



ORIGINAL

## Psychometric properties of the Reward Probability Index in a Colombian sample

Pablo L. Reyes-Buitrago <sup>a</sup>, Javier M. Bianchi <sup>a,\*</sup>, Juan C. Suárez-Falcón <sup>b</sup>, Francisco J. Ruiz <sup>a</sup>

<sup>a</sup> Fundación Universitaria Konrad Lorenz, Bogotá, Colombia

<sup>b</sup> Universidad Nacional de Educación a Distancia (UNED), Madrid, Spain

Received 23 May 2022; accepted 19 October 2022

### KEY WORDS

RPI,  
response-contingent  
positive reinforcement  
(RCPR),  
behavioral activation,  
depression

**Abstract** **Introduction:** This study analysed the psychometric properties of the Reward Probability Index (RPI) in an online Colombian sample with 1129 participants. **Method:** To conduct a cross-validation study, the sample was randomly divided into two subsamples. An exploratory factor analysis was conducted with the first subsample yielding a two-factor structure. Then, the fit of this two-factor model was tested on the second subsample by conducting a confirmatory factor analysis. **Results:** This model obtained a good fit to the data and measurement invariance across gender was observed. The RPI also showed good internal consistency according to both Cronbach's alpha and McDonald's omega, scoring .88 in both cases. The RPI demonstrated convergent construct validity given its correlations with other related measures such as the Environmental Reward Observation Scale ( $r = .81$ ), and the full version of the Behavioral Activation Scale for Depression ( $r = .71$ ). **Conclusions:** The RPI showed good psychometric properties in this Colombian sample.

© 2023 Fundación Universitaria Konrad Lorenz. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

### Propiedades psicométricas del Índice de Probabilidad de Recompensa en una muestra Colombiana

### PALABRAS CLAVE

RPI,  
reforzamiento positivo  
contingente a la respuesta  
(RCPR),  
activación conductual,  
depresión

**Resumen** **Introducción:** Este artículo tuvo como objetivo analizar las propiedades psicométricas del Índice de Probabilidad de Recompensa (RPI) en una muestra colombiana en línea con 1129 participantes. **Método:** Para realizar un estudio de validación cruzada, la muestra se dividió aleatoriamente en dos submuestras. Se realizó un análisis factorial exploratorio con la primera submuestra que arrojó una estructura de dos factores. Luego, se probó el ajuste de este modelo de dos factores en la segunda submuestra mediante la realización de un análisis factorial confirmatorio. **Resultados:** Este modelo obtuvo un buen ajuste a los datos y se observó invarianza de medida entre sexos. El RPI también mostró buena consistencia interna según el alfa de Cronbach y el omega de McDonald (.88 en ambos casos) y validez de constructo

\* Corresponding author.

Email: [javier.m.bianchi@konradlorenz.edu.co](mailto:javier.m.bianchi@konradlorenz.edu.co)

<https://doi.org/10.14349/rlp.2023.v55.1>

0120-0534/© 2023 Fundación Universitaria Konrad Lorenz. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

convergente dadas las correlaciones con otras medidas relacionadas como la Escala de Observación de Recompensa Ambiental ( $r = .81$ ), y la versión de la Escala de Activación Conductual para la Depresión ( $r = .71$ ). **Conclusiones:** el RPI mostró buenas propiedades psicométricas en esta muestra colombiana.

© 2023 Fundación Universitaria Konrad Lorenz. Este es un artículo Open Access bajo la licencia CC BY-NC-ND (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Response-Contingent Positive Reinforcement (RCPR) describes an increase in the frequency or duration of a behaviour as a result of a reinforcer being delivered. RCPR may increase the likelihood of behaviours due to an individual experiencing self-mastery or achievement following a behaviour (Wang & Feng, 2022). Thus, decreased environmental and social reinforcement may result in the extinction of “healthy” adaptive behaviours and, consequently, lead to the dysphoric mood and passivity that often characterize depression (Shindaini et al., 2022), with empirical evidence supporting the relationship between a low rate of RCPR and emotional distress (Fernández-Rodríguez et al., 2022).

The idea that low RCPR might produce dysphoria and depressive symptoms has prompted suggestions that, from a behavioural perspective, low RCPR may be a critical factor in the etiology and maintenance of clinical depression (Berman et al., 2022; Santos et al., 2021). Empirically-supported psychological interventions for depression, such as behavioural activation (BA), are based on this hypothesis. A low rate of RCPR is also present in other problems addressed by contextual and behavioural approaches, such as experiential avoidance disorder (Dahl et al., 2004) and behavioural avoidance and inhibition (Wagener & Blairy, 2015). RCPR also plays an important role in cases of substance abuse, especially in relation to the deprivation of environmental rewards and engagement in psychoactive substance-free activities (Mesheha et al., 2018).

Carvalho et al. (2011) developed the Reward Probability Index (RPI) to assess ability to experience environmental rewards during the months prior. Factor analyses revealed that its two-factor structure explained 41.1% of the variance (Carvalho et al., 2011). Factor 1 is called Reward Likelihood (11 items) and combines items from the “Potentially Reinforcing Events” and “Instrumental Behaviours” dimensions. Factor 2 is called Environmental Suppressors (9 items) and combines items from “Availability of Reinforcement” and “Aversive Stimuli in the Environment”. The RPI showed good internal consistency for the overall scale ( $\alpha = .88$ ), as well as for the dimensions (Reward Likelihood:  $\alpha = .80$ ; Environmental Suppressors:  $\alpha = .87$ ), and good convergent validity given the moderate to strong correlations with depression scores and measures of activity and avoidance (Carvalho et al., 2011).

Wagener and Blairy (2015) validated the French-language version of the RPI. They reported that the two-factor model showed a good fit. They assessed internal consistency using the Cronbach’s alphas of the two RPI factors. The subscales obtained good internal consistency (Reward Likelihood:  $\alpha = .87$ ; Environmental Suppressors:  $\alpha = .86$ ). In another study, Voss et al. (2021) conducted a psychometric evaluation of the RPI in a population of emerging adult drinkers in the United

States using two studies with different samples (Study 1:  $n = 393$ ; Study 2:  $n = 602$ ). Study 1 conducted an exploratory factor analysis (EFA) that yielded a two-factor model. Study 2 applied several confirmatory factor analyses (CFAs) that showed that the two-factor model of a 17-item version (after removing items 4, 17, and 18) obtained the best fit. The two dimensions presented good internal consistency (Reward Likelihood:  $\alpha = .86$ ; Environmental Suppressors:  $\alpha = .84$ ).

