Humor in Life Scale (EHV): Development and Evidence of Validity and Reliability in Spanish Speakers

Escala de Humor ante la Vida (EHV): Desarrollo, Evidencias de Validez y Confiabilidad en Hispanohablantes

Nazira Calleja¹, Miriam Arrollo-Belmonte, & María Fernanda Valencia Camargo

Universidad Nacional Autónoma de México, México

Recibido 3 de marzo de 2022; aceptado 20 junio de 2022

Abstract

Despite the growing attention given to the study of humor, a measure in the Spanish language of people’s experience and enjoyment of humor in everyday life is still missing. The present study describes the development and validation of the EHV (from the Spanish Escala de Humor ante la Vida, Humor in Life Scale). In phase I, the items were developed using focus groups and interviews; their content validity was assessed through expert judgment. In phase II, the EHV was answered by two Mexican samples (N=1380), women and men, from 18 to 66 years; it was administered together with the Numeric Rating Scale of Humor and the Positive and Negative Affect Schedule, in both printed and electronic formats. The exploratory factor analysis supported the single factor structure and the confirmatory factor analysis showed adequate fit indices for the final eight-item scale; the factor structure was partially invariant between sexes. Reliability indices were satisfactory. Convergent and discriminant validity tests showed that the EHV is related to a global measure of humor and positive and negative affect. In sum, the results indicate that the EHV is a brief, valid and reliable measure to assess humor in life in Spanish speakers.

Keywords: Humor in life; Scale; Validity; Reliability; Psychometrics

Resumen

A pesar de la creciente atención que se ha otorgado al estudio del sentido del humor, no se contaba con un instrumento en español que evaluara la experiencia y el disfrute del humor en la vida diaria. La investigación...
Humor is present in our daily life and in most of our social interactions. It is a social phenomenon, a universal channel of communication (Betés de Toro, 2011). For Martin and Ford (2018), humor is a form of social play, profoundly influenced by culture. Humor is a very complex phenomenon. It can refer to the characteristics of a stimulus, the mental processes involved in the creation, perception, understanding, and appreciation of humor, or to the individual’s responses (Martin et al., 2003). Several theories have been proposed to explain it; however, (Ruch, 2008) argues that more theoretical and empirical work on the definition or foundation of the concepts is necessary.

Humor has been claimed to induce positive emotions, construct personal resources, enhance interpersonal relationships, strengthen bonds and give a sense of belonging to a group; it also helps to relieve tension, cope with stress, and reduce negative emotions, depression and anxiety (Ruch & Hofmann, 2017). In general, adaptive humor improves psychological and physical health (Martin et al., 2003; Kuiper & Harris, 2009) and it is associated with resilience (Coppari et al., 2018; Menoni & Klasse, 2007), positive affect (Cann et al., 2000) and life satisfaction (Caycho-Rodríguez et al., 2019; Schneider et al., 2018). Even in the recent Covid-19 crisis, humor has been one of the most frequent used strategies to make such a stressful experience more bearable (Cancelas-Ouviña, 2021) and its ensuing periods of confinement, has generated high levels of social stress on a global scale. In Spain, citizens were isolated in their homes and were not able to interact physically with family members, friends or co-workers. Different resources were employed to face this new stressful and unexpected situation (fitness, reading, painting, meditation, mindfulness, dancing, listening to music, playing instruments, cooking, etc.). During the early stages of Italy’s lockdown, individuals using humor to cope with troubled circumstances evaluated the Covid-19 humor as funnier and less aversive (Bischetti et al., 2021).

On average, men have higher humor production ability than women, as was found in a systematic quantitative meta-analysis on sex differences in hu-
HUMOR IN LIFE SCALE (EHV)

The relationship between religiousness and spirituality, and humor has also been studied. Saroglou (2002) found negative correlations of religiousness and religious fundamentalism with humor creation; no relation was found between religion and reported use of humor as coping. Marziali et al. (2008) also reported the lack of association between spirituality and coping humor in older adults.

