SUPLEMENTARY MATERIALS

The Hedonic and Arousal Affect Scale (HAAS): A brief adjective checklist to assess affect states

Study 1: Pilot item-selection

To develop the affective state scale (adjective format), a large pool of items was selected from different sources. First, for high arousal items, we used the adjectives with higher factor loadings in the PANAS (López-Gómez et al., 2015). Also, based on the literature reviewed, we introduced further affective experiences that are not included in the PANAS, such as 'vigorous', 'lively', and 'energetic' for the high positive affect dimension (Cohen et al., 2003; Gregg and Shepherd, 2009; Matthews et al., 1990; Norcross et al., 1984), or 'anxious' and 'uptight' for the high negative affect dimension (Cohen et al., 2003; Fernández et al., 2002).

Second, for low arousal items, we included adjectives used in the circumplex model of affect (Posner et al., 2005; Watson and Tellegen, 1985), the four dimensions of dispositional mood (Huelsman et al., 1998), the Activation and Safe/Content Affect Scale (Gilbert et al., 2008), the Cohen's emotions styles (Cohen et al., 2006, 2003), the Four-Dimension Mood Scale (Gregg and Shepherd, 2009), the Profile of Mood States (Fernández et al., 2002; Norcross et al., 1984), the Mood Adjective Checklist (Matthews et al., 1990), the Activation-Deactivation Adjective Check List (Thayer, 1986), the Modified Differential Emotions Scale (Fredrickson et al., 2003; Galanakis et al., 2016), the Multiple Affect Adjective Checklist (Lubin et al., 2001), the Mood Adjective Check List (Nowlis, 1965), and the Brief Mood Introspection Scale (Mayer and Gaschke, 2013). Furthermore, to assess the items' content validity, two authors (initials hidden for review) reviewed the fit between the adjectives and the low positive and negative affect constructs. Following Wild et al.'s (2005) recommendations, when no version of the item was available in a published instrument, an adjective translation-back translation procedure was carried out by the same authors, keeping those adjectives in which there was an agreement.

From the initial pool of items, 24 adjectives were finally selected (6 for each affect x arousal category) with consensus from the authors by following these criteria: (1) exclusion of synonym adjectives (e.g., scare-afraid-fearful) to avoid artificially inflated correlations within a factor (Bech, 1990; Boyle et al., 2015), which is a problem that has been detected in several well-known scales of depression (Fried, 2017); (2) exclusion of antonym adjectives (e.g., happy-unhappy/sad or tranquil-nervous) to avoid correlated error terms; (3) exclusion of extreme adjectives in the continuum (e.g., we used 'tired' instead of 'exhausted', or 'active' instead of 'euphoric') as they are less common in daily experiences; (4) exclusion of items not expressing affect but domains like cognitive and meta-cognitive processes of affect (e.g., positive-negative, good-bad, or pleasant-unpleasant), or personality constructs (e.g., 'shyness', 'optimistic', 'pessimistic', etc.); (5) inclusion of clinically-relevant adjectives (e.g., 'guilty', 'sad', 'calm') as well as adjectives relevant to experimental tasks (e.g., 'tired' or 'bored'). As three items per factor is the minimum number of items required for adequate model identification (Brown, 2015; Kline, 2015), the goal was to build a 12-item scale (i.e., 3 items per factor) as, furthermore, some authors have suggested that, for brief scales, the optimal trade-off between reliability and validity is found with two to four items by scale (Burisch, 1997).

Data analysis

Two CFA models were run using LISREL 8.8 (Jöreskog and Sörbom, 2006, 1996), first for a two-factor model (positive vs. negative affect: Pos and Neg factors) and second for a four-factor model (positive-negative affect x high-low arousal: Pos High, Pos Low, Neg High and Neg Low factors). The four-factor structure is the hypothesized model and the two-factor structure, a nested model (Brown, 2015), is used for comparison purposes. A nested model contains a subset of the free parameters of another model (i.e., the parent four-factor model). For the 24-item form, the two-factor (nested) model consists of 49 freed parameters: 24 factor loadings, 24 error variances (one per item), and 1 correlation between factors (each factor variance is fixed to 1 to define the latent variable metric). The four-factor (parent) model consists of 54 freed parameters: 24 factor loadings, 24 item errors, and 6 correlations between factors (all of them with a factor variance fixed to 1). In both cases, the number of elements of the input matrix is 24*(24 + 1)/2 = 300, so the two-factor model has 251 df (i.e., 300-49), and the four-factor model has 246 df (i.e., 300-54). We used a χ^2 difference test (χ^2_{diff}; 251 – 246 = 5 df) to compare the model fit of the two and four-factor solutions. For the 12-item short-form, the number of elements of the input matrix was 12(12 + 1)/2 = 78, so the two-factor model had 53 df (i.e., 78 – 25), the four-factor model had 48 df (i.e., 78 – 30), and the χ^2 difference test had 53 – 48 = 5 df. Complementarily, given that the extraction of additional factors can produce overfitting, we have also used the comparative indices Expected Cross Validation Index (ECVI) and Akaike Information Criterion (AIC), which are more conservative approaches.