In sum, RCPR seems to predict and correlate with the acquisition and maintenance of affective disorders (Fu et al., 2021). It is also related to evidence-based therapies such as BA, and its usefulness in problems related to substance use is being explored. However, to the best of our knowledge, the psychometric properties of the RPI have not been analysed in Spanish-speaking countries. Collado et al. (2014) conducted the only study that mentions a Spanish version of the RPI, in which the RPI was translated for a preliminary study that analysed the efficacy of BA in Latinos in the US. Obtaining evidence of the psychometric properties of the Spanish version of the RPI would provide an assessment tool that would enrich research and clinicians’ assessments for BA in Spanish-speaking countries. For this purpose, the present study aimed to analyse the psychometric properties, factor structure, and gender invariance of the RPI in a large Colombian sample.

## Method

### Participants

The sample consisted of 1129 adult participants ( $M = 27.56$  years,  $SD = 8.92$ ; 74.2% women, 25.3% men, .44% did not report gender) from 28 Colombian departments corresponding to the five regions of Colombia (Bogotá, Atlantic, Central, Eastern, and Pacific). They participated voluntarily in the research and provided informed consent. Some participants were excluded because they did not meet the age criteria ( $n = 26$ ) or because they did not complete the entire survey ( $n = 6$ ). Exclusion criteria included the presence of psychosis or a current diagnosis of a severe psychiatric or neurological disorder. None of the participants stated that they suffered from these types of disorder and so no participants were excluded for this reason. Participants did not receive financial compensation.

Participants were mainly employed (39.4%), single (72.9%), with university education (48.9%), at a medium socioeconomic level (42.9%), and from the Bogotá region (46.5%). Thirty-four participants (3%) reported a history of hospitalization for mental health reasons.

## Instruments

**Reward Probability Index (RPI).** The RPI (Carvalho et al., 2011) is a 20-item self-report measure designed to approximate RCPR by evaluating the magnitude of environmental rewards. The scale assesses the four dimensions of RCPR across two factors: Reward Likelihood and Environmental Suppressors. Participants rate each item on a 4-point Likert scale (1 = *strongly disagree*; 4 = *strongly agree*) for the past few months, with higher scores indicating higher reward levels. The psychometric properties of the RPI were established through three studies (Carvalho et al., 2011). The measure demonstrated strong internal consistency ( $\alpha = .88$ ) and good two-week test-retest reliability ( $r = .69$ ). Convergent validity was established from the strong correlations shown between measures of activity, avoidance, reward, and depression ( $r > .60$ ).

**Environmental Reward Observation Scale (EROS).** The EROS (Armento & Hopko, 2007) is a scale that assesses environmental reward using ten items that are responded to on a 4-point, Likert-type scale. The scoring range is 10-40, with higher scores indicating higher perception of reinforcement. The scale seeks to detect the magnitude of reinforcing events, the availability of reinforcement in the environment, and the individual's ability to elicit such reinforcement. The EROS presents strong internal consistency ( $\alpha = .88$ ) and excellent one-week test-retest reliability ( $r = .85$ ) (Armento & Hopko, 2007). The validation study for the Colombian population (Valderrama et al., 2016) found appropriate levels of internal consistency ( $\alpha = .87$ ) and a one-factor internal structure explaining 46.82% of the variance.

**Behavioral Activation For Depression Scale (BADS).** The BADS (Kanter et al., 2007) assesses the behaviours responsible for change (i.e., behaviours that should lead to increased contact with RCPR) according to the BA treatment model. The instrument has 25 items comprising four dimensions: Activation (Items 3, 4, 5, 11, 12, and 23), Avoidance/Rumination (Items 8, 9, 10, 13, 14, 15, 24, and 25), Difficulties at Work or School (Items 1, 2, 6, 21, and 22), and Social Impairment (Items 16, 17, 18, 19, and 20). The BADS also provides a total score (Kanter et al., 2007) and uses a 7-point, Likert-type scale. The scoring range is 0-150, with higher scores indicating greater activation. In the Colombian validation (Leal et al., 2016), the BADS showed appropriate levels of internal consistency ( $\alpha = .92$ ) and acceptable levels in its dimensions ( $\alpha > .81$ ). Significant correlations were observed with the Zung's Self-Assessment Scale for Depression ( $r = -.726$ ). CFA indicated that the four-factor model obtained an acceptable fit.

**Zung's Self-rating Depression Scale (Zung-SDS).** The Zung-SDS (Zung, 1965) consists of 20 items that are responded to on a 4-point, Likert-type scale. The scoring range is 20-80, with scores equal to or greater than 40 considered an indicator of the probable presence of a clinically-relevant major depressive episode (Campo-Arias et al., 2005). The validation study of the Colombian version of the instrument showed an adequate internal consistency ( $\alpha = .85$ ), sensitivity of 88.6% (95% CI [82.9, 98.9]) and specificity of 74.8% (95% CI [68.4, 80.2]) (Campo-Arias et al., 2005).

## Procedure

The present study was conducted as part of a master's thesis and followed the ethical guidelines of the Psychology Research Center at Fundación Universitaria Konrad Lorenz. The research adhered to the Declaration of Helsinki of ethical principles for medical research involving human participants.

**Phase I. Authorization.** Authorization was obtained from the original authors of the RPI for its use for academic and research purposes. The permission was requested and granted via email from the researchers who designed the scale (Carvalho et al., 2011) and those who had translated the version used in studies with Spanish speakers (Collado et al., 2014). The Spanish translation of the RPI was used in the study by Collado et al. (2014), but to the best of our knowledge, it has not been submitted to psychometric analyses.

**Phase II. RPI adaptation.** A pilot test was conducted with a group of 60 undergraduates to detect problems of clarity in each of the items of the RPI. Appropriate adjustments were subsequently made according to the observations in the pilot sample.

