Journal of Personality Assessment

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/hjpa20

Development and Validation of a Short Version of the Reinforcement Sensitivity Theory of Personality Questionnaire (RST-PQ-S)

Michele Vecchione & Philip J. Corr

To cite this article: Michele Vecchione & Philip J. Corr (2020): Development and Validation of a Short Version of the Reinforcement Sensitivity Theory of Personality Questionnaire (RST-PQ-S), Journal of Personality Assessment, DOI: 10.1080/00223891.2020.1801702

To link to this article: https://doi.org/10.1080/00223891.2020.1801702

JOURNAL OF PERSONALITY ASSESSMENT
https://doi.org/10.1080/00223891.2020.1801702

Development and Validation of a Short Version of the Reinforcement Sensitivity Theory of Personality Questionnaire (RST-PQ-S)

Michele Vecchione[1] and Philip J. Corr[2]

[1]Department of Social and Developmental Psychology, Sapienza University of Rome; [2]Department of Psychology, City University of London

ABSTRACT

We conducted three studies (total $n = 998$) aimed at developing and validating a shortened version of the Reinforcement Sensitivity Theory of Personality Questionnaire (i.e., the RST-PQ-S). In Study 1 ($n = 341$), twenty-two items were selected from the original RST-PQ, based on both theoretical and empirical criteria. Confirmatory factor analysis (CFA) on this short-form supported the six-factor structure comprising: FFFS, BIS, and four BAS components (Reward Interest, Goal Drive-Persistence, Reward Reactivity, and Impulsivity). The six scale scores showed adequate levels of internal consistency. Construct validity was supported by correlations with established personality measures. In Study 2 ($n = 340$), CFA results were cross-validated in an independent sample and construct validity was supported by correlations with BIS/BAS scales. In Study 3 ($n = 317$) test-retest correlations showed acceptable-to-good levels of temporal stability over a four-week interval. Results revealed a substantial overlap with original, full-length RST-PQ, providing evidence for the comparability of the two versions. Latent State-Trait analyses showed that the items of the RST-PQ-S mostly capture interindividual differences that are stable across situations. Taken together, findings indicate that the RST-PQ-S provides an efficient, valid and reliable alterative to the longer RST-PQ.

ARTICLE HISTORY
Received 9 January 2020
Accepted 23 June 2020

KEYWORDS
Reinforcement sensitivity theory; short-scale; revised RST

The reinforcement sensitivity theory (RST) of personality (Corr, 2008; Gray & McNaughton, 2000) postulates three major neuropsychological systems underlying individual differences in reaction to classes of repulsor and attractor stimuli (Corr & McNaughton, 2012): The Fight-Flight-Freeze System (FFFS), the Behavioral Inhibition System (BIS), and the Behavioral Approach System (BAS).

According to revised RST (McNaughton & Corr, 2004), the FFFS is activated by frustrating, punishing, and threatening stimuli that require immediate defensive action. When there is no motivation to approach a danger, this system elicits *active avoidance*, *flight*, and, for threats that cannot be easily avoided, *freezing*. Its activity is accompanied by the emotions of fear and, at very short (actual or perceived) defensive distance, panic. In contrast, the BIS is activated by conflicting stimuli, which elicit *cautious approach* to assess potential threat or *passive avoidance* (Gray & McNaughton, 2000). It is responsible for passively avoiding of and the withholding entry to a potentially dangerous field (Corr, 2013). Its activity is accompanied by feelings of anxiety, with increased vigilance toward the source of potential threat. Different components of the BIS functioning have been identified (Corr, 2008): Motor interruption (i.e., the inhibition of ongoing behaviors); risk assessment (i.e., cautious approach aimed at gathering information on potential threats); rumination of past unpleasant events and worry about possible future dangers; obsessive thoughts (i.e., emotional and cognitive engagement with unavoidable dangers); and disengagement (i.e., behavioral withdrawal from situations where the threat cannot be avoided and has to be approached).

The BAS is activated by conditioned and unconditioned attractor stimuli, such as cues for reward, non-punishment and escape from punishment (Corr & McNaughton, 2012), and it elicits approach behavior. As discussed by Corr and McNaughton (2008), the main function of the BAS is to move the organism along a spatio-temporal gradient toward a final biological reinforcer. In the conceptualization proposed by Corr (2008), four distinct but related BAS processes can be distinguished: *Reward Interest* refers to the initial motivation to seek out potentially rewarding stimuli; *Goal-Drive Persistence* is concerned with actively pursuing desired goals, especially when reward is not immediately available. This dimension is conceptually similar, although not identical, to Carver and White (1994) Drive scale; *Reward Reactivity* relates to excitement at doing things well and winning. It closely resembles Carver and White (1994) Reward Responsiveness scale; *Impulsivity* refers to acting fast without thinking in the final stage of reward capture, similarly to Carver and White (1994) Fun-Seeking scale. The theoretical and empirical bases for this multidimensional structure are detailed elsewhere (Corr, 2008, 2016).

CONTACT Michele Vecchione michele.vecchione@uniroma1.it Department of Social and Developmental Psychology, Sapienza University of Rome, Rome, Italy.

Supplemental data for this article is available online at https://doi.org/10.1080/00223891.2020.1801702.

Although several questionnaires have been proposed to assess individual differences in reactions to attractor and repulsor stimuli (e.g., Carver & White, 1994; Jackson, 2009; Reuter et al., 2015; Torrubia et al., 2001), a comprehensive measure of FFFS, BIS and BAS, as conceptualized in revised RST, has only recently been made available (for a review of existing RST-based questionnaires, see Corr, 2016, and Krupić et al., 2016). This measure, the Reinforcement Sensitivity Theory of Personality Questionnaire (RST-PQ, Corr & Cooper, 2016), comprises 65 items, divided over FFFS, BIS and BAS.

The FFFS and the BIS are unidimensional scales modeled on several facet domains, which capture different aspects of the underlying dimensions. The facet domains of FFFS (10 items) include *flight*, *active avoidance*, and *freeze*. The facet domains of BIS (23 items) include *motor planning interruption*, *cautious risk-assessment*, *obsessive thoughts*, and *behavioral disengagement*. The BAS include four related but separate scales: *Reward Interest* (7 items), *Goal-Drive Persistence* (7 items reflecting the facets of *goal planning* and *drive-persistence*), *Reward Reactivity* (10 items), and *Impulsivity* (8 items).

Corr and Cooper (2016) report five validation studies in their development of the RST-PQ. A confirmatory factor analysis (Study 3, $n = 831$) supported the expected six-factor structure, although the goodness of fit indices revealed mixed findings (i.e., the RMSEA indicated acceptable fit, whereas the CFI did not reach the recommended cut-off). As expected, BIS and FFFS showed a moderate positive correlation and the four BAS scales were positively related to each other, except for Goal-Drive Persistence and Impulsivity, which showed a near zero correlation. All scales showed high levels of internal consistency (Study 3, $n = 831$), with Cronbach's alpha reliability coefficients ranging from .74 (Impulsivity) to .93 (FFFS). These factors had meaningful patterns of relations with Carver and White (1994) BIS/BAS scales and other personality scales conceptually related to RST constructs, including the Big-5 personality traits, trait anxiety and fear (Study 4, $n = 362$).

After the original validation study in the U.K., a growing number of studies from different research fields have adopted the RST-PQ (e.g., Bacon et al., 2018; Beaton et al., 2017; De Pascalis & Scacchia, 2019; Jiang & Tiliopoulos, 2014). Moreover, the RST-PQ has been translated into German (Pugnaghi et al., 2018), Polish (Wytykowska et al., 2017), and Swedish (Eriksson et al., 2019).

In sum, although research using the RST-PQ is still in its infancy, this instrument displays promising psychometric properties and replicates across countries and in different languages. However, due to its length, it is relatively time consuming and often it is not convenient to include in studies that use multiple measurement instruments or adopt complex experimental paradigms. This fact only serves to hamper its use in such contexts. Moreover, it is less than optimal for use in large-scale surveys, where space is expensive and highly limited. In addition, with such a large number of items, difficulties may be encountered in the assessment of model fit. Factor analytic procedures on 65 manifest variables requires a very large sample size to be performed. Most importantly, adequate fit might be unrealistic for lengthy scales (e.g., Floyd & Widaman, 1995; Yang et al., 2010) and, even with an adequate sample size, goodness of fit statistics could be biased when the model being estimated is large (Moshagen, 2012).

