Original Article

Validation of the Danish Version of Perseverative Thinking Questionnaire (PTQ) – Introducing the PTQ Short Version

Stine Bjerrum Moeller[1,2], Pia Veldt Larsen[3], Ida-Marie T. P. Arendt[1,2], Thomas Ehring[4], Nina Reinholt[5], Morten Hvenegaard[6], Anne Bryde Christensen[7], and Sidse Arnfred[7,8]

[1]Department of Psychology, University of Southern Denmark, The National Center of Psychotraumatology, Odense, Denmark
[2]Southern Denmark Psychiatry, Department of Trauma and Torture Survivors, Vejle, Denmark
[3]Mental Health Services in the Region of Southern Denmark, Vejle, Denmark
[4]Department of Psychology, LMU Munich, Germany
[5]Research Unit for Psychotherapy and Psychopathology, Region Zealand Mental Health Services, Slagelse, Denmark
[6]Mental Health Center Copenhagen, Copenhagen University Hospital, Denmark
[7]Psychiatric Research Unit, Region Zealand Mental Health Services, Slagelse, Denmark
[8]Department of Clinical Medicine, Faculty of Health, University of Copenhagen, Denmark

Abstract: *Introduction:* The Perseverative Thinking Questionnaire (PTQ) measures repetitive negative thinking and has been translated and validated in several countries. However, the PTQ has not been translated and validated in a Danish clinical sample. The aim of this paper was to evaluate the psychometric properties of the PTQ in a Danish clinical population and to introduce a shorter, 9-item version of the PTQ, the PTQ-9. *Methods:* Participants were recruited from a multicenter randomized clinical trial (N = 251). They all completed the PTQ and World Health Organization 5-item Wellbeing Index (WHO-5). Participants were further assessed with the 6-item Hamilton Anxiety Rating Scale (HAM-A6) and the 6-item Hamilton Depression Rating Scale (HAM-D6). *Results:* The 2nd order factor model with one higher order general factor and three lower order factors showed the best model fit for the PTQ and the PTQ-9. Both versions showed good internal consistency and the expected correlations with the constructs used for validation. Furthermore, using WHO-5 as primary outcome, both the PTQ and PTQ-9 versions were able to discriminate between treatment responders and nonresponders. *Conclusion:* The PTQ and the PTQ-9 showed satisfying psychometric properties in a Danish clinical sample, including sensitivity to change, and could be used to evaluate psychotherapeutic treatment. To minimize the burden for the patients, the PTQ-9 may be recommended for clinical use over the PTQ.

Keywords: repetitive negative thinking, clinimetrics, Perseverative Thinking Questionnaire, transdiagnostic, depression and anxiety disorders

Negative thinking is a natural part of life. It is usually triggered when mood is low or when we have not reached the valued goals we intended to. However, negative thinking is also well-known to be involved in mental distress. In particular, *repetitive* negative thinking has been linked to psychopathology, including depression and anxiety disorders (Harvey et al., 2004).

Repetitive negative thinking (RNT) has been identified as a process of prolonged and recurrent negative thinking about oneself, feelings, personal concerns, and upsetting experiences (Watkins, 2008). Initially, the response styles theory conceptualized RNT as a psychopathological process linked to depression, labeled depressive rumination, and defined it as a process of repetitive thinking about the symptoms, causes, circumstances, meanings, implications, and consequences of depressed mood and distress (Nolen-Hoeksema, 1991). Over the past decade, accumulating evidence has substantiated that RNT is a central transdiagnostic factor in the development and maintenance of a range of psychopathologies and contributes to intensifying and prolonging negative mood states, interference with problem-solving, and increased avoidance behaviors (Watkins & Roberts, 2020).

In the context of several disorders, RNT has been defined in disorder-specific ways (e.g., rumination for depression, worry for generalized anxiety disorder) which does not capture RNT independently of the content of the thinking. The first measure to assess RNT as a

transdiagnostic content-independent construct was the Perseverative Thinking Questionnaire (PTQ; Ehring et al., 2011). In the PTQ, RNT was defined as a style of thinking about one's problem (current, past, or future) or negative experiences (past or anticipated) with three key characteristics: (1) being repetitive, intrusive, and uncontrollable; (2) perceived as unproductive; and (3) capturing mental resources.

Brief Summary of the Validity Status of the PTQ

PTQ was developed in German and English (Ehring et al., 2011) and has been translated to Dutch (Ehring et al., 2012), Polish (Kornacka et al., 2016), French (Devynck et al., 2017), Turkish (Altan-Atalay & Saritas-Atalar, 2018), Persian (Kami et al., 2019), and Spanish (Valencia, 2020).

Factor Structure

In the initial validation study, PTQ was developed in a theory-based way including items covering the full scope of the definition of the construct. In an internet sample, a nonclinical sample, and a clinical sample, the best fitting model had one higher order general factor and three lower order factors (Core RNT Features, Perceived Unproductiveness of RNT, and RNT Capturing Mental Capacity; Ehring et al., 2011). This structure was reproduced in Dutch (Ehring et al., 2012), Turkish (Altan-Atalay & Saritas-Atalar, 2018), Iranian (Kami et al., 2019), and Australian, Dutch, and American (McEvoy et al., 2018) and with students and clinical patients with psychosis in the United Kingdom (Černis et al., 2016) and Iran (Kami et al., 2019). In the study of McEvoy et al. (2018), since the general factor explained 87% of the common variance, the authors concluded that PTQ is essentially a unidimensional construct.