**Phase III. Application of the RPI.** To conduct a psychometric analysis of the RPI, the instruments presented above were applied to the sample of 1129 participants. The sample was recruited through a referral chain or snowball sampling via social networks and the Internet. The questionnaires were applied through a survey platform, with participants completing them in an average of 25 minutes and 7 seconds. The sample was randomly divided into two groups ( $n_1 = 563$ ;  $n_2 = 566$ ). The size of each group was big enough for the analyses to be performed according to the guideline stating that a minimum of 10 to 15 participants must be recruited per item to perform the confirmatory factor analyses (Price, 2016).

## Data analysis

We examined the data before conducting the statistical analyses and found no missing values. The sample was then randomly divided into two sub-samples using SPSS 19®. For the first random sample ( $n = 563$ ), we conducted an exploratory factor analysis (EFA) using Factor 10.5 (Lorenzo-Seva & Ferrando, 2006). Using the polychoric correlation technique, we selected the robust diagonally weighted least square estimation method (robust DWLS) with direct oblimin rotation. We applied the optimal implementation of parallel analysis (PA) using minimum rank factor analysis (Timmerman & Lorenzo-Seva, 2011) and the Hull method using robust CFI and robust RMSEA (Lorenzo-Seva et al., 2011) to determine the number of factors to retain in EFA. We computed Unidimensional Congruence (UniCo), Explained Common Variance (ECV), and mean of item residual absolute loadings (MIREAL) indexes to assess essential unidimensionality. Values larger than .95 and .85 in UniCo and ECV, respectively, suggest that data can be treated as essentially unidimensional, whereas for MIREAL, a value lower than .30 suggests unidimensionality (Ferrando & Lorenzo-Seva, 2018). Lastly, we obtained corrected item-total correlations to analyse the discrimination item index.

For the second subsample ( $n = 566$ ), we adopted the robust DWLS estimation method using polychoric correlations to conduct the CFA using LISREL 8.71<sup>©</sup>. We computed the Satorra-Bentler chi-square test for both the single-factor model and the two-factor model, alongside the following goodness-of-fit indexes: (a) the root mean square error of approximation (*RMSEA*), (b) the comparative fit index (*CFI*), and (c) the non-normed fit index (*NNFI*), (d) the standardized root mean square residual (*SRMR*), (e) the expected cross-validation index (*ECVI*), and (f) the parsimony normed fit index (*PNFI*). According to Hu and Bentler (1999), *RMSEA* values of .08 represent a good fit, and values below .05 indicate that the model has a very good fit to the data. For the *SRMR*, values of less than .08 represent a reasonable fit, and values of less than .05 indicate a good fit. Concerning the *CFI* and *NNFI*, values greater than .90 indicate a good fit, and those greater than .95 denote that the model has a very good fit to data. Lastly, lower *ECVI* values indicate a better fit of the model, and higher *PNFI* values indicate a more parsimonious model.

The construct reliability and evidence of the convergent and discriminant validity of the measurement model for RPI were analysed according to the suggestions made by Brown (2015). We calculated the composite reliability coefficient (CR) to reveal construct reliability. CR values higher than .70 indicate high construct reliability and adequate internal consistency. We analysed three additional criteria to learn more about the convergent validity of the measurement model, namely (a) that factor loadings should be statistically significant (standardized loadings estimates should be .40 or higher, and ideally .70 or higher), (b) CR should be higher than .70 in each factor, and (c) the average variance extracted (AVE) should be equal to or higher than .50 for each construct of the RPI. Lastly, to assess discriminant validity we considered two criteria: (a) inter-construct correlations of below .80, these provide evidence that the discriminant validity is adequate; and (b) that the square root of AVE ( $\sqrt{AVE}$ ) of each factor should be greater than the inter-construct correlations between factors.

Following the guidelines provided by Jöreskog (2005) and Millsap and Yun-Tein (2004), we performed additional multi-group confirmatory factor analyses (CFAs) on the whole sample ( $n = 1129$ ) to test for metric, scalar, and strict invariances across gender. We compared the relative fit of four increasingly restrictive models: the multiple-group baseline model, the metric invariance model, the scalar invariance model, and the strict invariance model. We compared the *RMSEA*, *CFI*, and *NNFI* indexes of the nested models, and selected the more constrained model (i.e., second model vs. first model, third model vs. second model, fourth model vs. third model) if the following criteria suggested by Cheung and Rensvold (2002) and Chen (2007) were met: (a) the change in *RMSEA* ( $\Delta RMSEA$ ) was lower than .01 and (b) the changes in *CFI* ( $\Delta CFI$ ) and *NNFI* ( $\Delta NNFI$ ) were equal to or greater than -.01.

To explore the internal consistency of the RPI, we computed Cronbach's alpha and McDonald's omega using the MBESS package in R on the whole sample (Kelley & Lai, 2012; Kelley & Pornprasertmanit, 2016), while providing percentile bootstrap 95% confidence intervals (CI) (Viladrich et al., 2017). Descriptive data were calculated with SPSS 19<sup>©</sup>. Lastly, we calculated Pearson correlations between

the RPI and other scales to assess the validity evidence concerning other relevant constructs measured in this study. Correlations were interpreted according to the guidelines provided by Cohen (1988), which state that  $r$  between .10 and .30 indicates a small correlation,  $r$  between .30 and .50 a medium correlation, and  $r > .50$  a strong correlation. Values of  $r$  greater than .70 are considered very strong correlations and are needed to establish convergent validity (Chmielewski et al., 2016).

## Results

### Descriptive statistics of instruments

Table 1 presents the scale scores and internal consistency indexes that show the reliability of the instruments used for this sample.