Parameter estimation was performed using Robust Maximum Likelihood (RML), following the Yang-Wallentin et al. (2010) guidelines for conducting CFA with ordinal variables with LISREL. This estimation method uses a polychoric correlation matrix as input and, when run with LISREL, requires the previous estimation of the asymptotic variance-covariance matrix, resulting in listwise deletion of cases with missing values (Fokkema et al., 2013). To avoid listwise deletion, we selected only participants who had completed the 24 items (N = 259). For model fit evaluation, we used Root Mean Square Error of Approximation (RMSEA), Standardized Root Mean Square Residual (SRMR), Comparative Fit Index (CFI), and Tucker-Lewis Index (TLI-NNFI).

Results

Supplementary Table 1 shows descriptive statistics of the scores of the 24 items and the estimated factor loadings (λ_{ik}* standardized) for both the two-factor and the four-factor CFA models. The mean and standard deviation (SD) values of the items directly reflect how the scale was constructed. Thus, items with positive affect showed high mean values, and items with negative affect showed low mean values. There were 5 items within the block of items with negative affect (5, 11, 13, 23, and 24) that showed an asymmetric response pattern, with a positive skewness index > 2. This kind of response pattern could lead to some estimation problems when the CFA is conducted (i.e., underestimation of some parameters). Nevertheless, in a simulation study, Yang-Wallentin et al. (2010) showed that robust estimation methods, such as RML, are not influenced by the shape of the distribution in similar asymmetry conditions as those observed in this study.

Supplementary Table 1. Descriptive statistics of the scores in the 24 items and estimated factor loadings (two and four-factor models) of the HAAS.

Items	*Spanish translation*[1]	*Affect / Arousal*	Mean (SD)	Skewness	Kurtosis	*Factor loadings (24 items)*		*Factor loadings (12 items)*	
						2-factor model	*4-factor model*	*2-factor model*	*4-factor model*
Interested	Interesado por las cosas	Pos High	3.0 (.89)	-.92	.94	.625	.642	.630	.646
Active	Activo/a	Pos High	2.7 (.83)	-.24	-.22	.562	.654	.562	.636
Vigorous	Vigoroso/a	Pos High	2.0 (1.08)	-.28	-.49	.602	.696		
Lively	Animado/a	Pos High	2.8 (.89)	-.39	-.38	.849	.851		
Energetic	Lleno/a de energía	Pos High	2.3 (1.05)	-.23	-.50	.857	.903		
Strong	Fuerte	Pos High	2.3 (.97)	-.32	-.08	.784	.807	.741	.785
Calmed	Calmado/a	Pos Low	2.4 (.95)	-.22	-.22	.569	.799	.537	.565
Tranquil	Tranquilo/a	Pos Low	2.5 (1.01)	-.38	-.14	.538	.764		
Peaceful	En paz	Pos Low	2.4 (1.18)	-.46	-.61	.725	.777		
At ease	A gusto	Pos Low	2.8 (0.90)	-.43	-.08	.844	.798	.842	.866
Content	Contento/a	Pos Low	2.5 (0.92)	-.38	.08	.850	.685	.862	.844
Serene	Sereno/a	Pos Low	2.6 (1.00)	-.50	-.11	.717	.885		
Jittery	Agitado/a.	Neg High	1.2 (1.08)	.67	-.39	.562	.716		
Anxious	Ansioso/a.	Neg High	.9 (1.03)	.97	.25	.684	.777	.664	.690
Irritable	Irritable.	Neg High	.5 (.83)	2.18	4.84	.804	.801	.834	.861
Scared	Temeroso/a.	Neg High	.5 (.81)	1.55	2.31	.730	.756		
Uptight	Tenso/a.	Neg High	.8 (.95)	1.00	.23	.813	.910		
Guilty	Culpable.	Neg High	.4 (.87)	2.21	4.31	.807	.759	.808	.797
Bored	Aburrido/a.	Neg Low	.3 (.71)	2.52	6.24	.523	.586	.521	.584
Exhausted	Fatigado/a.	Neg Low	.9 (1.00)	.83	-.19	.633	.682		
Sad	Triste.	Neg Low	.4 (.76)	2.05	4.58	.840	.844	.833	.887
Dull	Apagado/a.	Neg Low	.5 (.84)	1.82	3.07	.793	.883		
Tired	Cansado/a.	Neg Low	1.3 (1.05)	.64	-.34	.576	.628	.510	.502
Downcast	Abatido/a.	Neg Low	.4 (.75)	2.49	6.53	.879	.865		