However, others have found better fit for other models. For example, Kornacka et al. (2016) found evidence for a single order three-factor model among Polish students with the same three factors as the original model, but with some items from the original factor structure validation not covered. In a French community sample using exploratory factor analysis, Devynck et al. (2017) validated a bifactor model with one common factor, labeled RNT, and three subfactors (repetitive characteristic of RNT, the intrusiveness of RNT, and the mental resources effect of RNT) composed of 10 items. In that study, the model was validated in a mixed clinical sample as well.

Reliability

Internal consistency has been tested by all the above studies with Cronbach's α scores for the different subscales and populations ranging between α = .64–.97, while a 4-week test–retest has been explored with Pearson correlations ranging between r = .66–.69 (Ehring et al., 2011), r = .60–.67 (Altan-Atalay & Saritas-Atalar, 2018), and a 3-week test–retest with r = .72–.81 (Kami et al., 2019).

Validity

In the original validation study, the PTQ Total score was found to correlate positively with worry r = .70 (Penn State Worry Questionnaire), rumination r = .72 (Response Styles Questionnaire; RSQ), anxiety r = .64 (State Trait Anxiety Inventory), and depression r = .54 (Beck Depression Inventory; Ehring et al., 2011). Studies have consistently shown markedly higher correlation between PTQ and the brooding subscale of the RSQ than with the reflection subscale (Devynck et al., 2017; Ehring et al., 2011, 2012). This pattern has been repeated in the other validation studies mentioned above. Furthermore, PTQ has shown to differentiate between clinical and nonclinical samples among patients with persecutory delusions compared with healthy volunteers (Černis et al., 2016) and inpatients with psychosis compared to the general population (Kami et al., 2019).

Finally, there is evidence that PTQ is sensitive to change; for example, in their prevention trial, Topper et al. (2017) found that PTQ scores were significantly reduced following rumination-focused preventive interventions (d > .62 at 3 m follow-up) but not in a control condition not receiving any intervention. However, a more systematic investigation of the sensitivity to change issue is warranted.

The Current Study

We wanted to focus particularly on a clinical population and on the sensitivity to change in a Danish validation of the PTQ. Moreover, when developing self-report measures, it is worthwhile attempting to compress the number of items, all while retaining psychometric properties and clinical validity (Cella et al., 2013). Limiting the burden on participants regarding completion time and cognitive load may increase the possibility that participants want to participate in research projects. Moreover, instruments may well be better suited for clinical practice when they are made brief and concise. To aid clinicians to routinely use measures in their practice, brief rating scales are specifically recommended (Aboraya et al., 2018). As the factor structure of the PTQ has been replicated several times in earlier studies, and the PTQ-core

subscale of the PTQ contains concomitant items, a shortened version omitting matching items from the PTQ-core subscale seems to be a reasonable alternative strategy for shortening the scale rather than using exploratory factor analyses, as in Devynck et al. (2017).

Therefore, the current study assessed the factor structure of the PTQ in a clinical population of patients with a primary DSM-5 diagnosis of either social anxiety disorder, panic disorder, agoraphobia, or major depressive disorder, receiving group cognitive behavior therapy in three hospital-based outpatient mental health clinics. We investigated the validity of existing factor models, as well as the shortened version of the PTQ with the same factor structure as the original scale, and further explored the psychometric properties of both versions by investigating the internal reliability, convergent validity, and sensitivity to change.

Method

Measures

The PTQ (Ehring et al., 2011) is a scale constructed to measure repetitive negative thinking across mental disorders. It consists of 15 items evaluating (1) the core characteristics of RNT, identified as the repetitiveness of RNT, the intrusiveness of RNT, and the difficulty of disengaging from RNT (PTQ-core with nine items); (2) the perceived unproductiveness of RNT (PTQ-unproductiveness with three items); and (3) RNT capturing of mental resources (PTQ-mental capacity with three items).

The response to each item is given using a 5-point Likert scale from 0 = *never* to 4 = *almost always*. A higher score reflects a higher level of repetitive negative thinking. Two independent forward translations of the PTQ into Danish were made by two clinical researchers, Morten Hvenegaard (MH) and Stine Bjerrum Moeller (SBM). A reconciled version was created. The reconciled version was translated back to original language (English) by a native English-speaking researcher and clinical psychologist. The back-translation was compared to the original English version PTQ, and minor language corrections were made in the Danish version, ensuring semantic correspondence. The Danish version was tested by four patients in treatment for anxiety, psychotic disorder, and depression by MH. Patients' feedback was taken into consideration when deciding on the final version.