**Table 1.** Statistics and internal consistency coefficients of the instruments for each sample

Instrument	$\alpha$	Sample 1		Sample 2	
		M	SD	M	SD
<b>BADS</b>					
Activation	.78	27.72	7.87	28.19	7.63
Avoidance/ Rumination	.88	20.19	11.58	21.55	11.98
Work/ School	.82	11.77	7.38	12.06	7.37
Social	.86	9.79	7.61	10.04	7.90
<b>Total</b>	.93	93.97	27.71	92.54	28.29
<b>RPI</b>	.88	58.51	8.15	58.47	7.93
<b>EROS</b>	.82	28.36	4.84	28.39	4.87
<b>Zung-SDS</b>	.88	37.19	10.25	37.01	9.82

Note: Zung-SDS: Zung Self-report Depression Scale; EROS: Environmental Reward Observation Scale; BADS: Behavioral Activation Depression Scale; RPI: Reward Probability Index. M: Media; SD: Standard Deviation;  $\alpha$ : Cronbach's alpha

### Psychometric Quality of the Items

Table 2 shows the items of the RPI, their translations into Spanish, and the corrected item-total correlations found in the first random sample ( $n = 563$ ). All items showed good discrimination, with corrected item-total correlations ranging from .32 (Item 19) to .64 (Item 11). Cronbach's alpha was .88 (95% CI [.87, .89]), whereas McDonald's omega coefficient was .88 (95% CI [.87, .89]) for the total sample. The internal consistency of both dimensions was good (Reward Probability:  $\alpha = .85$ , 95% CI [.84, .87],  $\omega = .86$ , 95% CI [.84, .87]; Environmental Suppressors:  $\alpha = .83$ , 95% CI [.82, .85],  $\omega = .83$ , 95% CI [.82, .85]).

### Validity evidence based on internal structure

**Dimensionality.** We conducted an EFA on the RPI using the first random sample. The Bartlett's test statistic was statistically significant at 5928.7(190),  $p < .001$ , and the result of the Kaiser-Meyer-Olkin test was very good (.90).

**Table 2.** Item description of the RPI (Spanish), factor loadings from EFA for the first random sample, and corrected item-total correlations

Item	RP Factor loading EFA	ES Factor loading EFA	Corrected item-total Sample 1 <i>n</i> = 563
1. I have many interests that bring me pleasure [Tengo muchos intereses que me producen satisfacción/placer].	.48		.44
2. I make the most of opportunities that are available to me [Aprovecho al máximo las oportunidades que se me presentan].	.61		.45
3. My behaviors often have negative consequences [A menudo mis comportamientos tienen consecuencias negativas].		.34	.46
4. I make friends easily [Hago amistades fácilmente].	.83		.43
5. There are many activities that I find satisfying [Hay muchas actividades que me parecen satisfactorias].	.67		.58
6. I consider myself to be a person with many skills [Me considero una persona con muchas habilidades].	.72		.51
7. Things happen that make me feel hopeless or inadequate [Suceden cosas que me hacen sentir sin esperanza o inapropiado].		.57	.59
8. I feel a strong sense of achievement [Me siento muy satisfecho con mis logros].	.55		.53
9. Changes have happened in my life that have made it hard to find enjoyment [Ha habido cambios en mi vida que me dificultan disfrutar de las cosas].		.57	.46
10. It is easy to find good ways to spend my time [Me resulta fácil encontrar buenas maneras de utilizar mi tiempo].	.54		.51
11. I have the abilities to obtain pleasure in life [Tengo habilidades para obtener satisfacción en mi vida].	.72		.64
12. I have few financial resources, which limits what I can do [Tengo pocos recursos económicos, y eso limita las cosas que puedo hacer].		.47	.39
13. I have had many unpleasant experiences [He tenido muchas experiencias desagradables].		.77	.51
14. It seems like bad things always happen to me [Pareciera como si las cosas malas siempre me sucedieran a mí].		.77	.59
15. I have good social skills [Tengo buenas habilidades sociales].	.86		.54
16. I often get hurt by others [Frecuentemente otras personas me lastiman].		.89	.56
17. People have been mean or aggressive toward me [Las personas han sido malas o agresivas conmigo].		.83	.55
18. I have been very capable in jobs I have had [He sido competente en los empleos que he tenido].	.55		.34
19. I wish I could find a place to live that brought more satisfaction to my life [Me gustaría poder encontrar un lugar para vivir que trajera más satisfacción a mi vida].		.52	.32
20. I have many opportunities to socialize with people [Tengo muchas oportunidades de socializar con otras personas].	.75		.51

Note. RP= Reward Probability; ES= Environmental Suppressors; Values lower than .30 are excluded.

The PA suggested that two factors be extracted, which accounted for 51% of the variance. The results of the Hull method when applied to robust CFI and RMSEA also suggested that two common factors be retained to obtain a very good fit ( $CFI = .970$  y  $RMSEA = .067$ ). Table 2 also shows that factor loadings were good for all items, with Reward Probability from .479 (Item 1) to .864 (Items 15) and Environmental Suppressors from .336 (Item 3) to .887 (Items 16). Values of UniCo (.925), ECV (.765), and MIREAL (.314) suggest that the data should not be treated as essentially unidimensional. In conclusion, the results of the EFAs we conducted suggest that RPI can be treated as a two-dimension measure.

To corroborate the EFA results, we conducted a CFA on the second random sample ( $n = 566$ ) to analyse the single-factor model and the two-factor model. The fit of the single-factor model was poor ( $\chi^2/df = 9.724$ ,  $p < .001$ ;  $RMSEA = .124$ ,  $CFI = .885$ ,  $NNFI = .871$ ; and  $SRMR = .111$ ), whereas the two-factor model produced clearly better results ( $\chi^2/df = 3.440$ ,  $p < .001$ ;  $RMSEA = .066$ ,  $CFI = .968$ ,  $NNFI = .964$ ; and  $SRMR = .067$ ). The ECVI and PNFI values also support the superiority of the two-factor model (see Table 3). The results of the completely standardized solution of the two-factor model can be seen in Figure 1.