Humor is strongly affected by culture. Cross-cultural studies of humor have shown that people in different societies make jokes and amuse themselves in different ways, and that what they find to be humorous also differs. For example, Thorson et al. (1997), when comparing responses from Croatian and American students to a humor scale, found differences in constructions of sense of humor between the two samples. Carbelo-Baquero et al. (2006) also reported some differences in the constructions of sense of humor between Spanish respondents, who scored higher on coping humor, and Americans, who tended to score higher on humor creativity.

Even though the study of humor as a field of psychology has rapidly expanded in the last decades (Martin & Ford, 2018), understanding the use of humor as a strategy to cope with life is still a challenge. Several multidimensional approaches to humor in everyday life have been proposed, studying a great variety of humor styles: cheerful, witty, deriding, amused, sarcastic, self-directed, canned (Heintz, 2017); socially warm, cold, reflective, boorish, competent, inept, earthy, repressed, benign and mean-spirited (Craik et al., 1996); enjoyment of humor, laughter, verbal humor, laughing at yourself, and humor under stress (McGhee, 1999, 2010). Nevertheless, unidimensional instruments that explicitly focus on the extent to which people experience and enjoy humor in the circumstances of everyday life, are yet to be developed, particularly in Mexico. In fact, very few psychological studies on humor have been conducted in this country (v.g., Heintz et al., 2020; Palomar et al., 2011; Villarreal et al., 2012). This lack is especially conspicuous since Mexican society prides itself for having a mocking humor that arises repeatedly in daily life (Portilla, 1997); Mexican humor is present in all aspects of life, even and perhaps especially those dimensions which other societies may consider taboo, like politics, sex, and death.

The present article is aimed at developing and validating the Humor in Life Scale (in Spanish, Escala de Humor ante la Vida, EHV). Even though there are a few existing instruments for measuring humor in Spanish, most of them have been developed initially with participants from English-speaking countries (mainly university students), so that, even when translated and adapted, they contain items that may feel unrelatable to Spanish speakers. One example of this is the item “Coping by using humor is an elegant way of adapting” of the Multidimensional Sense of Humor Scale (Thorson & Powell, 1993), translated as Hacer frente a la vida mediante el humor es una manera elegante de adaptarse (Carbelo-Baquero et al., 2006). To adapt with elegance is a concept that makes sense for English speaking peoples, but in Mexico, elegance is circumscribed to aesthetic taste, it is not used as a way to describe how people deal with life. For Mexican participants this item is confusing and unrelatable. In addition, scales on humor tend to focus on the ways in which people appreciate, produce or interpret humor; in contrast, the EHV was intended to measure, comprehensively, how people experience humor in their everyday life.

Humor in life was defined as finding humor in everyday life, enjoying humor and its positive effects (good mood, calm, and relaxation), having a playful and cheerful attitude, laughing at oneself, and using humor under difficult situations. The new instrument was developed in Phase I and the validity of the EHV was investigated in Phase II. The following elements were assessed: (1) its factor structure with EFA and CFA, (2) its metric invariance across sex, (3) its reliability, (4) its convergent validity with a global evaluation of humor, and (5) its construct validity with positive and negative affect and with similarities between groups having or not spiritual or religious beliefs and having or not a partner. A different sample was used for different validity tests.
Phase I. Development of the EHV

Focus groups and interviews were used to develop items for measuring the construct of humor in life. It was sought to create items that reflected the construct definition as outlined earlier, taking special care to capture how Mexican people create and enjoy humor in everyday life and how they use it to deal with challenging circumstances.

The items’ formulation followed three principles. First, the wording of the items had to reflect the respondents’ natural way of speaking. Second, items had to be written in the first person to involve the respondents personally, they also had to be formulated in an evaluative, emotional way. Third, they had to be positively worded, since Dalal and Carter (2009) have shown that when a scale contains both positively and negatively worded items, false factor solutions are generated. A pool of potential items was generated ($k = 30$). Experts (two psychologists with knowledge on the study of humor, and two with expertise on psychometry) assessed, on a five-point scale, the intelligibility and appropriateness of the items, as well as if they covered all relevant aspects of the construct (Presser & Blair, 1994). The degree of agreement among the experts was quantified into coefficients $V$ (Aiken, 1980). The results of $V$ values and confidence intervals showed that there were 14 items that did not reach the criterion of .70 (Charter, 2003); therefore, these items were deleted. The lower limit of the confidence interval for the remaining 16 items was .74 or more, indicating evidence of content validity.