Nine items for the shortened version were selected based on clinical judgment, aiming for the shortened scale to retain the same factor structure as the full PTQ: The original PTQ-core subscale includes nine items covering the three themes of repetitiveness of negative thinking, intrusiveness of negative thinking, and difficulty disengaging from negative thinking (as described above). One item was selected from each theme (Item 1, 3, 7), as the other six items in the PTQ-core subscale (Item 2, 6, 8, 11, 12, 13) were very similar to the selected items, thus seen to add a limited amount of new information (e.g., Item 2: "Thoughts intrude into my mind," vs. Item 7: "Thoughts come to my mind without me wanting them to"). No changes were made to the PTQ-unproductive and PTQ-mental capacity subscales.

The World Health Organization-5 (WHO-5; Topp et al., 2015) is a rating scale measuring well-being, consisting of five positively phrased items scored from 5 = *all of the time* to 0 = *none of the time*. In a systematic review, Topp et al. (2015) found good construct validity and clinical validity covering the dimension of subjective well-being across various settings and diagnoses and sensitivity in regard to being able to capturing improvement in well-being. The authors concluded that the scale is among the most widely used questionnaires assessing subjective psychological well-being as well as a sensitive screening tool for depression. For relevant clinical change, we used the same criteria as Reinholt et al. (2022), identifying a change score of 10 scale points as a minimal clinically important difference.

The 6-item Hamilton Anxiety Rating Scale (HAM-A; Bech, 1981) and the 6-item Hamilton Depression Rating Scale (HAM-D; Carrozzino et al., 2020) are observer-rated instruments measuring anxiety and depression symptoms. Psychology students were trained and supervised in administering the instruments by telephone interview.

Sample

We extracted data (N = 251) from a multicenter, single-blinded, parallel, noninferiority, randomized clinical trial, comparing two types of cognitive behavioral therapy (CBT): disorder-specific group therapy and transdiagnostic group therapy for patients with depression and anxiety disorders (Arnfred et al., 2017). The study was registered on the ClinicalTrials.gov website (ID NCT02954731). Data of all participants from this trial with baseline data on the PTQ were included (see Reinholt et al., 2022) for details on procedures within the trial.

The trial had the following inclusion criteria: (1) a principal DSM-5 diagnosis of depression (MDD; single episode or recurrent), social anxiety disorder, or agoraphobia/panic disorder (Ag/PD); (2) age 18–65 years; and (3) sufficient knowledge of the Danish language. Exclusion criteria were (1) risk of suicide evaluated as high or moderate according to clinicians or assessment researchers, (2) an eating disorder with a body mass index < 18, (3) bipolar disorder, (4) alcohol or drug dependency, or (5) diagnosed with a cluster A or B (DSM-5) personality disorder.

Statistical Analyses

Descriptive statistics on participant characteristics were computed. We used baseline data on the full sample (N = 251) for all analyses except the sensitivity to change-analysis which used end-of-treatment data (N = 151) in addition to baseline data. Analyzing the full sample including both intervention arms was possible as the parent study showed no significant differences between the two interventions on the primary or secondary outcomes at end of treatment and further since neither of the interventions (standard and transdiagnostic group CBT) specifically targeted repetitive negative thinking (Reinholt et al., 2022). The PTQ comprises 15 items and three factors; even with low levels of communality, a sample size of N = 251 is sufficient to obtain an excellent agreement between sample and population solutions (coefficient of congruence $K \geq 0.98$; Mundfrom et al., 2005).

For descriptive purposes, Spearman correlations (ρ) were calculated between all pairs of the 15 PTQ items measured on a 5-point Likert scale. A number of different factor structures of the original 15 PTQ items were assessed using confirmatory factor analyses (CFAs) with Satorra–Bentler corrected maximum likelihood estimators (Satorra & Bentler, 1994). First, three factor models with no correlated uniquenesses were considered: $M01_{15}$, a one-factor model with no correlated uniquenesses; $M02_{15}$, a three factor model with no correlated uniquenesses; and $M03_{15}$, a second order model with one higher order general factor and three lower order factors (in the following abbreviated as 1:3 factors) with no correlated uniquenesses. The three subfactors comprised PTQ-core, PTQ-unproductiveness, and PTQ-mental capacity. Second, we considered three factor models allowing for correlating uniquenesses: $M1_{15}$, a one-factor model with seven correlated uniquenesses; $M2_{15}$, a three factor model with five correlated uniquenesses; and $M3_{15}$, a second order model with 1:3 factors with five correlated uniquenesses. The factor models $M1_{15}$–$M3_{15}$ are shown in Supplementary Figure 1 (https://zenodo.org/record/8340719). Uniqueness is defined as 1 – communality, and the correlated uniquenesses were identified by a judgmental approach combined with assessing model fit and modification indices. Only clinically relevant correlations were included, only if the modification index of the correlated uniqueness exceeded 10, and if including the correlated uniqueness in the model improved the model fit in terms of increase in comparative fit index, decrease in root-mean-square error of approximation (RMSEA), and decrease in Akaike's information criteria (AIC). All CFAs were conducted using structural equation models (SEM); the final models were conducted using SEM with Satorra–Bentler standard errors.