The current research

The present research aims to develop and validate a short, 22-item, Italian version of the RST-PQ scale (i.e., the RST-PQ-S). We believe that such a shortened version would represent a useful alternative when practical constraints make the use of the full version either not feasible or recommended. In developing the RST-PQ-S, we followed standard recommendations for short-scale construction and validation (e.g., Rammstedt & Beierlein, 2014; Smith et al., 2000; Ziegler et al., 2014; for a thorough discussion of merits and limitations of shortened scales, see Kruyen et al., 2013, and Sandy et al., 2014).

The items of the short-form were selected in a first sample, based on both theoretical and empirical criteria. The psychometric properties of the RST-PQ-S were investigated in terms of factor structure and internal consistency, in accordance with classical test theory. Construct validity was assessed through correlations of scale scores with established personality variables conceptually linked to RST constructs. The factor structure and the internal consistency were further examined in a second, independent sample, along with correlations with Carver and White (1994) BIS/BAS scales. In a third, longitudinal study, reliability in terms of temporal stability was evaluated through test-retest correlations. This study also assessed the overlap between the short and the extended version, and compared their ability to predict a criterion variable (i.e., test anxiety). Moreover, it provided an estimate of the reliable proportion of variance in the RST-PQ-S items that can be attributable to state and trait components.

Study 1

The main aim of this study was to develop a short-form of the RST-PQ (i.e., the RST-PQ-S), drawing a subset of 22 items from the original, full version: Five items were selected to tap the FFFS and BIS scales; three items were selected for each of the four BAS dimension: Reward Interest, Goal-Drive Persistence, Reward Reactivity, and Impulsivity. This procedure provides a short-form with an adequate number of items per factor, which maintain approximately the same length of other instruments devised to assess similar constructs (e.g., Aluja & Blanch, 2011; Carver & White, 1994; Jackson, 2009).

After the items had been selected, we aimed to provide initial evidence about the psychometrics properties of the short-form, in terms of factor structure, internal consistency, and construct validity. Factor structure was examined by means of Confirmatory Factor Analyses (CFA). Internal consistency was evaluated using McDonald's (1999) omega,

a model-based reliability coefficient (i.e., the proportion of the total variance accounted for by the latent factor). We reported preliminary information on Cronbach's alpha, because this was the criteria used in the original validation study of the RST-PQ (Corr & Cooper, 2016). However, although alpha is the most widely used measure of internal consistency, omega represents a more appropriate approach (Chen et al., 2012).

Construct validity was assessed by examining correlations with personality measures included by Corr and Cooper (2016) in the original validation study (i.e., the Big-5 factors, trait anxiety, and fear toward aversive stimuli). We expected to replicate the pattern found in that study. Specifically, the four BAS scales were expected to be positively related to extraversion. Goal-Drive Persistence and Impulsivity should be related to high and low Conscientiousness, respectively. Furthermore, both FFFS and BIS should be positively related to neuroticism. However, BIS should be predominantly related to anxiety, whereas FFFS should be related mostly to fear.

Materials and methods

Participants and procedures

A total of 341 individuals (60% female, M_{age} = 28.14, SD = 8.86) participated in the study. Of these, 150 (46% females, M_{age} = 34.51, *SD* = 9.99) were Italian adults recruited from the general population by two psychology students as part of their master thesis requirement. Each student was briefed on the general aims of the research and instructed how to administer an online questionnaire, which included, in the following order, the RST-PQ and measures of the Big-5 personality factors, trait anxiety, and fear toward aversive stimuli, along with other measures that are not relevant to this study. Participation was voluntary, with no compensation. The other 191 participants were psychology students (72% females, M_{age} = 23.14, *SD* = 2.28), who completed the RST-PQ in class, through an online survey platform, as part of a course assignment at the University of Rome. Students were free to participate in the research and received course credit. All respondents were informed about the aim of the study and signed a consent form.

Measures

Participants completed an Italian version of the full, 65-item, version of the RST-PQ. For each item, they were asked how accurately it describes them on a 4-point Likert scale, ranging from *not at all* to *highly*. The order in which the items were presented was randomized across participants. The complete list of items is reported in the Online Supplementary material (Table S1). The instrument was adapted in Italian following recommended guidelines for translating and adapting questionnaires (e.g., Beaton et al., 2000). The forward translation (i.e., from English to Italian) was performed by the first author and a professional translator, who worked independently from each other. The two versions were then compared and merged into a single forward translation. A back-translation into the original language was performed by an Italian native speaker who was fluent in English. The original and back-translated versions were compared and checked by the second author. The final version was administered to a small sample of 5 students to check for comprehensiveness and clarity.

The subsample of 150 adults additionally completed several personality measures, including the Big Five personality traits, trait anxiety, and fear toward aversive stimuli. The Big-5 factors were assessed with the Mini-IPIP (Donnellan et al., 2006). We used the Italian translation of the items, available at https://ipip.ori.org/newItemTranslations.htm. The scale includes 20 items, 4 for each of the five personality dimensions (Extraversion, Agreeableness, Conscientiousness, Neuroticism, and Openness to Experience). For each item, respondents were asked to indicate how much it accurately described them using a 5-point Likert scale ranging from *strongly disagree* to *strongly agree*. In the current study, the alpha reliability coefficients of the five factors ranged from .61 (neuroticism) to .70 (extraversion).

The generalized propensity to be anxious was assessed with the Italian adaptation (Pedrabissi & Santinello, 1989) of the Trait Anxiety Scale, taken from the State-Trait Anxiety Inventory (STAI) Form-Y2 (Spielberger et al., 1983). The scale includes 20 item asking participants to assess how "in general" they felt, using a 4-point Likert scale, ranging from *almost never* to *almost always*. In the present study, Cronbach's alpha for the composite score was .86.

Fear toward several potentially aversive stimuli (e.g., *cemeteries*, *closed spaces*, *speaking in front of a group*) was assessed by using the Fear Survey Schedule (FSS; Wolpe & Lang, 1977). The version used in this study (Arrindell, 1980) includes 52 items tapping five domains: Agoraphobia (13 items), fear of animal (4 items), fear of sexual or aggressive situations (8 items), social fear (13 items), and fear of tissue damage (12 items). The items have been adapted to Italian by using the same procedure as for the RST-PQ. For each statement, respondents were asked to indicate the intensity of their fear, using a 4-point Likert scale, ranging from *no fear* to *very much fear*. Five composite scores were derived by summing item responses within each domain. In the present study, alpha coefficients ranged from .75 (fear of sexual or aggressive situations) to .85 (tissue damage fear).

Results and discussion

Preliminary results

Three hundred and fourteen participants (92% of the sample) provided complete data for the RST-PQ. For the remaining 8% (n = 27), mean number of missing items per participant was 6.74. Mahalanobis distance (with a $p < .001$ criteria) was used to inspect data for the presence of multivariate outliers (Tabachnick & Fidell, 2001). We identified and deleted six outliers, for a remaining sample size of 335. Univariate skewness ranged from −.63 to .72, with a mean of .33 (*SD* = .21); univariate kurtosis ranged from −1.44 to .59, with a mean of .55 (SD = .35). Mardia's test for

multivariate kurtosis was statistically significant (z = 9.27, $p < .001$), which indicated a moderate deviation from multivariate normality.

Item selection

The item selection procedure was performed in a two-step approach, which entails a combination of empirical and theoretical considerations (e.g., Smith et al., 2000). In the first step, we performed a set of CFAs on each of the six dimensions separately, using Mplus version 7.4 (Muthén & Muthén, 2010). Parameters were estimated by means of full-information maximum likelihood, with robust standard errors and rescaled test statistics (missing = FIML and estimator = MLR). This procedure allowed us to deal with missing data and non-normal distributions (the same procedure was used for subsequent analyses described below). We inspected factor loadings and modification indices to identify items with inadequate properties. We dropped items with weak (< .40) loadings on the target factor or highest correlated uniqueness. This ensured acceptable levels of internal consistency, while avoiding items with redundant content or wording. At the same time, we maintained an adequate number of items for each facet domain, thus preserving the content validity of the scale (Haynes et al., 1995; Smith et al., 2000). This led to the removal of 27 items: 3 from the FFFS scale (30% of the total), 12 from BIS (52%), 2 from Reward Interest (29%), 2 from Goal Drive-Persistence (29%), 5 from Reward Reactivity (50%), and 3 from Impulsivity (37%).