The response format for the EHV was a seven-point scale ranging from 1 = Disagree to 7 = Strongly agree; an asymmetric format was used to avoid the ceiling effect. Higher levels indicate a greater presence of humor in life.

During the pilot application of the scale, the think-aloud technique was used to determine the thoughts that led participants to their responses (Collins, 2003). This technique helped to assess whether participants fully understood the items in the intended way. Based on the information gathered, the items were adapted linguistically and they were reduced to an 11-item scale.

Phase II. Validation of the EHV

Participants

Two samples from the general population were obtained, aged 18 to 66 years. The first sample consisted of 550 people and the second of 830. The participants were recruited online and offline. The data of the two samples were collected four months apart. The demographic characteristics of the participants are presented in Table 1. Because the samples were used for different validity tests, Table 1 also provides an overview of the analyses conducted in each one.

Instruments

Humor in Life Scale (EHV). The EHV is the scale developed in Phase I. It measures the presence of humor in everyday life, the enjoyment of humor and its positive effects, the playful and cheerful attitude, the laughing at oneself, and the humor under difficult situations. The EHV is integrated by 11 items with seven-point response options, from 1=Disagree to 7=Strongly agree. Higher scores indicate a greater presence of humor in life.

Numeric Rating Scale of Humor (NRSH). To obtain a global evaluation of humor, the participants were asked to indicate their position along a continuum between two end-points. The question was: “There are people who have a great sense of humor, that is, they laugh often and see the funny side of things. On a scale from 0 to 10, how much sense of humor would you say you have? 0 means Not at all and 10 means A lot.” This instrument was created for the present study based on the visual analog scale technique.

The Positive and Negative Affect Schedule (PANAS, Watson et al., 1988). This scale is widely used in the field of psychology for both clinical and non-clinical populations. The inventory contains two mood scales, 10 positive affect (PA) items (v.gr., enthusiasm, activation, interest, and pride) and 10 negative affect (NA) items (v.gr., anger, contempt, disgust, guilt, fear,
nervousness). Low PA is characterized by sadness and lethargy, and low NA is a state of calmness and serenity. In Mexico, Robles & Páez (2003) validated a Spanish version and reported a good internal reliability (for PA, $\alpha = .90$, for NA, $\alpha = .85$). In the current study, the participants rated the degree to which they generally experience each emotion on a scale from 1=Never to 7=Always. The scale, which originally had five options, was expanded to seven options, in order to maintain consistency with the other measures used. In this study, the CFA showed that the bifactorial model provided the best overall model fit when one item was eliminated for PA factor (12. Alert) and three for NA factor (4. Upset; 8. Hostile; 11. Irritable). The fit indices were: $\chi^2/df=2.688$, CFI=.973; SRMR=.047, RMSEA=.055, CI .90%; .047-.064. The interfactor correlation was -.381 and the global Cronbach’s alpha =.921.

**Psychometric analysis**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Sample 1 n=550</th>
<th>Sample 2 n=830</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (range: 18 to 66 years)</td>
<td>Mean=33.06±8.45</td>
<td>Mean=29.55±7.72</td>
</tr>
<tr>
<td>Women</td>
<td>69.5%</td>
<td>60.8%</td>
</tr>
<tr>
<td>Single</td>
<td>51.9%</td>
<td>65.6%</td>
</tr>
<tr>
<td>With children</td>
<td>39.4%</td>
<td>38.6%</td>
</tr>
<tr>
<td>Job</td>
<td>69.1%</td>
<td>49.8%</td>
</tr>
<tr>
<td>Students</td>
<td>44.5%</td>
<td>51.8%</td>
</tr>
<tr>
<td>Higher education</td>
<td>82.8%</td>
<td>86.4%</td>
</tr>
<tr>
<td>With spiritual or religious beliefs</td>
<td>64.2%</td>
<td>59.9%</td>
</tr>
<tr>
<td>Exploratory Factor Analysis (EFA)</td>
<td>Confirmatory Factor Analysis (CFA)</td>
<td></td>
</tr>
<tr>
<td>Convergent validity</td>
<td>Metric invariance</td>
<td></td>
</tr>
<tr>
<td>Construct validity</td>
<td>Cronbach’s $\alpha$</td>
<td></td>
</tr>
<tr>
<td>Composite reliability</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Procedure**