**Table 3.** Goodness-of-fit indexes for the single-factor and two-factor models

Goodness-of-fit indicators	Single-factor model	Two-factor model
S-B $\chi^2$ (df)	1653.202 (170)	581.419 (169)
RMSEA [90% CI]	.124 [.119, .130]	.0657 [.060, .072]
CFI	.885	.968
NNFI	.871	.964
SRMR	.111	.067
ECVI	3.068	1.174
PNFI	.781	.850

**Construct reliability and convergent and discriminant validity of the measurement model.** The CR values were higher than .70 for both factors (Reward Probability = .89 and Environmental Suppressors = .87), which indicates that the RPI showed high construct reliability. Regarding convergent validity, all factor loadings were statistically significant, with six standardized loadings higher than .70 and 13 higher than .40. Only Item 19 showed a factor loading slightly below the cutoff of .40 (lambda  $\lambda = .39$ ). The AVE values were lower than the cutoff of .50 for the two constructs (Reward Probability = .44, Environmental Suppressors = .43). However, the convergent validity of the constructs can be considered acceptable when the CR values are higher than .70 and AVE values higher than .40 (Huang et al., 2013).

Additionally, all CR values were very high, and the AVE values were higher than .40, which supports the convergent validity of the RPI. The RPI was also shown to have good discriminant validity because the inter-construct correlation was below .80 (correlation between Reward Probability and Environmental Suppressors = .55). Moreover, the square roots of the AVE values of each factor were higher than the

inter-construct correlation between this latent variable and any other factor ( $\sqrt{AVE}$  of Reward Probability = .66 and  $\sqrt{AVE}$  Environmental Suppressors = .66, higher than .55).

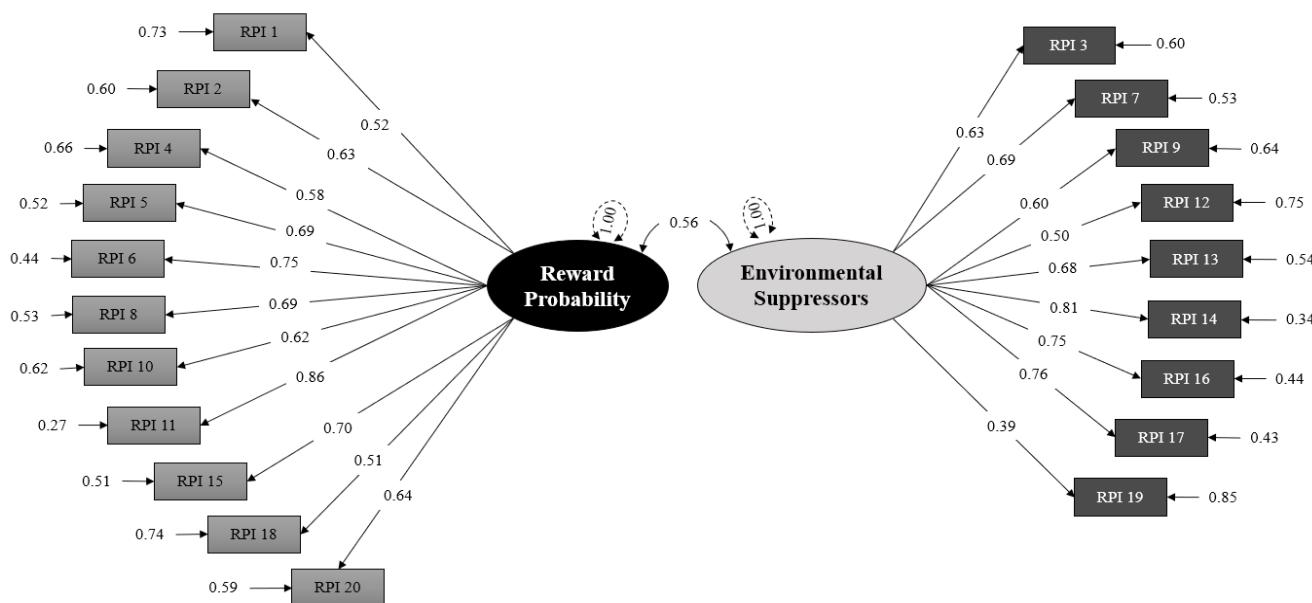
**Measurement invariance.** Table 4 presents the results of the measurement invariance analysis. The measurement invariance at metric, scalar, and strict levels was supported across gender because changes in RMSEA, CFI, and NNFI were below .01.

**Table 4.** Metric, scalar, and strict invariance across gender ( $N = 1129$ )

Model	RMSEA	$\Delta$ RMSEA	CFI	$\Delta$ CFI	NNFI	$\Delta$ NNFI
MG Baseline model	.068		.966		.962	
Metric invariance	.069	-.000	.964	-.002	.962	.000
Scalar invariance	.068	.001	.963	-.001	.963	.001
Strict invariance	.067	.001	.963	.000	.964	.001

#### Validity evidence based on relationships with other variables

The correlations that the RPI obtained with other relevant constructs are theoretically coherent and statistically significant:  $p < .001$  (see Figure 2 and Table 5). The RPI subscales showed correlations with the other assessed constructs that conformed to the expected directions and sizes. We observed strong correlations between the RPI total score and other measures of activation and avoidance ( $BADS_{activation} = .57$ ,  $BADS_{avoidance/rumination} = -.60$ ), environmental reward (EROS = .81), and depression (Zung-SDS = -.75). These correlations supported the convergent validity of the RPI. The relationships were largely consistent across both RPI factors.



**Figure 1.** Completely standardized solution of the two-factor model

**Table 5.** Pearson correlations between the scores for RPI and its two dimensions and the dimensions of BADS

Measure	Reward Probability [CI95%]	Environmental Suppressors [CI95%]	RPI [CI95%]
<b>BADS</b>			
Activation	.55*** [.50, .59]	-.42*** [-.47, -.37]	.57*** [.52, .60]
Avoidance/Rumination	-.39*** [-.44, -.34]	.63*** [.59, .67]	-.60*** [-.63, -.56]
Work/School	-.40*** [-.45, -.35]	.50*** [.45, .54]	-.52*** [-.56, -.48]
Social	-.68*** [-.50, -.41]	.56*** [.52, .60]	-.59*** [-.63, -.55]

Note. BADS: Behavioral Activation Depression Scale (dimensions: Activation, Avoidance/Rumination, Work/School and Social); \*\*\* $p < .001$ .