Similarly, the following factor structures of the proposed 9-item PTQ short version were assessed: $M01_9$, a one-factor model with no correlated uniquenesses; $M1_9$, a one-factor model with one correlated uniqueness; $M2_9$, a three-factor model with no correlated uniquenesses; and $M3_9$, a second order model with 1:3 factors with no correlated uniquenesses. The factor models, $M1_9$–$M3_9$, are shown in Supplementary Figure 2 (https://zenodo.org/record/8340719).

The goodness of fit of all above factor structures were assessed by means of Satorra–Bentler scaled χ^2 test statistic (χ^2_{SB}, cutoff: $\chi^2_{SB}/df < 3$), Satorra–Bentler scaled root-mean-square error of approximation ($RMSEA_{SB}$, cutoff: ≤0.06), RMSEA (90% CI; RMSAS cutoff: ≤0.06), standardized root-mean-square residual (SRMR, cutoff: ≤0.08), Satorra–Bentler scaled comparative fit index (CFI_{SB}, cutoff: ≥ 0.95; Shi et al., 2019), and AIC (lowest AIC preferred; Akaike, 1998). Cronbach's α (Novick & Lewis, 1967) was used to assess the internal consistency of all factors and subfactors, and the measurement precisions of the total scales of the 15-item PTQ version and 9-item PTQ short version were compared using the relative 95% confidence interval (CI) of the mean, i.e. the ratio between the width of the 95% CI of the mean and the scale length (Kruyen et al., 2013).

Construct validity was evaluated by examining correlations between three depression and anxiety scales (WHO-5, HAM-D, and HAM-A) and each of the factor scales (PTQ Total, PTQ-core, PTQ-unproductiveness, and PTQ-mental capacity) based on the original 15-item PTQ as well as the 9-item PTQ short version. Spearman correlations were used to account for slight left skewness of the PTQ subscales in the sample (test of skewness: all $p < .05$; D'Agostino et al., 1990).

The parent study used a change score of 10 scale points on the WHO-5 as recommended by Topp et al. (2015) as a minimal clinically important difference. Therefore, to test PTQ's sensitivity to change for each of the four PTQ scales in both the 15-item PTQ version and the 9-item PTQ short version, independent t tests were used to compare the change score from baseline to follow-up between patients with clinical improvement in WHO-5 of at least 10 scale points and patients with less than 10 scale points' improvement. Cohen's d was calculated for each of the sensitivity to change comparisons (Cohen, 1988).

All analyses were performed in STATA 17.0 (Statacorp, Texas, USA).

Results

The sample consisted of 251 patients, predominantly female, single, and predominantly out of current work or education (patient characteristics are shown in Table 1). Notably, a majority of the sample was younger than

30 years. Primary diagnoses were depression, social anxiety, and panic disorder/agoraphobia, with a mean duration of almost 5 years. The *M* (*SD*) of the 15-item PTQ Total was 40.6 (10.0), which is within 1 *SD* of mean scores reported for patients suffering from depression (PTQ Total: $M = 37.56$, $SD = 9.99$) or anxiety disorders (PTQ Total: $M = 35.93$, $SD = 13.60$; Ehring et al., 2011), as well as in patients with psychosis (PTQ Total = 44.6, $SD = 9.7$; Černis et al., 2016). For the 9-item PTQ Total score, the *M* was 24.4 (5.9).

Assessment of Factor Structures

The pairwise correlations between the original 15-item PTQ-items within each factor were fair to moderate (Spearman correlations between $r = .40$ and $r = .70$, apart from items 9 and 14 in the PTQ-unproductiveness factor with $\rho = 0.25$; see Supplementary Table 1 at https://zenodo.org/record/8340719). The correlations between the factors were moderate between $r = .53$ and $r = .64$ (Supplementary Tables 2a [15-item PTQ] and 2b [9-item PTQ]).

The standardized factor loadings for the PTQ Total factor and each of the subfactors PTQ-core, PTQ-unproductiveness, and PTQ-mental capacity were substantial and statistically significant (all loadings ≥ 0.45, all $p < .001$) for both the 15-item PTQ and the 9-item PTQ short version (see Supplementary Tables 3a–3b at https://zenodo.org/record/8340719).

The CFA goodnesses of fit of the various factor models of the original 15-item PTQ with no correlated uniquenesses were unsatisfactory, but when allowing for five correlated uniquenesses, the model fits of the 3-factor and the 2nd order 1:3 factor model were acceptable, with the 2nd order model having the best fit. The 1-factor model was only close to acceptable when allowing for seven correlated uniquenesses (Table 2 and Supplementary Figure 1 at https://zenodo.org/record/8340719). For the 9-item PTQ short version, the goodness of fit of the 1-factor model was unsatisfactory even when allowing for one correlated uniqueness; however, the model fits of both the 3-factor model and the 2nd order 1:3 factor model were good (Table 2 and Supplementary Figure 2 at https://zenodo.org/record/8340719).

Measurement precision of the original and short version of the PTQ total showed no differences in the width of the 95% CI divided by scale length between the original and short version of the PTQ (see Supplementary Table 5 at https://zenodo.org/record/8340719).