In the second step, we performed an Exploratory Factor Analysis (EFA) on the whole set of remaining 38 items. The six-factor solution had the following fit indices: $\chi^2(522) = 681.33$, $p < .001$, CFI = .951, TLI = .930, RMSEA = .030, 90% CI [.023, .036], SRMR = .031.[1] The resulting pattern of rotated factor loadings is reported in the Online Supplementary material (Table S2). Results of EFA were used to select the 22 items with the best psychometric properties, with the constrain that the target domains of the original RST-PQ were adequately represented in the short form (Haynes et al., 1995; Smith et al., 2000). This was accomplished by: (a) selecting the best marker variables (i.e., items with highest ratio between primary and secondary loadings) for each thematic facet; and (b) maintaining approximately the same proportion of items for each facet as in the original RST-PQ (Corr & Cooper, 2016). This approach might decrease the internal consistency of the short form but it affords sufficient construct coverage (Kruyen et al., 2013).

In the final version of the short form (RST-PQ-S), the FFFS scale included five items measuring the facets of *flight* (2 items), *active avoidance* (1 item), and *freeze* (2 items). The BIS scale includes five items measuring the facets of *motor planning interruption* (1 item), *cautious risk-assessment* (1 item), *obsessive thoughts* (2 items), and *behavioral disengagement* (1 item). The four BAS scales (Reward Interest, Goal-Drive Persistence, Reward Reactivity, and Impulsivity) are composed by three items each. The Goal-Drive Persistence scale includes items reflecting the two facet domains of *goal planning* (2 items) and *drive-persistence* (1 item) (Corr & Cooper, 2016).

[1]We interpreted CFI and TLI values > .90 (Bentler, 1990), RMSEA values < .08 (Browne & Cudeck, 1993), and SRMR values < .06 (Hu & Bentler, 1998) as indicating an adequate fit.

Table 1. Descriptive statistics and scale correlations of the RST-PQ-S (Study 1).

Scales	1.	2.	3a.	3b.	3c.	3d.
1. FFFS		.23*	−.10	−.01	.13	−.05
2. BIS			−.11	−.17*	.03	.10
3. BAS						
3a. BAS - Reward Interest				.25*	.26*	.32*
3b. BAS - Goal-Drive Persistence					.27*	−.06
3c. BAS - Reward Interest						.21*
3d. BAS - Impulsivity						
M	2.49	2.39	2.61	3.00	3.16	2.25
SD	.61	.66	.60	.59	.55	.65
Skewness	.04	.06	−.03	−.23	−.36	.20
Kurtosis	−.45	−.58	−.10	.10	−.51	−.58
Cronbach's Alpha	.60	.80	.77	.72	.61	.67
Average inter-item correlation	.24	.44	.52	.47	.35	.41

Note. Correlations among RST-PQ-S scales were tested for significance according to the Bonferroni-Holm procedure, which correct for multiple comparisons.
*adjusted p < .05.

Descriptive statistics

Table 1 reports descriptive statistics, Cronbach's alphas, and average inter-item correlations for the six scales of the RST-PQ-S. Alpha reliability coefficients ranged from .60 (FFFS) to .80 (BIS), showing acceptable to good levels of internal consistency. Overall, alphas were lower with respect to those reported for the full-length scale in the original validation study (Table 3, Study 3, Corr & Cooper, 2016). This is not unexpected, given the reduced number of items in the RST-PQ-S. Average inter-item correlations were within the recommended range of .15 − .50 (Briggs & Cheek, 1986; Clark & Watson, 1995), with exception of the Reward Interest scale, which fall slightly above (.52). Taken together, these estimates suggest that the items of the short form are relatively homogeneous, while avoiding unnecessary redundancy (Robins et al., 2007).

Internal validity and reliability

A six-factor CFA model on the 22 items of the RST-PQ-S, allowing the correlations among the factors, had adequate fit: $\chi^2(194) = 313.090$, $p < .001$, CFI = .921, TLI = .906, RMSEA = .04, 90% CI [.034, .051], SRMR = .057. Standardized factor loadings are shown in the upper panel of Table 2. They were all significant ($p < .001$) and greater than the .40 criterion suggested by Brown (2006). Intercorrelations among the six latent factors are reported in the lower panel of Table 2. As expected, BIS and FFS were moderately correlated and they showed non-significant or small, negative relations with the BAS scales. Correlations among the BAS scales were all positive and statistically significant, with the exception of the one between Impulsivity and Goal-Drive Persistence. This result is consistent with

Table 2. Standardized factor loadings of the RST-PQ-S (CFA in Study 1).

Items	F1: FFFS	F2: BIS	F3: BAS-RI	F4: BAS-GDP	F5: BAS-RR	F6: BAS-Imp
19. I would run quickly if fire alarms in a shopping mall started ringing	.42					
39. I would instantly freeze if I opened the door to find a stranger in the house	.58					
46. I would leave the park if I saw a group of dogs running around barking at people	.42					
48. I would freeze if I was on a turbulent aircraft	.59					
58. I would not hold a snake or spider	.44					
6. I sometime feel "blue" for no good reason		.65				
43. I often worry about letting down other people		.50				
49. My behavior is easily interrupted		.68				
50. It's difficult to get some things out of my mind		.87				
57. I often wake up with many thoughts running through my mind		.66				
14. I regularly try new activities just to see if I enjoy them			.74			
18. I get carried away by new projects			.70			
32. I am always finding new and interesting things to do			.74			
12. I am motivated to be successful in my personal life				.76		
41. I am very persistent in achieving my goals				.74		
65. I will actively put plans in place to accomplish goals in my life				.55		
16. Good news makes me feel over joyed					.58	
25. I get a special thrill when I am praised for something I've done well					.55	
36. I always celebrate when I accomplish something important					.66	
27. I sometimes cannot stop myself talking when I know I should keep my mouth closed						.53
28. I often do risky things without thinking of the consequences						.66
38. I find myself doing things on the spur of the moment						.73
Factor correlations and omega reliability coefficients	1.	2.	3.	4.	5.	6.
1. FFFS	*.62*					
2. BIS	.33*	*.81*				
3. BAS – RI	−.18	−.13	*.77*			
4. BAS – GDP	−.07	−.28*	.35*	*.73*		
5. BAS – RR	.16	.01	.38*	.39*	*.63*	
6. BAS – Imp	−.12	.13	.47*	−.09	.35*	*.68*

FFFS = Fight-Flight-Freeze System; BIS = Behavioral Inhibition System; RI = Reward Interest; GDP = Goal-Drive Persistence; RR = Reward Reactivity; Imp = Impulsivity. Significant correlations among latent factors ($p < .05$, Bonferroni–Holm-adjusted) are marked with asterisks. Reliability coefficients (McDonald's omega) are reported on the main diagonal of the correlation matrix.

U.K. (Corr & Cooper, 2016) and German (Pugnaghi et al., 2018) data using the full RST-PQ.

Based on CFA model parameters, we estimated omega reliability coefficients for the six RST-PQ-S factors. One of the main advantage of using omega rather than alpha for assessing the reliability of a shortened scale is that omega is not affected by scale length (Rammstedt & Beierlein, 2014; for a thorough discussion about properties of omega, as well as criticisms against alpha, interested readers are referred to Raykov, 1997, Revelle & Zinbarg, 2009, and Sijtsma, 2009). Conceptually, omega is similar to alpha: both provide an empirical estimate of the ratio of the true score variance to the observed score variance. The two coefficients have the same metrics and can, therefore, be interpreted by using the same standards. Results are reported in Table 2 (lower panel). Coefficients were in the range of .62 (FFFS) − .81 (BIS), thus suggesting adequate levels of internal consistency.

Construct validity

Table 3 reports Pearson correlations of the RST-PQ-S with personality measures, age and gender. As expected, FFFS, BIS and BAS were correlated in meaningful ways with the Big-5 factors. Both the FFFS and BIS were positively related to Neuroticism, although the correlation was larger for BIS. Reward Interest was significantly and positively related to Extraversion; Goal-Drive Persistence was positively related to Conscientiousness; Impulsivity was positively related to Extraversion, and negatively related to Conscientiousness.

As regard the other personality measures, the BIS showed a large positive correlation with trait anxiety. FFFS, by contrast, was more correlated than BIS to fear in all examined domains, with the exception of social fear, which showed a stronger positive correlation with BIS. This was expected, as social 'fear' in RST terms is a conflict stimuli and thus should be related to the BIS and anxiety. Our results parallel what was found in the original validation study (Corr & Cooper, 2016) and are consistent with the theory underlying the construction of the RST-PQ. Finally, gender and age were only moderately related to the RST-PQ-S scales. Females scored higher on FFFS and, to a lesser extent, on BIS and Reward Reactivity. Age was significantly correlated to BIS, with younger individuals having higher scores than older adults. These results are broadly in line with those found in the UK (Corr & Cooper, 2016) and German (Pugnaghi et al., 2018) samples.