For its administration, the battery of items was developed in both printed and electronic versions. The ethical guidelines established by the Mexican Society of Psychology (Sociedad Mexicana de Psicología & Hernández, 2017) were followed in its application. The printed version was applied in public places (parks, commercial centers, hospital waiting rooms). Once the purpose of the study was explained, participants were invited to participate voluntarily and with the understanding that they could withdraw at any time; they were informed that their responses would be treated anonymously and confidentially; finally, they were asked to express their informed consent to participate in the study. Google Forms and E-survey creator platforms were used for electronic applications. Questionnaires were disseminated through social networks and by email. The questionnaires took approximately 15 min to complete. Data collection was performed just prior to the start of the COVID-19 lockdown.

**Statistical and psychometric analysis**

The following statistical analyses were performed in SPSS 22: means and standard deviations, skewness, kurtosis, corrected item-total correlations (corrected homogeneity coefficient, HIc), and alpha if item deleted. Reliability was obtained through internal consistency (Cronbach’s $\alpha$) and composite reliability index, for which Raykov & Shrout (2002) suggest a minimum value of 0.70. To examine the factor
structure and factor loads, as well as to determine the percentage of variance explained by the factor, an EFA was performed using the maximum likelihood method (Schermelleh-Engel et al., 2003), previously verifying that the Kaiser-Meyer-Olkin (KMO) index of sample adequacy were > .80 and Bartlett’s sphericity tests had p < .05 (Hair et al., 2019). The factor structure of the EHV was analyzed by generating models and using confirmatory factor analysis (CFA) in AMOS 22. The model tests were based on maximum likelihood estimation. To assess the quality of the models the following statistics were used: chi-square statistics, the comparative fit index (CFI), standardized root mean square residual (SRMR), and the root mean square error of approximation (RMSEA) with its corresponding confidence interval. CFI values close to .95 and above, SRMR values of .08 or lower and RMSEA values of .06 or lower were applied as indicative of a good fit to the data (Browne & Cudeck, 1993; Hu & Bentler, 1999). The average variance extracted (AVE) was calculated from \( \lambda \) obtained in CFA; as recommended by Fornell & Larcker (1981), its value must be above .50. To assess the EHV invariance across sexes (Vandenberg & Lance, 2000) tests of group mean differences, invariance of structural parameter estimates, a model that allowed the parameters to be freely estimated (configurational model or baseline) was compared with models that constrained factor loadings (weak measurement invariance model), then intercepts (strong measurement invariance model) and finally unique error variances (strict measurement invariance model) were calculated. Chi-square difference scores (\( \Delta \chi^2 \)), and the changes in CFI and RMSEA were used to compare the models. A strong invariance is supported when \( \Delta \text{CFI} \leq 0.01, \Delta \text{RMSEA} \leq 0.015 \) and \( \Delta \chi^2 \) results with p > .05 (Cheung & Rensvold, 2002). The internal consistency of the EHV was obtained (Nunnally & Bernstein, 1994), as well as the corrected homogeneity coefficient (HIC), alpha if the item is deleted and factor loads obtained through the exploratory factor analysis (with KMO = .933 and Bartlett’s sphericity test: \( \chi^2(55)=5873.425, p < .001 \), which yielded a one-dimensional structure, with a total explained variance of 55.97%. The internal consistency index was \( \alpha = .933 \).

The data obtained for the items HV10 and HV11 showed indices out of range, so it was decided to eliminate them. Therefore, for the next analysis, the EHV comprised of nine items (HV1 through HV9).

The unifactorial structure of the EHV was corroborated in the CFA. However, a better model fit was found without item HV9. Besides, the modification indices, to improve the fit, suggested associations between the error of this item with the errors of five of the remaining eight items. Therefore, it was decided to leave the scale integrated by eight items (HV1 through HV8). Table 3 shows the indices of the models with and without item HV9. The AVE obtained was .572, which exceeded the minimum criterion proposed of .50.

