. V., & Schulteenberg, S. E. (2016). Meaning, resilience, and traumatic stress after the Deepwater horizon oil spill: A study of Mississippi coastal residents seeking mental health services. *Journal of Clinical Psychology*, 72(12), 1264–1278. https://doi.org/10.1002/jclp.22232

Akbari, B., Soghra, F. S., & Reza, G. J. (2019). Canonical analysis of the relationships of religiosity, hope, and optimism with the meaning of life and quality of life in spinal cord injury patients. *Journal of Religion and Health*, 7(1), 11–19.

Armas, M., Castedo, A., & Cabaco, A. S. (2018). Fiabilidad y validez del “Purpose in life” (PIL) en una muestra clinica. *European Journal of Health Research*, 4, 43–51. https://doi.org/10.30552/ejhr.v4i1.89

Arslan, G., & Allen, K. A. (2021). Exploring the association between coronavirus stress, meaning in life, psychological flexibility, and subjective well-being. *Psychology, Health & Medicine*, 1-12. https://doi.org/10.1080/13548506.2021.1876892

Asparouhov, T., & Muthén, B. (2014). Multiple-group factor analysis alignment. *Structural Equation Modeling: A Multidisciplinary Journal*, 21(4), 495–508. https://doi.org/10.1080/10705511.2014.919210

Atak, H., & Çök, F. (2008). The Turkish version of inventory of the dimensions of emerging adulthood (the IDEA). *International Journal of Educational and Pedagogical Sciences*, 2(4), 392–396.

Bader, M., Jobst, L. J., Zettler, I., Hilbig, B. E., & Mosshagen, M. (2021). Disentangling the effects of culture and language on measurement noninvariance in cross-cultural research: The culture, comprehension, and translation bias (CCT) procedure. *Psychological Assessment*, 33(5), 375–384. https://doi.org/10.1037/pas0000989

Baltaci, Ö. (2021). Prediction of fear of COVID-19: Meaning in life and psychological resilience. *Educational Policy Analysis and Strategic Research*, 16(2), 46–58.

Batthyány, A., & Gutman, D. (2006). *Empirical research in logotherapy and meaning-oriented psychotherapy*. Zeig, Tucker & Thieben.

Baumeister, R. F., Vohs, K. D., Aaker, J. L., & Garbinsky, E. N. (2013). Some key differences between a happy life and a meaningful life. *The Journal of Positive Psychology*, 8(6), 505–516. https://doi.org/10.1080/17439760.2013.830764

Bieda, A., Hörschfeld, G., Schönfeld, P., Brajovskaja, J., Zhang, X. C., & Margraf, J. (2017). Universal happiness? Cross-cultural measurement invariance of scales assessing positive mental health. *Psychological Assessment*, 29(4), 408–421.

Boer, D., Hanke, K., & He, J. (2018). On detecting systematic measurement error in cross-cultural research: A review and critical reflection on equivalence and invariance tests. *Journal of Cross-Cultural Psychology*, 49(5), 713–734. https://doi.org/10.1177/00223891177409042

Borsboom, D. (2006). When does measurement invariance matter? *Medical Care*, 44(11), S176–S181. https://doi.org/10.1097/01.0000000000001433

Boyraz, G., Lightsey Jr., O. R., & Can, A. (2013). The Turkish version of the meaning in life questionnaire: Assessing the measurement invariance across Turkish and American adult samples. *Journal of Personality Assessment*, 95(4), 423–431. https://doi.org/10.1080/00223891.2013.765882

Brown, T. A. (2015). *Confirmatory factor analysis for applied research* (Second ed.). Guilford Publications https://books.google.com.pe/books?id=TL2BQAQBAJ

Brunelli, C., Bianchi, E., Murr, L., Monfomosso, P., Bosi, M., Gangieri, L., et al. (2012). Italian validation of the purpose in life (PIL) test and the seeking of noetic goals (SONG) test in a population of cancer patients. *Supportive Care in Cancer*, 20(11), 2775–2783. https://doi.org/10.1007/s00520-012-1399-6

Byrne, B. M., & van de Vijver, F. J. R. (2017). El enfoque de alineación de máxima verosimilitud para evaluar de forma aproximada la invarianza de medida: Una aplicación intercultural paradigmática. *Psicothema*, 29(4), 539–551. https://doi.org/10.7334/psicothema2017.178

Byrne, B. M., & Watkins, D. (2003). The issue of measurement invariance revisited. *Journal of Cross-Cultural Psychology*, 34(2), 155–175. https://doi.org/10.1177/0022022102250225
Minkov, M. (2018). A revision of Hofstede’s model of national culture: Old evidence and new data from 56 countries. Cross Cultural & Strategic Management, 25, 231–256.