Study 2

After selecting the items of the RST-PQ-S and examining their psychometric properties in a first sample, we conducted a second, independent, study. The CFA model tested in Study 1 provided only weak evidence for the psychometric properties of the instrument, due to potential effects of capitalization on chance. Therefore, the results needed to be

Table 3. Correlations between the RST-PQ-S and personality measures (Study 1).

	RST-PQ-S factors					
	FFFS	BIS	BAS-Reward Interest	BAS-Goal-Drive-Persistence	BAS-Reward Reactivity	BAS-Impulsivity
Five-factor model						
Extraversion	−.04	−.13	.27*	.19	.24	.27*
Agreeableness	.07	.05	.17	−.01	.16	.09
Conscientiousness	.07	−.23	−.09	.32*	.14	−.32*
Neuroticism	.33*	.67*	−.10	−.10	.11	.12
Openness	−.09	.08	.20	.06	.07	.21
STAI						
Trait Anxiety	.18	.75*	−.03	−.24	−.12	.20
Fear Survey Schedule						
Tissue fear	.53*	.25	−.08	−.09	.06	−.02
Social fear	.33*	.59*	.02	−.16	−.09	.09
Agoraphobia	.43*	.35*	−.14	−.13	−.05	−.02
Sex fear	.36*	.13	−.14	−.10	−.13	−.11
Animal fear	.54*	.39*	.06	−.08	.17	.12
Age	−.07	−.31*	.05	.04	−.13	−.05
Gender	−.34*	−.20*	.03	−.08	−.17*	.06

Note. * $p < .05$, Bonferroni–Holm adjusted.

cross-validated with independent data. We expected to replicate the six-factor structure observed in the first study.

Moreover, we further assessed the internal consistency of scale scores and examined their correlations with Carver and White (1994) BIS/BAS Scale, the most widely used RST measure. We expected to replicate the pattern of correlations found in the original validation study (Corr & Cooper, 2016). Accordingly, the BAS dimensions were expected to be positively related to one other. In accordance with revised RST (Corr, 2008), we expected that BAS Drive, which concerns the persistence in the pursuit of desired goals, would show the highest correlation with Goal-Drive Persistence. We expected that BAS Reward Responsiveness, which reflects a positive response to the occurrence of reward, would show the highest correlation with Reward Reactivity. Finally, we expected that BAS Fun-Seeking, which concerns the desire for new rewards and the willingness to approach them on the spur of the moment, would show the highest correlation with Impulsivity. Furthermore, we expected that Carver and White BIS scale would be positively related with RST-PQ-S BIS and, to a lesser extent, with FFFS.

It should be acknowledged, however, that Carver and White (1994) scales were based on Gray's (1987) original theory. Moreover, they were designed to assess BIS and BAS, but not FFFS. Since the theory has undergone important theoretical revisions (McNaughton & Corr, 2004), researchers have explored the possibility to obtain proxy measures of revised RST, by adapting Carver and White (1994) BIS/BAS scales (e.g., Corr & McNaughton, 2008; Johnson et al., 2003; Poythress et al., 2008). Among others, Heym et al. (2008) provided support for a model that distinguishes two factors within the original BIS scale, by separating items conceptually related to anxiety (BIS) from items that appear to tap fear (FFFS). We examined correlations of RST-PQ-S also with this alternative model, referred to as Heym et al. (2008) model.

Materials and methods

Participants and procedures

The study was part of a larger project on personality assessment that was conducted at the University of Rome (Italy). Three hundred and forty psychology students (64% females, $M_{age} = 27.48$, $SD = 12.72$) completed online a self-report questionnaire that included the RST-PQ-S, Carver and White BIS/BAS scales and other measures not relevant to this study. Each respondent was informed about the aim of the study and signed a consent form. Participation was voluntary, with no compensation.

Measures

Respondents completed the 22 items of the short-form developed in Study 1, and the Carver and White (1994) BIS/BAS Scales, a 20-item instrument to assess dispositional BIS and BAS sensitivities. We used the Italian adaptation of the instrument (Leone et al., 2002). Seven items were originally designed to measure BIS (e.g., "*I feel worried when I think I have done poorly on something*"). According to Heym et al. (2008), these items can be split into BIS (4 item – e.g., "*I feel worried when I think I have done poorly on something*") and FFFS (3 item – e.g., "*Even if something bad is about to happen, I rarely experience fear or nervousness*").

Thirteen items were designed to assess BAS, which comprises three subscales: Drive (4 items – e.g., "*I go out of my way to get things I want*"), Fun-Seeking (4 items - "*I crave excitement and new experiences*"), and Reward Responsiveness (5 items - "*When good things happen to me, it affects me strongly*"). All items are on a 4-point Likert scale, ranging from *strongly disagree* to *strongly agree*. Cronbach's reliability coefficients in the present sample were .85 for the original BIS scale, .73 and .74 for the BIS and FFFS scales adapted to reflect the revised theory (Heym et al., 2008), .77 for Drive, .79 for Fun-Seeking, and .85 for Reward Responsiveness.

Results and discussion

Preliminary results

The vast majority of participants (98%, $n = 332$) provided complete data for the RST-PQ-S. Based on Mahalanobis distance, two outliers ($p < .001$) were identified and excluded

Table 4. Correlations between the RST-PQ-S and BIS/BAS scales (Study 2).

		RST-PQ-S factors				
	FFFS	BIS	BAS-Reward Interest	BAS-Goal-Drive Persistence	BAS-Reward Reactivity	BAS-Impulsivity
BIS/BAS scales						
BIS – original Carver and White's (1994) scale	.48*	.68*	−.07	.02	.18*	−.05
BIS – revised Heym et al. (2008) scale	.44*	.65*	−.04	.03	.18*	−.04
FFFS – revised Heym et al. (2008) scale	.48*	.64*	−.10	.00	.16	−.04
BAS: Drive	.05	−.07	.38* (.11)	.55* (.47*)	.33* (.05)	.35* (.11)
BAS: Reward Responsiveness	.17*	.11	.31* (.06)	.44* (.27*)	.54* (.46*)	.23* (-.04)
BAS: Fun-Seeking	−.11	.04	.49* (.34**)	.15 (-.27*)	.25* (-.02)	.50* (.39*)

Note. * p < .05, Bonferroni–Holm-adjusted. Coefficients outside the parenthesis represent bivariate (zero-order) correlations. Coefficients in parentheses are partial correlations controlling for Carver and White (1994) BAS subscales.

from further analysis. Univariate skewness and kurtosis were in the range −.97 to .82 and −1.34 to .81, respectively. Mardia's test for multivariate kurtosis ($z = 2.01$, $p < .05$) suggests that the normality assumption was only mildly violated.

Internal validity and reliability

The six-factor model yielded close to acceptable fit on all criteria except for the TLI, which was slightly below the recommended threshold of .90: $\chi^2(194) = 357.45$, $p < .001$, CFI = .909, TLI = .892, RMSEA=.050 (.042, .058), SRMR = .057. Standardized loadings ranged from .40 to .85, providing support for the convergent validity of the scales. The full list of factor loadings and intercorrelations among factors are available in the Online Supplementary material (Table S3). Findings substantially replicated the pattern found in Study 1. McDonald's omega coefficients were .63 for FFFS, .82 for BIS, .79 for Reward Interest, .74 for Goal-Drive Persistence, .64 for Reward Reactivity, and .68 for Impulsivity. This indicates acceptable to high levels of internal consistency for the RST-PQ-S scale scores in the present sample.

Correlations with Carver and White's BIS/BAS original model

Table 4 reports zero-order Pearson's correlations between the RST-PQ-S and Carver and White (1994) BIS/BAS scales. As shown, the two BIS scales correlated highly. The FFFS also correlated, although less strongly, with Carver and White BIS. The BAS scales were all positively correlated with each other, except for Goal-Drive Persistence and Fun-Seeking. Since Carver and White BAS scales (Drive, Reward Responsiveness and Fun-Seeking) were substantially correlated (with coefficients in the range of .45–.55), we also examined their pattern of partial correlations with the RST-PQ-S BAS. Results are shown in the parentheses of Table 4. We found that Goal-Drive Persistence was mostly related to BAS Drive, Reward Reactivity was mostly related to BAS Reward Responsiveness, and Impulsivity was mostly related to BAS Fun-Seeking. This result aligns with revised RTS and provides empirical support to Corr's (2008) theoretical conceptualization of the four RST-PQ BAS components.