The metric invariance between women and the men was assessed in sample 2. A model that allowed the factor loadings to be freely estimated was compared with a model that constrained the factor loadings across the two groups, then this model was compared with one that constrained the intercepts in addition to the factor loadings, and finally this model in turn was compared with one that also constrained the residuals. The test for metric invariance (Table 4) led to a non-significant chi-square difference test in the first and second comparisons, yet the difference was significant in the third comparison. However, it has been recognized that strict measurement invariance tests are excessively restrictive (Bentler, 2006). In all cases, \( \Delta \text{CFI} \) were less than 0.01 and \( \Delta \text{RMSEA} \) were less than 0.015, so, together, these data showed that the factor structure is invariant across the sex groups.
Results indicated that men and women were not significantly different in mean humor in life scores ($M_{men}=5.30$, $SD=1.14$; $M_{women}=5.16$, $SD=1.16$, $t(549)=1.360$, $p=.174$, $d=0.121$).

**Reliability**

The internal consistency (Cronbach’s alpha) obtained for the EHV was .915 and the composite reliability index (omega) was .914, indicating satisfactory reliability.

**Convergent validity**

To establish the convergent validity of the EHV, it was correlated with the Numeric Rating Scale of Humor ($M=7.59$, $SD=1.80$, range=0 to 10). The correlation index was $r=.625$. 
Construct validity

To establish the construct validity, the EHV was correlated with the positive affect (PA) and negative affect (NA) subscales of the PANAS. According to the nomological network of humor in life, this construct should be positively related to PA and negatively related to NA. This hypothesized pattern was obtained, although the positive index was stronger than the negative one. Results showed that the participants with high scores in humor in life also had high scores in PA ($r=.574$, $p<.001$); the opposite was true for the NA variable ($r=-.205$, $p<.001$).

Two indicators that theoretically are not linked to humor are: having spiritual or religious beliefs and having a partner. Using mean differences between groups, it was found that humor in life scores were not influenced by either of these two variables ($M_{\text{with spiritual or religious beliefs}} = 4.94$, $SD = 1.24$; $M_{\text{without spiritual or religious beliefs}} = 4.90$, $SD = 1.20$, $t(811) = 0.454$, $p = .650$, $d = 0.032$; $M_{\text{with partner}} = 4.89$, $SD = 1.25$; $M_{\text{without partner}} = 4.93$, $SD = 1.20$, $t(820) = 0.470$, $p = .639$, $d = 0.033$).

Discussion

In this research, a measure capturing humor in life was developed. Based on a literature review, it was identified that, despite the growing research attention directed towards sense of humor (Martin & Ford, 2018), understanding the use of humor as a strategy to cope with life is still a challenge. It was also found that most of the instruments applied in Spanish-speaking populations, particularly in Mexico, have been scales designed and validated in languages other than Spanish.

To develop the EHV items, focus groups, interviews, and the think-aloud technique were used, which allowed us to capture typically Mexican forms of expression, such as: *Procuro tomar la vida con buen humor* (I try to live life with a good sense humor) or *Ver reír a otras personas me pone de buenas* (Seeing other people laugh puts me in a good mood). Compared to the translation of some items from English to Spanish, the EHV items were simpler and more understandable; for example, the item “I have often found that my problems have been greatly reduced when I tried to find something funny in them” of the Coping Humor Scale (Martin et al., 2003), translated as *Con frecuencia he encontrado que mis problemas se reducen en gran parte cuando he intentado encontrar algo divertido en ellos*, could correspond to the much briefer and simple item *Tengo la capacidad de ver el lado gracioso de las situaciones* (I am able to see the funny side of things) of the EHV.
Its validity was assessed using data from two independent but equivalent samples. In one of two samples, EFA revealed that the hypothesized unidimensional structure was supported. In the second sample, which was used to perform CFA, the model with eight items and one factor had good fit indices. All items loaded high on the latent dimension. The EHV yielded satisfactory reliability indices. Moreover, the latent structure in men and women was invariant, indicating the robustness of the scale.

Another aim of this research was to develop a measure that was related to other measure of humor. It was shown that the EHV had high positive correlations with the Numeric Rating Scale of Humor (NRSH). This result provide evidence for good convergent validity, since the EHV is not identical to that measure of humor, but still related to it (Carlson & Herdman, 2012).