Table 1. Patient characteristics

Characteristic	*n* (%)
Gender	
Men	94 (37.5)
Women	157 (62.6)
Age in years	
18–29	138 (55.6)
30–41	60 (24.2)
42–66	65 (20.2)
Education level	
Primary or secondary school	163 (64.9)
Vocational or university	88 (35.1)
Occupation	
Full or part time or student	95 (38.3)
Sick leave for more than 3 months	103 (41.5)
Other	50 (20.2)
Marital status	
Cohabiting	99 (39.9)
Living alone or other	149 (60.1)
Diagnosis	
Depression	120 (47.8)
Social anxiety	73 (29.1)
Panic disorder/agoraphobia	58 (23.1)
Duration of diagnosis in months, *M* (SD)	57.3 (97.4)

Note. $N = 251$; *SD* = standard error.

Internal Consistency of the PTQ Scales

For both the 15-item and the 9-item PTQ versions, McDonald's ω were between .76 and .92 (Cronbach's α .73–.92) for the PTQ Total, the PTQ-core scale, and the PTQ-mental capacity scale, while McDonald's ω was only .65 (Cronbach's $\alpha = 0.62$) for the PTQ-unproductiveness scale (see Supplementary Table 4 at https://zenodo.org/record/8340719).

Construct Validity and Sensitivity to Change of the PTQ Scales

The PTQ in both the original and the shortened version showed the expected and consistent correlations with the constructs used for validation, namely subjective well-being (negative correlations ranging between −.31 and −.41), depression (positive correlations ranging between .30 and .40), and anxiety (positive correlations ranging between .27 and .34) (Table 3).

Comparing the clinically improved patients (at least 10 scale point improvement on the WHO-5) on the PTQ with nonimproved patients (less than 10 scale point improvement on the WHO-5) revealed that PTQ could

Table 2. CFA model fit

Tested models	χ^2_{SB} (*df*)	$RMSEA_{SB}$	RMSEA (90%CI)	SRMR	CFI_{SB}	AIC
Based on original 15-item PTQ						
1-factor model, no correlated uniquenesses	369 (90)	0.111	0.123 (0.111, 0.134)	0.070	0.839	8,684.1
1-factor model, 7 correlated uniquenesses[a]	176 (83)	0.067	0.075 (0.062, 0.089)	0.051	0.947	8,470.5
3-factor model, no correlated uniquenesses	302 (87)	0.099	0.110 (0.098, 0.122)	0.065	0.876	8,609.8
3-factor model, 5 correlated uniquenesses[b]	165 (82)	0.063	0.072 (0.058, 0.085)	0.048	0.952	*8,459.7*
2nd order model 1:3 factors, no correlated uniquenesses	302 (87)	0.099	0.110 (0.098, 0.122)	0.065	0.876	8,609.8
2nd order model 1:3 factors, 5 correlated uniquenesses[c]	160 (82)	0.062	0.070 (0.056, 0.084)	0.046	0.955	**8,454.6**
Based on 9 item PTQ short version						
1-factor model, no correlated uniquenesses	91 (27)	0.097	0.105 (0.084, 0.127)	0.056	0.910	5,429.4
1-factor model, 1 correlated uniqueness[d]	71 (26)	0.083	0.090 (0.067, 0.113)	0.050	0.937	5,407.7
3-factor model, no correlated uniquenesses	48 (24)	0.063	0.070 (0.045, 0.093)	0.041	0.967	**5,386.9**
2nd order model 1:3 factors, no correlated uniquenesses	48 (24)	0.063	0.070 (0.045, 0.093)	0.041	0.967	**5,386.9**

Note. χ^2_{SB} (*df*) = Satorra–Bentler scaled χ^2 test statistic (degrees of freedom); $RMSEA_{SB}$ = Satorra–Bentler scaled root-mean-square error of approximation; RMSEA = root-mean-square error of approximation (90% CI); SRMR = standardized root-mean-square residual; CFI_{SB} = Satorra–Bentler scaled comparative fit index; AIC = Akaike's information criterion; uniqueness = 1 – communality. [a]Allowing for correlated uniquenesses between Items 2 & 7, Item 8 & 9, Items 10 & 15, Items 1 & 6, Items 2 & 12, Items 7 & 12, Items 5 & 10. [b]Allowing for correlated uniquenesses between Items 2 & 7, Item 8 & 9, Items 1 & 6, Items 2 & 12, Items 7 & 12. [c]Allowing for correlated uniquenesses between Items 2 & 7, Item 8 & 9, Items 2 & 12, Items 7 & 12 and between Item 9 and the PTQ core factor. [d]Allowing for correlated uniquenesses between Items 1 & 3. 1. When not allowing for correlated uniquenesses, the 3-factor model and the 2nd order model with three subfactors are reparameterisations of one another and therefore have identical fit. 2. AICs can only be compared between models based on the same items. 3. For each of the two item sets, the minimum AIC is highlighted in bold face, while AICs less than 10 units above the minimum. AIC are highlighted in italics.