To compare the observed pattern of zero-order correlations with the one reported in the original validation study (Corr & Cooper, 2016, Study 4, $n = 362$), we calculated Cohen's q for each pair of corresponding correlations. This is an effect size measure of the difference between two correlations. Cohen's (1987) recommended benchmarks for q values are: <.1 = no effect; .1 to .3 = small effect; .3 to .5 = medium effect; >.5 = large effect. The differences we observed were negligible to small: the *q's* were in the range .00 − .24 ($M = .10$, $SD = .07$). We can, therefore, conclude that the short and full and scale showed a similar pattern of correlations with Carver and White's BIS/BAS scales.

Correlations with Heym et al. model

As reported in Table 4, both the RST-PQ FFFS and BIS scales correlated in similar ways with Heym et al. (2008) BIS and FFFS scales. Correlations between homologous dimensions (e.g., between the two BIS scales) were higher than correlations between heterologous scales (e.g., between the RST-PQ-S BIS and Heym et al. FFFS). Differences, however, were negligible. Thus, Heym et al. (2008) adapted version of Carver and White's BIS/BAS scales seems to capture only partially the distinction between FFFS and BIS, as conceptualized in the revised RST. In this regard, it should be noted that Heym et al. (2008) model provides only a proxy measure for the revised RST. Carver and White's scales, indeed, were theoretically conceived to assess BIS and BAS only. Moreover, several authors have raised concerns about the psychometric properties of the 3-item scale for the assessment of FFFS (e.g., Beck et al., 2009; Dissabandara et al., 2012; Maack & Ebesutani, 2018). Further studies are needed to confirm this result in other samples, using both the short and full version of the RST-PQ.

Study 3

The aims of this two-wave longitudinal study were threefold. First, we assessed the reliability of the RST-PQ-S in terms of temporal stability, through test-retest correlations. This is a recommended procedure for assessing the reliability of shortened scales (e.g., Rammstedt & Beierlein, 2014). Second, we assessed the degree to which the short and the full forms of the RST-PQ share similar psychometric properties – this is a critical issue in the process of short scales construction (e.g., Heene et al., 2014; Smith et al., 2000; Ziegler et al., 2014). Specifically, we examined the degree of overlap between the two forms, and compared their

criterion validity with regard to a measure of test anxiety, namely the disposition to experience concern about one's own performance in evaluative settings (Schwarzer et al., 1987; Zeidner, 1998). Two main manifestations of test anxiety have been identified (Sarason, 1961), commonly referred to as emotionality (i.e., heightened physiological activity), and worry (i.e., self-deprecating ruminations). Given the target population of the present study (university students), this appears a relevant criterion for assessing the concurrent validity of the RST-PQ-S. While there is preliminary evidence to suggest that test anxiety is related to sensitivity to punishment (Hagopian & Ollendick, 1994), the present study is the first to test this relation within the revised RST. We expected that, among the RST-PQ scales, test anxiety would be mostly related to BIS. According to the theory (Corr, 2008), this defensive system is activated by conflict or uncertainty, such as when there is an approach-avoidance situation. This is the case of academic settings, where students must approach a situation that presents signals of both reward (obtaining high grades) and punishment (emotional distress). Test anxiety should be unrelated or weakly related to the other defensive system, the FFFS, which is mediated by fear, not anxiety. Moreover, the FFFS is activated when there is no motivation to approach aversive stimuli that can be avoided (Gray & McNaughton, 2000). This does not describe what typically happens in academic settings. Finally, test anxiety was expected to be unrelated with BAS.

Third, a Latent State-Trait analysis was applied with the aim of estimating: (a) the reliability of the RST-PQ-S at the item-level; and (b) the extent to which the reliable source of variance in items responses reflect stable individual differences and systematic effects of the situation. In this regard, we expected a substantial degree of trait variance, in line with the assumption that RST constructs reflect sources of variation in neuropsychological systems that are stable over time (Corr, 2008; Pickering & Corr, 2008).

Materials and methods

Participants and procedures

Three hundred and seventeen university students (72% females, M_{age} = 21.28, SD = 3.30) completed online the 65 items of the full RST-PQ and a self-report measure of test-anxiety. Respondents were part of a larger project on personality assessment that was conducted at the University of Rome (Italy). The sample included two cohorts, comprising respectively 134 (78% females, M_{age} = 21.43, SD = 3.97) and 183 (68% females, M_{age} = 21.29, SD = 2.89) university students attending the same course in two subsequent semesters. The two cohorts have similar distribution with respect to age, t (315) = .35, p = .73, and gender, $\chi^2(1)$ = 3.72, p = .05, standardized residuals in the range −1.2 to 1.1. Overall, 75% of the sample (n = 238) completed the 22 items of the short form four weeks later. Students earned course credit in exchange for participation.

Measures

Respondents completed the full length version of the RST-PQ at Time 1, and the short-form at Time 2. Test anxiety was also assessed at Time 1, through the Test Anxiety Inventory (TAI, Spielberger, 1980). The TAI is a 20-item scale designed to measure individual differences in worry and emotionality, the two main dimensions of test anxiety (Liebert & Morris, 1967). The *Worry* subscale (8 items) refers to a cognitive component that includes intrusive and ruminative thoughts related to the consequences of failure. The *Emotionality* subscale (8 items) concerns affective and physiological reactions, such as tension and heart rate, which are experienced by the individual in the evaluative situation.

We used an Italian adaptation of the scale, that was prepared by employing a back-translation procedure (such as that described for the RST-PQ in Study 1). For each item, respondents were asked to report how frequently they experience specific symptoms of anxiety before, during, and after examinations, on a 4-point Likert scale (from *almost never* to *almost always*). Examples of items include: "*Thoughts of doing poorly interfere with my concentration on tests*' (Worry), and "*During tests I feel very tense*" (Emotionality). Cronbach's reliability coefficients in the present sample were .89 for Worry, and .92 for Emotionality.

Results and discussion

Preliminary results

Participants with complete data on the RST-PQ-S were 99.3% at Time 1, and 99.2% at Time 2. Via the Mahalanobis distance, two multivariate outliers (p < .001) were identified and excluded from subsequent analyses. Univariate skewness ranged from −.70 to .71 at Time 1, and from −.57 to .58 at Time 2. Univariate kurtosis ranged from −1.17 to .40 at Time 1, and from −.98 to .13 at Time 2. Mardia's test for multivariate kurtosis was significant at Time 2 (z = 7.30, p < .001), but not at Time 1 (z = 1.44, p = .15).

Internal validity and reliability

The six-factor CFA model was estimated at Time 1 because the available sample is larger. The model yielded adequate fit: $\chi^2(194)$ = 297.09, p <.001, CFI = .920, TLI = .905, RMSEA = .043 (.033, .052), SRMR = .059. Standardized loadings were in the range .39 − .78. Table 5 (left panel) reports omega reliability coefficients and test-retest correlations. Omegas ranged from .60 (FFFS) to .79 (BIS). Four-week test-retest reliability coefficients ranged from .67 (Reward Reactivity) to .82 (BIS). These findings suggest acceptable to good levels of reliability, both in terms of internal consistency and temporal stability.

Overlap with the extended RST-PQ

To assess the degree of overlap between the short (s) and the full (f) scale, we first calculated Pearson's correlations

Table 5. Reliability of the RST-PQ-S and overlap with the full-length scale (Study 3).

	Reliability of the RST-PQ-S		Correlation with the RST-PQ
	φ	r_{s1s2}	r_{f1s2}
FFFS	.60	.79	.76
BIS	.79	.82	.78
BAS-Reward Interest	.77	.70	.65
BAS- Goal Drive-Persistence	.70	.81	.78
BAS-Reward Reactivity	.65	.67	.65
BAS-Impulsivity	.65	.75	.65

Note. φ = omega reliability coefficient; r = Pearson correlation coefficient; s1 = short scale at T1; s2 = short scale at T2; f1 = full scale at T1.

between the two forms at the first measurement occasion (i.e., r_{f1s1}). Coefficients for the six scales were all $\geq$ .84 (M = .89; SD = .04). This appears to suggest a large amount of overlapping variance between the short and full versions. We should consider, however, that correlations might be overestimated, since part of the items are included in both forms (Smith et al., 2000).