The fact that both positive and negative affect were found to be correlated to humor in life (the former strongly and positively; the latter negatively and moderately), supports the EHV construct validity. Positive affect has to do with enthusiasm, activation, interest, and pride, whereas negative affect has to do with anger, guilt, fear, and nervousness (Watson et al., 1988). This finding lends support to the assumption that humor is related to experienced affect or mood (Cann et al., 2000), which could be explained by the relief humor theory, which states that people experience humor and laugh to reduce stress (Meyer, 2000). Humor might generate positive affect and might limit the negative emotional impact of unpleasant, stressful events.

Another verification of the construct validity of EHV, derived from the nomological network of humor, consisted of test differences in humor in life among different groups of participants. In line with previous findings (Marziali et al., 2008; Saroglou, 2002), no differences were obtained between individuals who have spiritual or religious beliefs and those who do not. There was also no difference between people with and without a partner. Although no studies were found that specifically related the presence/absence of a partner to humor in life, this lack of relationship can be inferred from the studies on the marital status and humor style (Saroglou et al., 2010), the role of humor in the choice of partner (Torres & Cano, 2019), and the effect of partners’ humor style similarity on the quality of their relationship (Barelds & Barelds-Dijkstra, 2010). Further research could confirm this hypothesis.

Previous studies have reported that men have higher humor production ability (Greengross et al., 2020) and that they use more aggressive humor than women (Edwards & Martin, 2010). We did not find sex differences in humor in life. Moake et al. (2018) state that gender role expectations might cause people to negatively evaluate women who use aggressive forms of humor, but normative expectations would allow both men and women to use common forms of humor (i.e., humor in life). Sex differences need to be addressed in future research.

**Limitations and suggestions for future research**

This study has some limitations that need to be considered. First, the two samples used self-reported data exclusively. Employing the same source of data (in this case, self-report) may inflate the associations between the constructs (Podsakoff et al., 2003). However, since variance that is shared by the predictors because of common methods of data collection gets attributed to neither predictor in multiple regression analysis (Cohen & Cohen, 2003), the observed relations likely result in true covariation between constructs.

Second, the EHV focuses on the presence of humor in everyday life. Thus, the instrument may not be appropriate for measuring humor reaction, production or appreciation or humor styles, for example. Due to the diversity of constructs of sense of humor, this distinction seems necessary.

Third, we recruited two samples via online platforms and offline. We tried to ensure the quality of the study samples checking that the second sample resembles the first one in terms of age, sex, children, occupation (job/students), educational level and spiritual or religious beliefs. So, we assume that our
findings are generalizable and not prone to selection or response bias.

Fourth, we were not yet able to compare responses on the EHV with samples of other Spanish speaking countries nor with samples of non-Spanish speakers. This would, however, be an important further step in generalizing the validity of the EHV. Likewise, it is necessary to study how the EHV relates to other relevant established and validated humor scales.

**Conclusion**

In the present study the EHV-8 was developed and validated. This instrument shows advantages over other already existing measures: a) it assesses, in a single dimension and globally, the extent to which people experience and enjoy humor in everyday life which, although it has been measured previously, it has been studied in a diverse multiplicity of dimensions (cf., Craik et al., 1996; Heintz, 2017; McGhee, 1999, 2010); b) it is a short instrument (eight items), its quick and easy application generates a good acceptance in the respondents, and its results can be processed quickly; c) unlike most of the scales that evaluate humor in Spanish-speakers, which have been translated and adapted to this language, the EHV-8 has been developed in a Latin-American context, using various cognitive probing techniques; d) given that humor, as a psychosocial variable, has been relatively understudied in Mexico, this instrument will allow characterizing the presence of humor in the daily life of Mexicans, and to compare it with the levels of this variable in people of other countries in future studies. Additionally, the reliability, and the construct and convergent validity of the newly developed EHV-8 were strongly supported in two suitably-sized samples.

Having instruments with evidence of validity and reliability to measure humor is a fundamental step for the development of robust theories on humor. It is also necessary for generating effective interventions to foster the use of humor in daily life. Given the proven physical and psychological benefits of having a sense of humor in one’s life, such interventions are essential.

**References**