Table 3. Construct validity: Spearman correlations (ρ) of the PTQ with measures of depression and anxiety (all $p < .001$)

	Original 15 item PTQ				9-item PTQ short version			
Scale	Total	Core	Unproductiveness	Mental capacity	Total	Core	Unproductiveness	Mental capacity
WHO-5	−.40	−.36	−.38	−.31	−.41	−.34	−.38	−.31
HAM-D-6	.40	.35	.37	.32	.40	.30	.37	.32
HAM-A-6	.34	.30	.28	.30	.34	.27	.28	.30

Note. WHO-5 = World Health Organization 5-item Wellbeing Index; HAM-A-6 = 6-item Hamilton Anxiety Rating Scale; HAM-D-6 = 6-item Hamilton Depression Rating Scale.

discriminate between clinically improved and not improved patients. The shortened scale PTQ-9 showed similar results to the original scale (Table 4). For both scales, all effect sizes (Cohen's *d*) were above 0.6 indicating medium to large effect size.

Discussion

Aiming to provide evidence for the Danish validation of the PTQ in a clinical sample of patients with anxiety and depression, we found similar model fit as previous studies and good convergence validity and sensitivity to change using the WHO-5 as primary outcome. Moreover, testing a shortened version of the scale with the same factor structure provided better model fit and similar findings on measurement precision, convergence validity, and sensitivity to change.

Our study found mean scores of the PTQ comparable to mean scores reported from other studies with patients suffering from depression and anxiety and higher than nonclinical samples, e.g. undergraduate students (Ehring et al., 2011;[1] McEvoy et al., 2018). Therefore, evidence across studies suggests that PTQ scores are clinically informative within depressive and anxiety disorders.

In line with previous studies, we found that the best fitting model for the original 15-item PTQ was the 2nd order 1:3 factor model (Altan-Atalay & Saritas-Atalar, 2018; Černis et al., 2016; Ehring et al., 2012; Kami et al., 2019; McEvoy et al., 2018), although only when allowing for five correlated uniquenesses. However, unlike McEvoy et al. (2018), we found that the model fit of the one factor model

[1] These descriptives are taken from a corrigendum published by (McEvoy et al., 2021) as there were errors in the mean PTQ scores reported in the original manuscript.

Table 4. Sensitivity to change: change in PTQ according to clinical improvement in WHO-5 of at least 10 scale points

	WHO-5 improvement < 10, *n* = 61			WHO-5 improvement ≥ 10, *n* = 90			Difference of differences, *n* = 151		
PTQ scale	PTQ at baseline *M* (*SD*)	PTQ at follow-up *M* (*SD*)	Difference (follow-up - baseline) *M* (95% CI)	PTQ at baseline *M* (*SD*)	PTQ at follow-up *M* (*SD*)	Difference (follow-up - baseline) *M* (95% CI)	(Patients with WHO-5 improvement ≥ 10 – patients with WHO-5 improvement < 10) *M* (95% CI)	*p*	Cohen's *d*
Original 15 item PTQ									
Total (range: 0–60)	40.8 (10.1)	37.4 (11.2)	−3.4 (−6.3, −0.4)	41.2 (9.5)	28.7 (10.8)	−11.9 (−14.2, −9.7)	−8.6 (−12.2, −4.9)	<.001	−0.77
Core (range: 0–36)	25.2 (6.5)	23.5 (7.2)	−1.8 (−3.6, −0.1)	25.5 (6.0)	18.5 (6.8)	−6.8 (−8.2, −5.4)	−5.1 (−7.3, −2.8)	<.001	−0.74
Unproductive (range: 0–12)	8.1 (2.3)	7.4 (2.4)	−0.7 (−1.3, −0.1)	8.1 (2.3)	5.7 (2.4)	−2.2 (−2.8, −1.7)	−1.6 (−2.4, −0.8)	<.001	−0.64
Mental (range: 0–12)	7.4 (2.4)	6.5 (2.6)	−1.0 (−1.7, −0.2)	7.6 (2.4)	4.6 (2.5)	−2.9 (−3.5, −2.3)	−1.9 (−2.9, −0.9)	<.001	−0.65
9-item PTQ short version									
Total (range: 0–12)	24.3 (6.0)	22.1 (6.7)	−2.1 (−3.9, −0.4)	24.6 (5.7)	16.7 (6.5)	−7.6 (−8.9, −6.2)	−5.4 (−7.6, −3.2)	<.001	−0.81
Core (range: 0–12)	8.7 (2.3)	8.1 (2.6)	−0.5 (−1.2, 0.1)	8.9 (2.2)	6.4 (2.4)	−2.5 (−3.0, −1.9)	−1.9 (−2.7, −1.2)	<.001	−0.81
Unproductive (range: 0–12)	8.1 (2.3)	7.4 (2.4)	−0.7 (−1.3, −0.1)	8.1 (2.3)	5.7 (2.4)	−2.2 (−2.8, −1.7)	−1.6 (−2.4, −0.8)	<.001	−0.64
Mental (range: 0–12)	7.4 (2.4)	6.5 (2.6)	−1.0 (−1.7, −0.2)	7.6 (2.4)	4.6 (2.5)	−2.9 (−3.5, −2.3)	−1.9 (−2.9, −0.9)	<.001	−0.65