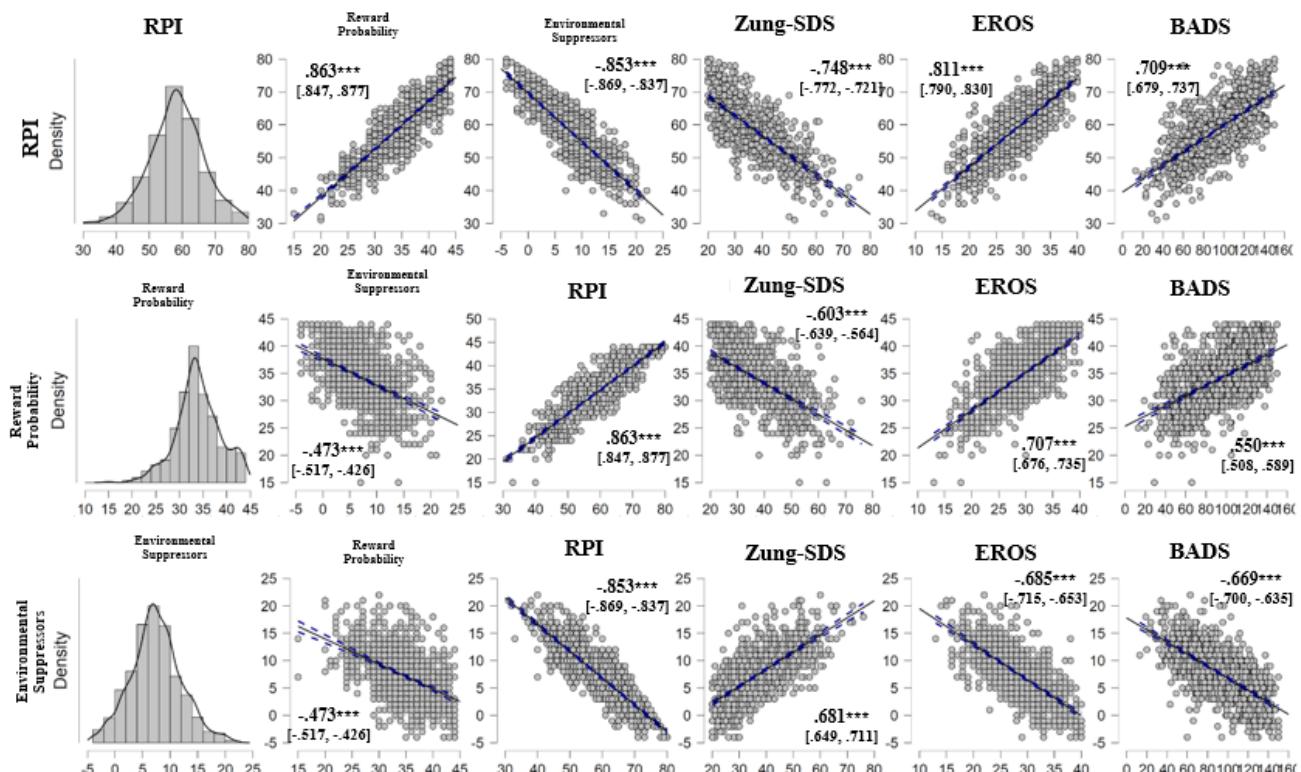
## Discussion

The objective of this research was to analyse the psychometric properties of the RPI in a Colombian sample. We administered the RPI and another related instrument to a sample of 1129 adult participants. The validity evidence ob-

tained through parallel analysis and Hull's method regarding internal structure suggests a bifactorial structure with adequate CFI and RMSEA values. All items showed good loadings on the dimensions Reward Likelihood ( $>.34$ ) and Environmental Suppressors ( $>.45$ ). The indexes of unidimensionality (UniCo, MIREAL, and ECV) did not support a single-factor model of the RPI. With regard to the CFA, the two-factor model presented in Carvalho et al. (2011) obtained a good fit to the data. The two-factor model also obtained a good fit in the French version of the RPI (Wagener & Blairy, 2015).

The measurement model of the RPI showed adequate construct reliability and convergent and discriminant validity. Specifically, the CFC was above .70, indicating that RPI has high construct reliability. In addition, all loadings were statistically significant (with values higher than .40), except for Item 19, which demonstrates convergent validity. As for discriminant validity, inter-construct correlations were lower than .80, and the values of the square root of AVE were higher than the correlation of the dimensions.

The complete RPI also obtained good internal consistency, according to  $\alpha$  and  $\omega$  coefficients (.88 and .89, respectively). The observed reliability was similar to that found by Carvalho et al. (2011) with the general population ( $\alpha = .88$ ). Likewise, the internal consistency of the two subscales was good (Reward Probability: .85 and .86; Environmental Suppressors: .83 in both indicators). Again, these results are similar to those reported in previous studies (Carvalho et al., 2011; Wagener & Blairy, 2015).

**Figure 2.** Pearson correlations between the RPI and other relevant self-report measures

Note: Zung-SDS: Zung Self-report Depression Scale; EROS: Environmental Reward Observation Scale; BADS: Behavioral Activation Depression Scale; RPI: Reward Probability Index; Pearson correlation ( $r_p$  [CI95%]), \*\*\* $p < .001$ .

This study is the first in providing empirical evidence of the measurement invariance of the RPI across gender. It therefore adds to the findings put forward by Voss et al. (2021) on emerging adult drinkers, who found factorial equivalence across ethnic groups and education levels. Adding further evidence of measurement invariance is relevant because it allows the RPI scores of different groups of participants to be compared (Greiff & Scherer, 2018).

The RPI also showed convergent construct validity given the high magnitude of the correlations (i.e.,  $> .70$ ) with similar and conceptually-related constructs such as the perception of reinforcement (EROS), activation (BADS-Total), and depression (Zung-SDS). As expected, significant and positive correlations were found with the EROS and BADS. This indicates that the higher Response-Contingent Positive Reinforcement (RCPR), the higher the perception of environmental reinforcement and the level of activation or engagement in "healthy behaviours" (Manos et al., 2010). In terms of the proposal that low RCPR may be a critical factor in the etiology and maintenance of clinical depression, significant, negative, and large correlations were found with depression (Carvalho & Hopko, 2011; Lewinsohn, 1974). Thus, exposing oneself to rewarding activities or events is related to lower levels of self-reported depression (Armento & Hopko, 2007). This evidence of convergent construct validity suggests that the RPI can be used as a clinical indicator, especially in the intervention of depression or other disorders, as well as studies which seek to recognize the availability of reinforcement, the ability to obtain reinforcement, and exposure to aversive events (Ryba et al., 2014).