[1] To maximize participants' identification with the adjectives, the format of the HAAS items includes gender ending as needed.

Study 3: Concurrent validity

Measures

Depression Anxiety Stress Scale (DASS-21; Daza et al., 2002; Lovibond & Lovibond, 1995): self-report measure of stress, anxiety, and depression symptoms. Participants are asked to rate whether each symptom was present during the past week (4-point Likert scale from 0 -"did not apply to me at all"- to 3 -"applied to me very much or most of the time"-).

Ruminative Response Style (RRS; Hervás, 2008; Nolen-Hoeksema & Morrow, 1991): self-report measure designed to assess rumination with two factors (cognitive reflection and brooding). While reflection is more associated with positive functioning, the brooding component has been found to be related to psychopathologies like depression or anxiety (Nolen-Hoeksema et al., 2008). Participants are asked to rate the frequency of ruminative thoughts when they are sad or depressed by using a 4-point Likert scale (from 1 = "almost never" to 4 = "almost always").

Pemberton Happiness Index (PHI; Hervás & Vázquez, 2013): self-report measure designed to assess integrative well-being, both hedonic and eudaimonic components of psychological well-being. It includes two subscales (remembered well-being and experienced well-being), but only the remembered well-being scale was included in this study. Responses are given on a scale from 0 (fully disagree) to 10 (fully agree).

Satisfaction with Life Scale (SWLS, Diener et al., 1985; Vázquez et al., 2013): self-report measure of global life satisfaction. Participants are asked to indicate their degree of agreement with each statement on a 7-point Likert scale (from 1 = strongly disagree to 7 = strongly agree).

Data analysis

First, a CFA using LISREL 8.8 (Jöreskog & Sörbom, 1996, 2006) was conducted based on the responses of the participants to each instrument, specifying the structural model recommended by previous research (i.e., three correlated factors for DASS-21, two correlated factors for RSS, and one factor for total scores in the PHI and the SWLS). Parameter estimation was performed using a polychoric correlation matrix and RML estimation method, except for the PHI instrument, whose responses (on an 11-point scale) were analyzed using the variance-covariance matrix and the Maximum Likelihood (ML) estimation method. Second, to assess concurrent validity, Pearson's correlations were calculated between each theoretical subscale (or total scale) of these instruments and each factor of the HAAS resulting from the two-dimension model. All correlations were calculated over the factor scores of each scale and subscale The correlation between factor scores and sum scores for the four factors of the HAAS were as follow: Pos High = .957, Pos Low = .969, Neg High = .958 and Neg Low = .838.

Study 4: Temporal invariance

Data analysis

Temporal invariance of the four-factor CFA model (12 items) was tested by simultaneously analyzing seven repeated measures of the HAAS (once a week) (Svetina et al., 2020; Wu & Estabrook, 2016). We used the lavaan (Rosseel, 2012) and the semTools (Jorgensen et al., 2018) R-packages, which allows for the implementation of the Wu & Estabrook (2016) model identification and delta parameterization to fit a baseline model (model 1: configural invariance), for a threshold equality model (model 2: metric invariance), and for a loading equality model (model 3: scale invariance; see Svetina et al., 2020 for technical details). When using the delta approach with ordinal data, the residual variances of the items are not allowed to be parameters of the models, so strict invariance cannot be tested. Nevertheless, to consider that the scores obtained over time are comparable, it is not necessary to test the equality of item residual variances (Brown, 2015). We tested equivalence between nested invariance models (configural, metric and scale) by a nonsignificant probability level of □2 (p-value > .05), a recommended approach to test invariance models with ordinal data (French and Finch, 2006 ; Svetina et al., 2020). In addition, we used RMSEA, CFI and TLI.