Note. *n* = number of observations; *SD* = standard error; 95% CI = 95% confidence interval; Cohen's *d*: small effect size: *d* = 0.2; medium effect size: *d* = 0.5; large effect size: *d* = 0.8; PTQ = Perseverative Thinking Questionnaire; WHO-5 = World Health Organization 5-item Wellbeing Index.

was not quite acceptable, and AIC indicated that the one factor model was inferior to the 2nd order 1:3 factor model, even when allowing for up to seven correlated uniquenesses. As it may prove problematic for a questionnaire if multiple versions with varying factor structure exist, it is worth noting that the PTQ-9 has the same factor structure with all items loading on the same factors as the original scale. Our only modification was that we used clinical inspection to remove redundant items from the PTQ-core subscale, which had more items that the other subscales, and tested the shortened 9-item scale for model fit and clinical utility compared with the original 15-item scale. For the simpler 9-item PTQ short version, the 2nd order 1:3 factor model provided a good fit to the data, while the model fit of the one-factor model was not acceptable. Providing an advantage over the original model, the 2nd order 1:3 factor model on the 9-item PTQ did not need correlated uniquenesses. Pairwise correlations between PTQ-items within each factor were fair to moderate, excepting one pair in the PTQ-unproductiveness factor with low correlation. A key benefit of the 9-item PTQ short version over the original model is that it maintains a similar level of measurement precision. This suggests that the condensation of the scale does not result in a loss of accuracy in identifying repetitive negative thinking. Thus, the short version offers the advantage of brevity without compromising the instrument's ability to accurately capture the intended psychological construct which is of paramount importance in clinical practice (Kemper et al., 2019).

Reliability scores were high, apart from the PTQ-unproductiveness scale (ω = .65, α = .62). Other validation studies also report the PTQ-unproductiveness scale having the lowest internal consistency, but none lower than α = .73 (Kami et al., 2019). Together, these results suggest some potential problems in the PTQ-unproductiveness subscale. Acknowledging that RNT is associated with both positive and negative metabeliefs (Kubiak et al., 2014; McEvoy & Mahoney, 2013; Weber & Exner, 2013), the patient who may be undecisive about whether RNT is productive or not, it might find it an ambivalent task to report on the unproductiveness of RNT.

We found similar patterns of correlations with subjective well-being, anxiety, and depression for the original 15 items PTQ and the 9-item shortened version, corroborating the construct validity of the shortened 9-item PTQ. The correlations between PTQ Total and subjective well-being, depression, and anxiety substantiated the validity of the scale. However, it was lower than in the first validation study using both clinical and nonclinical samples which showed no differences in the pattern of relationship between PTQ and anxiety/depression across sample type (Ehring et al., 2011) and in the general population study in Iran (Kami et al., 2019). However, these studies used self-report

assessment for anxiety and depression, while our study used clinician reported assessment by use of telephone interview.

We found robust evidence of sensitivity to change as the PTQ with medium to large effect sizes was able to discriminate between the group of patients identified as treatment responders and the other group of patients identified as nonresponders using WHO-5 as primary outcome in both the original 15-item version and the 9-item shortened version. These results corroborate the usefulness of the PTQ in psychotherapy research examining the processes of treatments specifically targeting RNT.

The use of only one sample in our study limits the generalizability of the psychometric properties across different respondent groups and settings. Moreover, the limited number of variables available to examine the convergent and discriminant validity of the translated PTQ poses a restriction on the exploration of the construct validity. That the study used blind observer-rated measures to assess depression and anxiety is a strength as it ammends the inherent risk of common method variance, when a study is comprised entirely of self-report measures. Also, it is a strength that the study sample was sufficiently large to conduct the factor analysis of the PTQ factor structure.

Conclusion

For the original 15-item PTQ, the 2nd order 1:3 factor model has the best model fit. The PTQ showed sensitivity to change and can therefore be used to evaluate treatment. Notably, as the 9-item PTQ showed psychometric properties and clinical sensitivity similar to, or even a little better, than the original 15-item PTQ, we recommend using the shortened PTQ to burden patients the least. Since the PTQ is a scale used in clinical settings, it would be valuable to produce cross-cultural norms for different target groups (e.g., general population and clinical groups) to guide clinical interpretation of PTQ scores.

References

Aboraya, A., Nasrallah, H. A., Elswick, D. E., Ahmed, E., Estephan, N., Aboraya, D., Berzingi, S., Chumbers, J., Berzingi, S., Justice, J., Zafar, J., & Dohar, S. (2018). Measurement-based care in psychiatry—Past, present, and future. *Innovations in Clinical Neuroscience*, *15*(11-12), 13–26.

Akaike, H. (1998). A new look at the statistical model identification. In E. Parzen, K. Tanabe, & G. Kitagawa (Eds.), *Selected papers of Hirotugu Akaike* (pp. 215–222). Springer.