Some limitations of the present study are worth mentioning. Firstly, the psychometric properties of the RPI were analysed in a nonclinical sample. Subsequent studies should analyse the functioning of the RPI in clinical samples in which RCPR is thought to play a relevant role. Secondly, the sample had a 3 to 1 ratio of women to men, meaning that the current findings can be better generalized to women. Secondly, the sample had a 3 to 1 ratio of women to men, meaning that the current findings can be better generalized to this population. However, this limitation is reduced given the measurement invariance that the two-factor model of the RPI demonstrated in this study. Thirdly, evidence of convergent construct validity was obtained only with measures related to depression. However, the RCPR could be useful in understanding problems related to ineffective instrumental behaviours and is related to other constructs such as exposure to aversive experiences, experiential avoidance (Dahl et al., 2004), and behavioural avoidance inhibition (Wagener & Blairy, 2015). Thus, exploring the relationship between the RPI and measures of these constructs might increase information on the convergent construct validity of the instrument. Fourthly, we have only explored the psychometric properties of the RPI in a sample from one Spanish-speaking country. Further studies should analyse the functioning of the RPI in other Spanish-speaking countries and analyse the measurement invariance across these samples. Lastly, this study did not obtain data concerning the sensitivity of the RPI to psychological interventions such as BA. Subsequent studies might analyse the treatment sensitivity of the RPI in interventions aimed at increasing RCPR.

To conclude, the present research reported on the adequacy of the two-factor model, the construct reliability

and convergent and discriminant validity, the measurement invariance across gender, and the convergent construct validity of the RPI in a large Colombian sample. Further studies should analyse the psychometric properties of the RPI in other Spanish-speaking countries and clinical samples.

## References

- Armento, M. E. A., & Hopko, D. R. (2007). The Environmental Reward Observation Scale (EROS): Development, validity, and reliability. *Behavior Therapy*, 38(2), 107-119. <https://doi.org/10.1016/j.beth.2006.05.003>
- Berman, N. C., Fang, A., Hoeppner, S. S., Reese, H., Siev, J., Timpano, K. R., & Wheaton, M. G. (2022). COVID-19 and obsessive-compulsive symptoms in a large multi-site college sample. *Journal of Obsessive-Compulsive and Related Disorders*, 33, 100727. <https://doi.org/10.1016/j.jocrd.2022.100727>
- Brown, T. A. (2015). *Confirmatory factor analysis for applied research*. Guilford publications.
- Campo-Arias, A., Díaz-Martínez, L. A., Rueda-Jaimes, G. E., & Barros-Bermúdez, J. A. (2005). Validación de la escala de Zung para depresión en universitarias de Bucaramanga, Colombia. *Revista Colombiana de Psiquiatría*, 34(1), 54-62. <https://bit.ly/3sJApz8>
- Carvalho, J. P., Gawrysiak, M. J., Hellmuth, J. C., McNulty, J. K., Magidson, J. F., Lejuez, C. W., & Hopko, D. R. (2011). The Reward probability index: design and validation of a scale measuring access to environmental reward. *Behavior Therapy*, 42(2), 249-262. <https://doi.org/10.1016/j.beth.2010.05.004>
- Carvalho, P. & Hopko, D. (2011). Behavioral theory of depression: Reinforcement as a mediating variable between avoidance and depression. *Journal of Behavior Therapy and Experimental Psychiatry*, 42(2), 154 - 162. <https://doi.org/10.1016/j.jbtep.2010.10.001>
- Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal*, 14(3), 464-504. <https://doi.org/10.1080/10705510701301834>
- Cheung, G.W., & Rensvold, R.B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal*, 9(2), 233-255. [https://doi.org/10.1207/S15328007SEM0902\\_5](https://doi.org/10.1207/S15328007SEM0902_5)
- Chmielewski, M., Sala, M., Tang, R., & Baldwin, A. (2016). Examining the construct validity of affective judgments of physical activity measures. *Psychological Assessment*, 28(9), 1128. <https://doi.org/10.1037/pas0000322>
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum.
- Collado, A., Castillo, S. D., Maero, F., Lejuez, C. W., & MacPherson, L. (2014). Pilot of the Brief Behavioral Activation Treatment for Depression in Latinos with limited english proficiency: Preliminary evaluation of efficacy and acceptability. *Behavior Therapy*, 45(1), 102-115. <https://doi.org/10.1016/j.beth.2013.10.001>
- Dahl, J., Wilson, K., & Nilsson, A. (2004). Acceptance and commitment therapy and the treatment of persons at risk for long-term disability resulting from stress and pain symptoms: A preliminary randomized trial. *Behavior Therapy*, 35(4), 785-801. [https://doi.org/10.1016/S0005-7894\(04\)80020-0](https://doi.org/10.1016/S0005-7894(04)80020-0)
- Fernández-Rodríguez, C., Coto-Lesmes, R., Martínez-Loredo, V., & Cuesta-Izquierdo, M. (2022). Psychological inflexibility, anxiety and depression: The moderating role of cognitive fusion, experiential avoidance and activation. *Psicothema*, 34(2), 240-248. <https://www.psicothema.com/pdf/4742.pdf>
- Ferrando, P. J., & Lorenzo-Seva, U. (2018). Assessing the quality and appropriateness of factor solutions and factor score estimates in exploratory item factor analysis. *Educational*