References

Bech, P., 1990. Measurement of psychological distress and well-being. Psychother. Psychosom. 54, 77–89. https://doi.org/10.1159/000288382

Boyle, G.J., Helmes, E., Matthews, G., Izard, C.E., 2015. Measures of affect dimensions, in: Measures of Personality and Social Psychological Constructs. Academic Press, pp. 190–224. https://doi.org/10.1016/B978-0-12-386915-9.00008-5

Brown, T., 2015. Confirmatory Factor Analysis for Applied Research. Second edition. Guilford Publications.

Burisch, M., 1997. Test Length and Validity revisited. Eur. J. Pers. 11, 303–315. https://doi.org/10.1177/014662168000400203

Cohen, S., Alper, C.M., Doyle, W.J., Treanor, J.J., Turner, R.B., 2006. Positive emotional style predicts resistance to illness after experimental exposure to rhinovirus or influenza A virus. Psychosom. Med. 68, 809–815. https://doi.org/10.1097/01.psy.0000245867.92364.3c

Cohen, S., Doyle, W.J., Turner, R.B., Alper, C.M., Skoner, D.P., 2003. Emotional style and susceptibility to the common cold. Psychosom. Med. 65, 652–657. https://doi.org/10.1097/01.PSY.0000077508.57784.DA

Daza, P., Novy, D.M., Stanley, M.A., Averill, P., 2002. The depression anxiety stress scale-21: Spanish translation and validation with a Hispanic sample. J. Psychopathol. Behav. Assess. 24, 195–205. https://doi.org/10.1023/A:1016014818163

Diener, E., Emmons, R.A., Larsen, R.J., Griffin, S., 1985. The Satisfaction with Life Scale. J. Pers. Assess. 49, 71–75.

Fernández, E., Fernández, C., Pesqueira, G., 2002. Adaptación al español del cuestionario "Perfil de los Estados de Ánimo" en una muestra de deportistas. Psicothema 708–713.

Fokkema, M., Smits, N., Kelderman, H., Cuijpers, P., 2013. Response shifts in mental health interventions: An illustration of longitudinal measurement invariance. Psychol. Assess. 25, 520–531. https://doi.org/10.1037/a0031669

Fredrickson, B., Tugade, M., Waugh, C., Larkin, G., 2003. What good are positive emotions in crisis? A prospective study of resilience and emotions following the terrorist attacks on the United States on September 11th, 2001. J. Pers. Soc. Psychol. 84, 365.

Fried, E.I., 2017. The 52 symptoms of major depression: Lack of content overlap among seven common depression scales. J. Affect. Disord. 208, 191–197. https://doi.org/10.1016/j.jad.2016.10.019

Galanakis, M., Stalikas, A., Pezirkianidis, C., Karakasidou, I., 2016. Reliability and validity of the modified differential emotions scale (mDES) in a Greek sample. Psychology 07, 101–113. https://doi.org/10.4236/psych.2016.71012

Gilbert, P., McEwan, K., Mitra, R., Franks, L., Richter, A., Rockliff, H., 2008. Feeling safe and content: A specific affect regulation system? Relationship to depression, anxiety, stress, and self-criticism. J. Posit. Psychol. 3, 182–191.

https://doi.org/10.1080/17439760801999461

Gregg, V.H., Shepherd, A.J., 2009. Factor structure of scores on the state version of the four dimension mood scale. Educ. Psychol. Meas. 69, 146–156. https://doi.org/10.1177/0013164408322002

Hervás, G., 2008. Adaptación al castellano de un instrumento para evaluar el estilo rumiativo. Rev. Psicopatología y Psicol. Clínica 13, 111–121. https://doi.org/10.5944/rppc.vol.13.num.2.2008.4054

Hervás, G., Vázquez, C., 2013. Construction and validation of a measure of integrative well-being in seven languages: the Pemberton Happiness Index. Health Qual. Life Outcomes 11, 66. https://doi.org/10.1186/1477-7525-11-66

Huelsman, T.J., Nemanick, R.C., Munz, D.C., 1998. Scales to measure four dimensions of dispositional mood: Positive energy, tiredness, negative activation, and relaxation. Educ. Psychol. Meas. 58, 804–819. https://doi.org/10.1177/0013164498058005006

Jöreskog, K., Sörbom, D., 2006. LISREL 8.8 for Windows. [Computer software]. IL: Scientific Software International, Lincolnwood.