Following the guidelines provided by Smith et al. (2000), a more conservative estimation was obtained by calculating the correlations between the full form measured at Time 1 and the short form measured at Time 2 (i.e., r_{f1s2}). Results are reported in the right panel of Table 5. Since these correlations are affected by measurement error, we compared them with test-retest reliability of the short scale (i.e., r_{s1s2}), which provide an upper limit for r_{f1s2}: the smaller the difference between r_{f1s2} and r_{s1s2}, the higher the overlap between the short- and the full-length scales.

As can observed in Table 5, correlations between the full-length scale at Time 1 and the short form at Time 2 were only marginally different from test-retest correlations: Cohen's q for the six scales ranged from .03 to .20 (M = .10; SD = .06). Therefore, we can conclude that there is adequate overlapping variance in the present sample between the two forms.

Criterion validity

The criterion validity of the RST-PQ-S was assessed by examining its relationship with students' feelings of anxiety. Two multiple regressions were conducted, including the six scale scores of the RST-PQ-S at Time 1 as predictors of Worry and Emotionality, respectively. Results, are shown in the left panel of Table 6. The BIS showed a significant relation to the cognitive component of test anxiety, whose manifestations are similar to those related to variations in the BIS functioning (e.g., worry, repetitive thoughts, anxious rumination; Pickering & Corr, 2008). The added contribution of FFFS to worry, by contrast, was not statistically significant. These results are consistent with the conceptual basis of revised RST, which views anxiety and fear as distinct aspects of defensive behavior (DeYoung, 2010; Gray & McNaughton, 2000).

Both the BIS and FFFS were significantly related to the emotional component of test anxiety. This is likely due to the difficulty in disentangling the physiological reactions that accompany anxiety (e.g. increased heart rate, sweating)

Table 6. Standardized regression weights and proportion of variance accounted in the Test Anxiety Inventory for the full- and the short-length RST-PQ (Study 3).

	RST-PQ-S		RST-PQ full	
	Worry	Emotionality	Worry	Emotionality
FFFS	.10	.26*	.13	.31*
BIS	.45*	.38*	.46*	.38*
BAS-Reward Interest	.03	−.09	.02	−.08
BAS- Goal Drive-Persistence	−.07	.04	−.08	.08
BAS-Reward Reactivity	.08	.09	.07	.04
BAS-Impulsivity	.02	−.08	.05	−.09
R^2	.24	.29	.30	.38

Note. * $p < .05$.

from those reflecting FFFS-mediated fear (Gray & McNaughton, 2000; Smilie et al., 2006). The BIS, however, made the largest unique contribution to the explanation of emotionality. This confirmed the prediction, providing further support for the revised RST. The BAS scales were unrelated to both components of test anxiety Taken together, the six scales of the RST-PQ-S accounted for 24% of the variance in Worry, and 29% in Emotionality.

To compare the criterion validity of the short form with that of the full version, the same analysis was replicated by using the full RST-PQ at Time 1 (Table 6, right panel). The pattern of relationship was similar to the one observed for the short form. Cumulatively, the six scales of the full RST-PQ accounted for a higher proportion of variance (30% in Worry, 38% in Emotionality) with respect to the RST-PQ-S scales. However, the expected relation between BIS and test anxiety was substantially the same for the two forms. In sum, although the amount of explained variance was higher for the full RST-PQ, the construct-relevant variance (i.e., the variance shared with constructs the scale is intended to measure) was substantially equivalent for the length and the short form.

Latent state-trait analysis

As a next step, we adopted the Latent State-Trait (LST) framework (Steyer et al., 1992) with the aim of investigating the extent to which scores on the RST-PQ-S items are due to dispositional rather than situational factors. Specifically, we adopted the Indicator-specific Trait (IT) factor model (Eid, 1996). In the IT approach, each item loads on two latent variables: (1) a trait factor, which reflect dispositional (stable) interindividual differences; and (2) a latent state residual factor, which reflects intraindividual differences between occasions of measurement due to the situation. Trait factors were allowed to correlate, whereas the latent state residual factors were assumed to be uncorrelated.

This approach has been preferred to other LST models, since it has shown to perform well under different conditions, and does not require the identification of a specific reference indicator, as in the case of the M − 1 model (see Geiser & Lockhart, 2012). Given the available size of the sample, we tested two separate models, one for the two defensive factors (FFFS/BIS), and one for the BAS. In order to increase the ratio between estimated parameters and number of cases, factor loadings were all fixed to 1.

Model fit was adequate for both FFFS/BIS, $\chi^2(131) = 158.90$, $p = .05$, CFI = .985, TLI = .979, RMSEA = .026 (.002, .039), SRMR = .036, and BAS, $\chi^2(190) = 211.11$, $p = .14$, CFI = .991, TLI = .987, RMSEA = .019 (.000, .032), SRMR = .035. Estimated model parameters have been used to decompose the total variability of each item into three components: *Consistency* (*Con*), which reflects the effect of stable trait factors; *occasion specificity* (*OSpe*), which reflects the effect of the situation; and *random error variance* (Eid & Diener, 2004). Reliability coefficients at the item-level, which reflects individual differences that are due to reliable sources of variance, can be obtained by taking the sum of the *Con* and *OSpe* coefficients (Geiser & Lockhart, 2012).

Results showed that the trait component of variance was substantially larger than the state component. Within each scale, the proportion of reliable variance due to trait and state components averaged, respectively: .68 and .03 for FFFS; .64 and .03 for BIS; .63 and .11 for Reward Interest; .71 and .03 for Goal-Drive Persistence; .57 and .07 for Reward Reactivity; and .61 and .03 for Impulsivity. Reliability coefficients were in the range .62–.78 for FFFS, .63–.71 for BIS, .71–.79 for Reward Interest, .71–.78 for Goal-Drive Persistence, .61–.67 for Reward Reactivity, and .58–.69 for Impulsivity. Results of the decomposition of variance at the item-level are detailed in the Online Supplementary material (Table S4).

Conclusions

Our research aimed to develop and validate a shortened Italian version of the RST-PQ (Corr & Cooper, 2016), a self-report scale for the assessment of the constructs postulated by revised RST (Corr, 2008). The items of the short form were selected from the original, full version. The short form is more parsimonious with respect to the original RST-PQ, in that it has a limited number of items (from three to five for each dimension), while maintaining similar psychometric properties with respect to the full version.

Confirmatory factor analysis in three independent samples supported the expected six-factor structure, comprising FFFS, BIS, and four BAS components (Reward Interest, Goal-Drive Persistence, Reward Reactivity, and Impulsivity). The six factors showed adequate levels of internal consistency throughout the studies, although omega coefficients were less than optimal for the FFFS, reward reactivity, and impulsivity scales, falling in the range of .60 − .70. It should be noted, however, that the overarching goal of the item selection procedure was not to maximize statistical criteria, such as items homogeneity or model fit. The primary interest, by contrast, was to retain the content validity of the original instrument. In this regard, we found a substantial overlap between the short and the full version, as revealed by high correlations between the corresponding scales, as well as by similar pattern and strength of association with several personality measures.

Finally, LST models suggested that the items of the RST-PQ-S mostly captures stable interindividual differences, showing moderate fluctuations across situations. This is the first study to assess state and trait components of RST constructs. These represent fundamental levels of description in the revised theory (Corr & McNaughton, 2008). Examining the behavioral and emotional correlates of FFFS, BIS and BAS at both state and traits levels may represents an important area for future investigations.

Taken together, results from three studies indicate that the RST-PQ-S represents an efficient, valid and reliable measure of RST constructs that can be fruitfully used in several contexts, such as health-related surveys or clinical research, where test length can be an issue. Of course, the short form is not intended to replace the original, full version. Due to its higher reliability, the 65-item RST-PQ is more appropriate for using in individual assessment, such as in applied settings, where test scores can have consequences for individuals.

There are some limitations in our research. First, respondents were recruited using convenience sampling. Second, the criterion validity of the RST-PQ-S was limited to a single construct (i.e., test anxiety). Future studies should examine the criterion validity of the short form with regard to multiple outcome measures, possibly in larger and representative samples. A further potential limitation of the study is that the items of the scale had to be translated into Italian, and we did not have the possibility to test its equivalence with the original, English version. Further studies should test the measurement invariance of the short form across countries and languages. In sum, although much has still to be done to complete the validation process of the RST-PQ-S, we believe that this measure has the potential to be used to good effect in future studies, enhancing the applicability of revised RST theory in empirical research.