- and *Psychological Measurement*, 78(5), 762-780. <https://doi.org/10.1177/0013164417719308>
- Fu, Z., Brouwer, M., Kennis, M., Williams, A., Cuijpers, P., & Bocking, C. (2021). Psychological factors for the onset of depression: A meta-analysis of prospective studies. *BMJ Open*, 11(7), e050129. <https://doi.org/10.1136/bmjopen-2021-050129>
- Greiff, S., & Scherer, R. (2018). Still comparing apples with oranges? *European Journal of Psychological Assessment*, 34(3), 141-144. <https://doi.org/10.1027/1015-5759/a000487>
- Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6(1), 1-55. <https://doi.org/10.1080/10705519909540118>
- Huang, C. C., Wang, Y. M., Wu, T. W., & Wang, P. A. (2013). An empirical analysis of the antecedents and performance consequences of using the Moodle platform. *International Journal of Information and Education Technology*, 3(2), 217-221. <https://doi.org/10.7763/IJIET.2013.V3.267>
- Jöreskog, K. G. (2005). Structural equation modeling with ordinal variables using LISREL. [https://ssicentral.com/wp-content/uploads/2021/04/lis\\_ordinal.pdf](https://ssicentral.com/wp-content/uploads/2021/04/lis_ordinal.pdf)
- Kanter, J., Mulick, P., Busch, A., Berlin, K., & Martell, C. (2007). The behavioral activation for depression scale (BADS): Psychometric properties and factor structure. *Journal of Psychopathology and Behavioral Assessment*, 29(3), 191-202. <https://doi.org/10.1007/s10862-006-9038-5>
- Kelley, K., & Lai, K. (2012). *MBESS: The MBESS R package* (Version 3.3.2) [Software]. <http://CRAN.R-project.org/package=MBESS>
- Kelley, K., & Pornprasertmanit, S. (2016). Confidence intervals for population reliability coefficients: Evaluation of methods, recommendations, and software for composite measures. *Psychological Methods*, 21(1), 69-92. <https://doi.org/10.1037/a0040086>
- Leal, D., Bianchi, J., & Villalba, J. (2016). *Propiedades psicométricas de la escala de activación conductual para la depresión (BADS)* [Unpublished master's thesis]. Fundación Universitaria Konrad Lorenz, Bogotá, Colombia.
- Lewinsohn, P. M. (1974). A behavioral approach to the treatment of depression. In R. M. Friedman & M. M. Katz (Eds.), *The psychology of depression: Contemporary theory and research* (pp. 157-185). Wiley.
- Lorenzo-Seva, U., & Ferrando, P.J. (2006). FACTOR: A computer program to fit the exploratory factor analysis model. *Behavioral Research Methods*, 38, 88-91. <https://doi.org/10.3758/BF03192753>
- Lorenzo-Seva, U., Timmerman, M. E., & Kiers, H. A. (2011). The Hull method for selecting the number of common factors. *Multivariate Behavioral Research*, 46(2), 340-364. <https://doi.org/10.1080/00273171.2011.564527>
- Manos, R. C., Kanter, J. W., & Busch, A. M. (2010). A critical review of assessment strategies to measure the behavioral activation model of depression. *Clinical Psychology Review*, 30(5), 547-561. <https://doi.org/10.1016/j.cpr.2010.03.008>
- Meshesha, L. Z., Utzelmann, B., Dennhardt, A. A., & Murphy, J. G. (2018). A behavioral economic analysis of marijuana and other drug use among heavy drinking young adults. *Translational Issues in Psychological Science*, 4(1), 65-75. <https://doi.org/10.1037/tps0000144>
- Millsap, R. E., & Yun-Tein, J. (2004). Assessing factorial invariance in ordered-categorical measures. *Multivariate Behavioral Research*, 39(3), 479-515. [https://doi.org/10.1207/S15327906MBR3903\\_4](https://doi.org/10.1207/S15327906MBR3903_4)
- Price, L. R. (2016). *Psychometric Methods: Theory into Practice*. Guilford Press.
- Ryba, M. M., Lejuez, C. W., & Hopko, D. R. (2014). Behavioral Activation for depressed breast cancer patients: The impact of therapeutic compliance and quantity of activities completed on symptom reduction. *Journal of Consulting and Clinical Psychology*, 82(2), 325-335. <https://doi.org/10.1037/a0035363>
- Santos, M. M., Puspitasari, A. J., Nagy, G. A., & Kanter, J. W. (2021). Behavioral activation. In A. Wenzel (Ed.), *Handbook of cognitive behavioral therapy: Overview and approaches* (pp. 235-273). American Psychological Association. <https://doi.org/10.1037/0000218-009>
- Shindaini, A. I., Irin A., Rahman, M. M., & Afrin, S. (2022). Assessment of mental health conditions among working class people during COVID-19: Insights from Dhaka South City, Bangladesh. *International Journal of Public Health*, 11(2), 518-526.
- Timmerman, M. E., & Lorenzo-Seva, U. (2011). Dimensionality assessment of ordered polytomous items with parallel analysis. *Psychological Methods*, 16(2), 209-220. <https://doi.org/10.1037/a0023353>
- Valderrama-Díaz, M. A., Bianchi-Salguero, J. M., & Villalba-Garzón, J. A. (2016). Validación de la Environmental Reward Observation Scale (EROS) en población colombiana. *Universitas Psychologica*, 15(4), 1-13. <https://doi.org/10.11144/Javeriana.upsy15-4.vero>
- Viladrich, C., Angulo-Brunet, A., & Doval, E. (2017). A journey around alpha and omega to estimate internal consistency reliability. *Anales de Psicología*, 33(3), 755-782. <https://doi.org/10.6018/analesps.33.3.268401>
- Voss, A. T., Floyd, R. G., Campbell, K. W., Dennhardt, A. A., MacKillop, J., & Murphy, J. G. (2021). Psychometric evaluation of the Reward Probability Index in emerging adult drinkers. *Psychology of Addictive Behaviors*, 35(4), 432-443. <https://doi.org/10.1037/adb0000712>
- Wagener, A. & Blairy, S. (2015). Validation and psychometric properties of the French versions of the Environmental Reward Observation Scale and of the Reward Probability Index. *Psychologica Belgica*, 55(2), 71-86. <https://doi.org/10.5334/pb.bg>
- Wang, X. and Feng, Z. (2022). A narrative review of empirical literature of behavioral activation treatment for depression. *Frontiers in Psychiatry*, 13, 845138. <https://doi.org/10.3389/fpsyg.2022.845138>
- Zung, W. (1965). A self-rating depression scale. *Archives of General Psychology*, 12(1), 63-70. <https://doi.org/10.1001/archpsyc.1965.01720310065008>