Jöreskog, K., Sörbom, D., 1996. LISREL 8: User's reference guide. IL: Scientific Software International, Chicago.

Kline, R., 2015. Principles and practice of structural equation modeling. Guilford publications.

López-Gómez, I., Hervás, G., Vazquez, C., 2015. Adaptación de la "Escala de Afecto Positivo y Negativo" (PANAS) en una muestra general española. Psicol. Conductual 23, 529–548.

Lovibond, P.F., Lovibond, S.H., 1995. The structure of negative emotional states: Comparison of the Depression Anxiety Stress Scales (DASS) with the Beck

Depression and Anxiety Inventories. Behav. Res. Ther. 33, 335–343. https://doi.org/10.1016/0005-7967(94)00075-U

Lubin, B., Van Whitlock, R., Reddy, D., Petren, S., 2001. A comparison of the short and long forms of the Multiple Affect Adjective Check List-Revised (MAACL-R). J. Clin. Psychol. 57, 411–416. https://doi.org/10.1002/jclp.1023

Matthews, G., Jones, D.M., Chamberlain, A.G., 1990. Refining the measurement of mood: The UWIST mood adjective checklist. Br. J. Psychol. 81, 17–42. https://doi.org/10.1111/j.2044-8295.1990.tb02343.x

Mayer, J., Gaschke, Y., 2013. Brief Mood Introspection Scale (BMIS). Measurement Instrument Database for the Social Science.

Nolen-Hoeksema, S., Morrow, J., 1991. A prospective study of depression and posttraumatic stress symptoms after a natural disaster: The 1989 of Loma Prieta earthquake. J. Pers. Soc. Psychol. 61, 115–121.

Nolen-Hoeksema, S., Wisco, B., Lyubomirsky, S., 2008. Rethinking Rumination. Perspect. Psychol. Sci. 3, 400–424. https://doi.org/10.1111/j.1745-6924.2008.00088.x

Norcross, J.C., Guadagnoli, E., Prochaska, J.O., 1984. Factor structure of the Profile of Mood States (POMS): Two partial replications. J. Clin. Psychol. 40, 1270–1277. https://doi.org/10.1002/1097-4679(198409)40:5<1270::AID-JCLP2270400526>3.0.CO;2-7

Nowlis, V., 1965. Factor analytic studies of the mood adjective checklist. Rochester University NY department of psychology.

Posner, J., Russell, J.A., Peterson, B.S., 2005. The circumplex model of affect: An integrative approach to affective neuroscience, cognitive development, and psychopathology. Dev. Psychopathol. 17, 715–734.

https://doi.org/10.1017/S0954579405050340

Thayer, R., 1986. Activation-Deactivation Adjective Check List: Current overview and structural analysis. Psychol. Rep. 58, 607–614. https://doi.org/10.2466/pr0.1986.58.2.607

Vázquez, C., Duque, A., Hervás, G., 2013. Satisfaction with life scale in a representative sample of Spanish adults: Validation and normative data. Span. J. Psychol. 16, 1–15. https://doi.org/10.1017/sjp.2013.82

Watson, D., Tellegen, A., 1985. Toward a consensual structure of mood. Psychol. Bull. 98, 219–235. https://doi.org/10.1037/0033-2909.98.2.219

Wild, D., Grove, A., Martin, M., Eremenco, S., McElroy, S., Verjee-Lorenz, A., Erikson, P., 2005. Principles of good practice for the translation and cultural adaptation process for patient-reported outcomes (PRO) measures: Report of the ISPOR task force for translation and cultural adaptation. Value Heal. 8, 94–104. https://doi.org/10.1111/j.1524-4733.2005.04054.x

Yang-Wallentin, F., Jöreskog, K., Luo, H., 2010. Confirmatory factor analysis of ordinal variables with misspecified models. Struct. Equ. Model. 17, 392–423. https://doi.org/10.1080/10705511.2010.489003