Data availability statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

References

Aluja, A., & Blanch, A. (2011). Neuropsychological behavioral inhibition system (BIS) and behavioral approach system (BAS) assessment: A shortened Sensitivity to Punishment and Sensitivity to Reward Questionnaire version (SPSRQ-20). *Journal of Personality Assessment*, *93*(6), 628–636. https://doi.org/10.1080/00223891.2011.608760

Arrindell, W. A. (1980). Dimensional structure and psychopathology correlates of the Fear Survey Schedule (FSS-III) in a phobic population: A factorial definition of agoraphobia. *Behaviour Research and Therapy*, *18*(4), 229–242. https://doi.org/10.1016/0005-7967(80)90080-7

Bacon, A. M., Corr, P. J., & Satchell, L. P. (2018). A reinforcement sensitivity theory explanation of antisocial behaviour. *Personality and Individual Differences*, *123*, 87–93. https://doi.org/10.1016/j.paid.2017.11.008

Beaton, D. E., Bombardier, C., Guillemin, F., & Ferraz, M. B. (2000). Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine*, *25*(24), 3186–3191. https://doi.org/10.1097/00007632-200012150-00014

Beaton, A. A., Mutinelli, S., & Corr, P. J. (2017). Fractionating negative and positive affectivity in handedness: Insights from the reinforcement sensitivity theory of personality. *Laterality*, *22*(4), 419–444. https://doi.org/10.1080/1357650X.2016.1213274

Beck, I., Smits, D. J. M., Claes, L., Vandereycken, W., & Bijttebier, P. (2009). Psychometric evaluation of the Behavioral Inhibition/Behavioral Activation System scales and the Sensitivity to Punishment and Sensitivity to Reward Questionnaire in a sample of eating disordered patients. *Personality and Individual Differences*, *47*(5), 407–412. https://doi.org/10.1016/j.paid.2009.04.007

Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin*, *107*(2), 238–246. https://doi.org/10.1037/0033-2909.107.2.238

Briggs, S. R., & Cheek, J. M. (1986). The role of factor analysis in the development and evaluation of personality scales. *Journal of Personality*, *54*(1), 106–148. https://doi.org/10.1111/j.1467-6494.1986.tb00391.x

Brown, T. A. (2006). *Confirmatory factor analysis for applied research*. Guilford.

Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (Vol. 154, pp. 136–162).Sage Focus Editions.

Carver, C. S., & White, T. L. (1994). Behavioral inhibition, behavioral activation, and affective responses to impending reward and punishment: The BIS/BAS scales. *Journal of Personality and Social Psychology*, *67*(2), 319–333. https://doi.org/10.1037/0022-3514.67.2.319

Chen, F. F., Hayes, A., Carver, C. S., Laurenceau, J.-P., & Zhang, Z. (2012). Modeling general and specific variance in multifaceted constructs: A comparison of the bifactor model to other approaches. *Journal of Personality*, *80*(1), 219–251. https://doi.org/10.1111/j.1467-6494.2011.00739.x

Clark, L. A., & Watson, D. (1995). Constructing validity: Basic issues in objective scale development. *Psychological Assessment*, *7*(3), 309–319. https://doi.org/10.1037/pas0000626

Cohen, J. (1987). *Statistical power analysis for the behavioral sciences*. Academic Press.

Corr, P. J. (2008). Reinforcement sensitivity theory (RST): Introduction. In P. J. Corr (Ed.), *The reinforcement sensitivity theory of personality* (pp. 1–43). Cambridge University Press. https://doi.org/10.1017/CBO9780511819384.002

Corr, P. J. (2013). Approach and avoidance behavior: Multiple systems and their interactions. *Emotion Review*, *5*(3), 285–291. https://doi.org/10.1177/1754073913477507

Corr, P. J. (2016). Reinforcement sensitivity theory of personality questionnaires: Structural survey with recommendations. *Personality and Individual Differences*, *89*, 60–64. https://doi.org/10.1016/j.paid.2015.09.045

Corr, P. J., & Cooper, A. J. (2016). The reinforcement sensitivity theory of personality questionnaire (RST-PQ): Development and validation. *Psychological Assessment*, *28*(11), 1427–1440. https://doi.org/10.1037/pas0000273

Corr, P. J., & McNaughton, N. (2008). RST and personality. In P. J. Corr (Ed.), *The reinforcement sensitivity theory of personality* (pp. 155–187). Cambridge University Press.

Corr, P. J., & McNaughton, N. (2012). Neuroscience and approach/avoidance personality traits: A two stage (valuation-motivation) approach. *Neuroscience and Biobehavioral Reviews*, *36*(10), 2339–2354. https://doi.org/10.1016/j.neubiorev.2012.09.013

De Pascalis, V., & Scacchia, P. (2019). The influence of reward sensitivity, heart rate dynamics and EEG-delta activity on placebo analgesia. *Behavioural Brain Research*, *359*, 320–332. https://doi.org/10.1016/j.bbr.2018.11.014

DeYoung, C. G. (2010). Mapping personality traits onto brain systems: BIS, BAS, FFFS, and beyond. *European Journal of Personality*, *24*, 404–407. https://doi.org/10.1002/per.780

Dissabandara, L. O., Loxton, N. J., Dias, S. R., Daglish, M., & Stadlin, A. (2012). Testing the fear and anxiety distinction in the BIS/BAS scales in community and heroin-dependent samples. *Personality and Individual Differences*, *52*(8), 888–892. https://doi.org/10.1016/j.paid.2012.01.023

Donnellan, M., Oswald, F., Baird, B., & Lucas, R. (2006). The Mini-IPIP scales: Tiny-yet-effective measures of the Big Five factors of personality. *Psychological Assessment*, *18*(2), 192–203. https://doi.org/10.1037/1040-3590.18.2.192

Eid, M. (1996). Longitudinal confirmatory factor analysis for polytomous item responses: Model definition and model selection on the basis of stochastic measurement theory. *Methods of Psychological Research Online*, *1*, 65–85.

Eid, M., & Diener, E. (2004). Global judgments of subjective wellbeing: Situational variability and long-term stability. *Social Indicators Research*, *65*(3), 245–277. https://doi.org/10.1023/B:SOCI.0000

Eriksson, L. J. K., Jansson, B., & Sundin, Ö. (2019). Psychometric properties of a Swedish version of the reinforcement sensitivity theory of personality questionnaire. *Nordic Psychology*, *71*(2), 134–145. https://doi.org/10.1080/19012276.2018.1516563

Floyd, F., & Widaman, K. F. (1995). Factor analysis in the development and refinement of clinical assessment instruments. *Psychological Assessment*, *7*(3), 286–299. https://doi.org/10.1037/1040-3590.7.3.286

Geiser, C., & Lockhart, G. (2012). A comparison of four approaches to account for method effects in latent state-trait analyses. *Psychological Methods*, *17*(2), 255–283. https://doi.org/10.1037/a0026977

Gray, J. A. (1987). *The psychology of fear and stress*. Cambridge University Press.

Gray, J. A., & McNaughton, N. (2000). *The neuropsychology of anxiety: An enquiry into the functions of the septo-hippocampal system*. Oxford University Press.

Hagopian, L. P., & Ollendick, T. H. (1994). Behavioral inhibition and test anxiety: An empirical investigation of Gray's theory. *Personality and Individual Differences*, *16*(4), 597–604. https://doi.org/10.1016/0191-8869(96)00064-5

Haynes, S. N., Richard, D. C. S., & Kubany, E. S. (1995). Content validity in psychological assessment: A functional approach to concepts and methods. *Psychological Assessment*, *7*(3), 238–247. https://doi.org/10.1037/1040-3590.7.3.238

Heene, M., Bollmann, S., & Bühner, M. (2014). Much ado about nothing, or much to do about something? Effects of scale shortening on criterion validity and mean differences. *Journal of Individual Differences*, *35*(4), 245–249. https://doi.org/10.1027/1614-0001/a000146

Heym, N., Ferguson, E., & Lawrence, C. (2008). An evaluation of the relationship between Gray's revised RST and Eysenck's PEN: Distinguishing BIS and FFS in Carver and White's BIS/BAS scales. *Personality and Individual Differences*, *45*(8), 709–715. https://doi.org/10.1016/j.paid.2008.07.013

Hu, L., & Bentler, P. M. (1998). Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. *Psychological Methods*, *3*(4), 424–453. https://doi.org/10.1037/1082-989X.3.4.424