Molcar, C. C., & Stuempfig, D. W. (1988). Effects of world view on purpose in life. The Journal of Psychology, 122(4), 365–371. https://doi.org/10.1080/00223980.1988.9915523

Morgan, J., & Farsides, T. (2009). Measuring meaning in life. Journal of Happiness Studies, 10(2), 197–214. https://doi.org/10.1007/s10902-007-9075-0

Muthén, B., & Asparouhov, T. (2014). IRT studies of many groups: Psychological support for a replicable two-factor model. Journal of Personality, 82(1), 159–182. https://doi.org/10.1111/j.1467-6494.2007.00484.x

Schulenberg, S. E., Baczkowski, B. J., & Buchanan, E. M. (2014). Measuring search for meaning: A factor-analytic evaluation of the seeking of noetic goals test (SONG). Journal of Happiness Studies, 15(3), 693–715. https://doi.org/10.1007/s10902-013-9446-7

Schulenberg, S. E., Smith, C. V., Drescher, C. F., & Buchanan, E. M. (2016). Assessment of meaning in adolescents receiving clinical services in Mississippi following the Deepwater horizon oil spill: An application of the purpose in life test-Short form (PIL-SF). Journal of Clinical Psychology, 72(12), 1279–1286. https://doi.org/10.1002/jclp.22240

Schumacker, R. E., & Lomax, R. G. (2015). A Beginner’s guide to structural equation modeling: Fourth edition (4th ed.). Routledge.

Shen, D. T. L. L. (1988). Reliability and factorial structure of the Chinese version of the purpose in life questionnaire. Journal of Clinical Psychology, 44(3), 384–392. https://doi.org/10.1002/1097-4679(198805)44:3<384::AID-JCLP2270440312>3.0.CO;2-1

Shun-Ami, A., & Bareket-Bojmel, L. (2021). What indicates your life is meaningful? A new measure for the indicators of meaning in life (3ML). Journal of Happiness Studies, 22(2), 625–644. https://doi.org/10.1007/s10902-020-00244-6

Simkin, H., Matrángolo, G., & Azzollini, S. (2018). Argentine validation of the purpose in life test/Validación argentina del test de Propósito en la Vida. Estudios de Psicología, 39(1), 104–126. https://doi.org/10.1080/02109395.2017.1407903

Singh, K., Junnarkar, M., Jaswal, S., & Kaur, J. (2016). Validation of meaning in life questionnaire in Hindi (MLQ-H). Mental Health, Religion & Culture, 19(5), 448–458. https://doi.org/10.1080/13674676.2016.1189759

Steger, M. F. (2006). An illustration of issues in factor extraction and identification of dimensionality in psychological assessment data. Journal of Personality Assessment, 86(3), 263–272. https://doi.org/10.1207/s15327752apa8603_03

Steger, M. F., Frazier, P., Oishi, S., & Kaler, M. (2006). The meaning in life questionnaire: Assessing the presence of and search for meaning in life. Journal of Counseling Psychology, 53(1), 80–93. https://doi.org/10.1037/0022-0167.53.1.80

Steger, M. F., Kawabata, Y., Shimai, S., & Otake, K. (2008a). The meaningful life in Japan and the United States: Levels and correlates of meaning in life. Journal of Research in Personality, 42(3), 660–678. https://doi.org/10.1016/j.jrp.2007.09.003

Steger, M. F., Kashdan, T. B., Sullivan, B. A., & Lorenz, D. (2008b). Understanding the search for meaning in life: Personality, cognitive style, and the dynamic between seeking and experiencing meaning. Journal of Personality, 76(2), 199–228. https://doi.org/10.1111/j.1467-6494.2007.00484.x

Steger, M. F., Oishi, S., & Kashdan, T. B. (2009). Meaning in life across the life span: Levels and correlates of meaning in life from emerging adulthood to older adulthood. The Journal of Positive Psychology, 4(1), 43–52. https://doi.org/10.1080/1743977080230127

Stoyles, G., Chadwick, A., & Caputi, P. (2015). Purpose in life and well-being: The relationship between purpose in life, hope, coping, and inward sensitivity among first-year university students. Journal of Spirituality in Mental Health, 17(2), 119–134. https://doi.org/10.1080/19349637.2015.985558