Jackson, C. J. (2009). Jackson-5 scales of revised Reinforcement Sensitivity Theory (r-RST) and their application to dysfunctional real world outcomes. *Journal of Research in Personality*, *43*(4), 556–569. https://doi.org/10.1016/j.jrp.2009.02.007

Jiang, Y., & Tiliopoulos, N. (2014). Individual differences in adult attachment and reinforcement sensitivity. *Personality and Individual Differences*, *68*, 205–210. https://doi.org/10.1016/j.paid.2014.04.022

Johnson, S. L., Turner, R. J., & Iwata, N. (2003). BIS/BAS levels and psychiatric disorder: An epidemiological study. *Journal of Psychopathology and Behavioral Assessment*, *25*(1), 25–36. https://doi.org/10.1023/A:1022247919288

Krupić, D., Corr, P. J., Ručević, S., Križanić, V., & Gračanin, A. (2016). Five reinforcement sensitivity theory (RST) of personality questionnaires: Comparison, validity and generalization. *Personality and Individual Differences*, *97*, 19–24. https://doi.org/10.1016/j.paid.2016.03.012

Kruyen, P. M., Emons, W. H., & Sijtsma, K. (2013). On the shortcomings of shortened tests: A literature review. *International Journal of Testing*, *13*(3), 223–248. https://doi.org/10.1080/15305058.2012.703734

Leone, L., Pierro, A., & Mannetti, L. (2002). Validità della versione italiana delle scale BIS/BAS di Carver e White (1994): Generalizzabilità della struttura e relazioni con costrutti affini. *Giornale Italiano di Psicologia*, *2*, 413–434. https://doi.org/doi: https://doi.org/10.1421/1245

Liebert, R. M., & Morris, L. W. (1967). Cognitive and emotional components of test anxiety: A distinction and some initial data. *Psychological Reports*, *20*(3), 975–978. https://doi.org/10.2466/pr0.1967.20.3.975

Maack, D. J., & Ebesutani, C. (2018). A re-examination of the BIS/BAS scales: Evidence for BIS and BAS as unidimensional scales. *International Journal of Methods in Psychiatric Research*, *27*(2), e1612https://doi.org/10.1002/mpr.1612

McDonald, R. P. (1999). *Test theory: A unified treatment*. L. Erlbaum Associates.

McNaughton, N., & Corr, P. J. (2004). A two-dimensional neuropsychology of defense: Fear/anxiety and defensive distance. *Neuroscience and Biobehavioral Reviews*, *28*(3), 285–305. https://doi.org/10.1016/j.neubiorev.2004.03.005

Moshagen, M. (2012). The model size effect in SEM: Inflated goodness-of-fit statistics are due to the size of the covariance matrix. *Structural Equation Modeling: A Multidisciplinary Journal*, *19*(1), 86–98. https://doi.org/10.1080/10705511.2012.634724

Muthén, L. K., & Muthén, B. (2010). *Mplus: Statistical analysis with latent variables*. Authors.

Pedrabissi, L., & Santinello, M. (1989). *STAI State-Trait Anxiety Inventory Forma Y Manuale*. Organizzazioni Speciali, Firenze.

Pickering, A., & Corr, P. (2008). J.A. Gray's reinforcement sensitivity theory (RST) of personality. In G. Boyle, G. Matthews, & D. Saklofske (Eds.). *The SAGE handbook of personality: Theory and assessment personality measurement and testing* (pp. 239–255). Sage.

Poythress, N. G., Edens, J. F., Landfield, K., Lilienfeld, S. O., Skeem, J. L., & Douglas, K. S. (2008). A critique of Carver and White's (1994) Behavioral Inhibition Scale (BIS) for Investigating Lykken's (1995) Theory of Primary Psychopathy. *Personality and Individual Differences*, *45*(4), 269–275. https://doi.org/10.1016/j.paid.2008.04.014

Pugnaghi, G., Cooper, A., Ettinger, U., & Corr, P. (2018). The psychometric properties of the German language Reinforcement Sensitivity Theory-Personality Questionnaire (RST-PQ). *Journal of Individual Differences*, *39*(3), 182–190. https://doi.org/10.1027/1614-0001/a000262

Rammstedt, B., & Beierlein, C. (2014). Can't we make it any shorter? The limits of personality assessment and ways to overcome them. *Journal of Individual Differences*, *35*(4), 212–220. https://doi.org/10.1027/1614-0001/a000141

Raykov, T. (1997). Scale reliability, Cronbach's coefficient alpha, and violations of essential tau-equivalence with fixed congeneric components. *Multivariate Behavioral Research*, *32*(4), 329–353. https://doi.org/10.1207/s15327906mbr3204_2

Reuter, M., Cooper, A. J., Smillie, L. D., Markett, S., & Montag, C. (2015). A new measure for the revised reinforcement sensitivity theory: Psychometric criteria and genetic validation. *Frontiers in Systems Neuroscience*, *9*, 38. https://doi.org/10.3389/fnsys.2015.00038

Revelle, W., & Zinbarg, R. (2009). Coefficients alpha, beta, omega, and the glb: Comments on Sijtsma. *Psychometrika*, *74*(1), 145–154. https://doi.org/10.1007/s11336-008-9102-z

Robins, R. W., Fraley, R. C., & Krueger, R. F. (Eds.). (2007). *Handbook of research methods in personality psychology*. Guilford Press.

Sandy, C. J., Gosling, S. D., & Koelkebeck, T. (2014). Psychometric comparison of automated versus rational methods of scale abbreviation: An illustration using a brief measure of values. *Journal of Individual Differences*, *35*(4), 221–235. https://doi.org/10.1027/1614-0001/a000144

Sarason, I. G. (1961). Test anxiety and the intellectual performance of college students. *Journal of Educational Psychology*, *52*(4), 201–206. https://doi.org/10.1037/h0049095

Schwarzer, R., Van Der Ploeg, H. M., & Spielberger, C. D. (1987). Test anxiety: An overview of theory and research. In C. D. Spielberger (Eds.), *Advances in test anxiety research* (pp. 3–9). Lawrence Erlbaum Associates Inc.

Sijtsma, K. (2009). On the use, the misuse, and the very limited usefulness of Cronbach's alpha. *Psychometrika*, *74*(1), 107–120. https://doi.org/10.1007/s11336-008-9101-0

Smith, G. T., McCarthy, D. M., & Anderson, K. G. (2000). On the sins of short-form development. *Psychological Assessment*, *12*(1), 102–111. https://doi.org/10.1037//1040-3590.12.1.102

Spielberger, C. D. (1980). *Test anxiety inventory: Preliminary professional manual*. Consulting Psychologists Press.

Spielberger, C. D., Gorsuch, R. L., Lushene, R. E., Vagg, P. R., & Jacobs, G. A. (1983). *Manual for the state-trait anxiety inventory: STAI (Form Y2)*. Consulting Psychologists Press.

Steyer, R., Ferring, D., & Schmidt, M. (1992). States and traits in psychological assessment. *European Journal of Psychological Assessment*, *8*, 79–98.

Tabachnick, B., & Fidell, L. (2001). *Using multivariate statistics*. Allyn & Bacon.

Torrubia, R., Avila, C., Molto, J., & Caseras, X. (2001). The Sensitivity to Punishment and Sensitivity to Reward Questionnaire (SPSRQ) as a measure of Gray's anxiety and impulsivity dimensions. *Personality and Individual Differences*, *31*(6), 837–862. https://doi.org/10.1016/S01918869(00)00183-5

Wolpe, J., & Lang, P. J. (1977). *Manual for the fear survey schedule*. EdITS.

Wytykowska, A., Fajkowska, M., Domaradzka, E., & Jankowski, K. S. (2017). Construct validity of the Polish version of the reinforcement sensitivity theory-personality questionnaire. *Personality and Individual Differences*, *109*, 172–180. https://doi.org/10.1016/j.paid.2016.12.054

Yang, C., Nay, S., & Hoyle, R. H. (2010). Three approaches to using lengthy ordinal scales in structural equation models: Parceling, latent scoring, and shortening scales. *Appl Psychol Meas*, *34*(2), 122–142. https://doi.org/10.1177/0146621609338592

Zeidner, M. (1998). *Test anxiety: The state of the art*. Plenum Press.

Ziegler, M., Kemper, C., & Kruyen, P. (2014). Short scales - Five misunderstandings and ways to overcome them. *Journal of Individual Differences*, *35*(4), 185–189. https://doi.org/10.1027/1614-0001/a000148