Altan-Atalay, A., & Saritas-Atalar, D. (2018). Psychometric qualities of Turkish version of Perseverative Thinking Questionnaire (PTQ). *Journal of Rational-Emotive & Cognitive-Behavior Therapy*, *36*(3), 252–266. https://doi.org/10.1007/s10942-018-0285-7

Arnfred, S. M., Aharoni, R., Hvenegaard, M., Poulsen, S., Bach, B., Arendt, M., Rosenberg, N. K., & Reinholt, N. (2017). Transdiagnostic group CBT vs. standard group CBT for depression, social anxiety disorder and agoraphobia/panic disorder: Study protocol for a pragmatic, multicenter non-inferiority randomized controlled trial. *BMC Psychiatry*, *17*(1), Article 37. https://doi.org/10.1186/s12888-016-1175-0

Bech, P. (1981). Rating scales for affective disorders: Their validity and consistency. *Acta Psychiatrica Scandinavica*, *64*(Suppl 295), 11–101. https://doi.org/10.1111/j.1600-0447.1981.tb09221.x

Carrozzino, D., Patierno, C., Fava, G. A., & Guidi, J. (2020). The Hamilton rating scales for depression: A critical review of clinimetric properties of different versions. *Psychother Psychosom*, *89*(3), 133–150. https://doi.org/10.1159/000506879

Cella, D., Hahn, E. A., Jensen, S. E., Butt, Z., Nowinski, C. J., & Rothrock, N. (2013). *Methodological issues in the selection, administration and use of patient-reported outcomes in performance measurement in health care settings*. [Paper presentation]. National Quality Forum (NQF): Patient Reported Outcomes (PROs) in Performance Measurement at Northwestern University, IL, USA.

Černis, E., Dunn, G., Startup, H., Kingdon, D., Wingham, G., Evans, N., Lister, R., Pugh, K., Cordwell, J., Mander, H., & Freeman, D. (2016). The Perseverative Thinking Questionnaire in patients with persecutory delusions. *Behavioural and Cognitive Psychotherapy*, *44*(4), 472–481. https://doi.org/10.1017/s1352465815000533.

Cohen, J. (1988). *Statistical power analysis for the behavioural sciences*. Routledge.

D'Agostino, R. B., Belanger, A., & D'agostino, R. B., Jr (1990). A suggestion for using powerful and informative tests of normality. *The American Statistician*, *44*(4), 316–321. https://doi.org/10.1080/00031305.1990.10475751

Devynck, F., Kornacka, M., Baeyens, C., Serra, É., Neves, J. F. D., Gaudrat, B., Delille, C., Taquet, P., Depraete, O., Tison, P., Sgard, F., Rousseau, A., & Romo, L. (2017). Perseverative thinking questionnaire (PTQ): French validation of a transdiagnostic measure of repetitive negative thinking. *Frontiers in Psychology*, *8*, Article 2159. https://doi.org/10.3389/fpsyg.2017.02159

Ehring, T., Raes, F., Weidacker, K., & Emmelkamp, P. M. (2012). Validation of the Dutch version of the Perseverative Thinking Questionnaire (PTQ-NL). *European Journal of Psychological Assessment*, *28*(2), 102–108. https://doi.org/10.1027/1015-5759/a000097

Ehring, T., Zetsche, U., Weidacker, K., Wahl, K., Schönfeld, S., & Ehlers, A. (2011). The Perseverative Thinking Questionnaire (PTQ): Validation of a content-independent measure of repetitive negative thinking. *Journal of Behavior Therapy and Experimental Psychiatry*, *42*(2), 225–232. https://doi.org/10.1016/j.jbtep.2010.12.003

Harvey, A. G., Watkins, E., & Mansell, W. (2004). *Cognitive behavioural processes across psychological disorders: A transdiagnostic approach to research and treatment*. Oxford University Press.

Kami, M., Moloodi, R., Mazidi, M., Ehring, T., Mansoori, A. K., Nodooshan, M. B., Mazinani, Z., Molavi, M., & Momeni, F. (2019). Measuring repetitive thinking in Iran: Psychometric properties of Persian version of Perseverative Thinking Questionnaire. *Personality and Individual Differences*, *148*, 101–109. https://doi.org/10.1016/j.paid.2019.05.012

Kemper, C. J., Trapp, S., Kathmann, N., Samuel, D. B., & Ziegler, M. (2019). Short versus long scales in clinical assessment: Exploring the trade-off between resources saved and psychometric quality lost using two measures of obsessive-compulsive symptoms. *Assessment*, *26*(5), 767–782. https://doi.org/10.1177/1073191118810057

Kornacka, M., Buczny, J., & Layton, R. L. (2016). Assessing repetitive negative thinking using categorical and transdiagnostic approaches: A comparison and validation of three polish language adaptations of self-report questionnaires. *Frontiers in Psychology*, *7*, Article 322. https://doi.org/10.3389/fpsyg.2016.00322

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

Kubiak, T., Zahn, D., Siewert, K., Jonas, C., & Weber, H. (2014). Positive beliefs about rumination are associated with ruminative thinking and affect in daily life: Evidence for a metacognitive view on depression. *Behavioural and Cognitive Psychotherapy*, *42*(5), 568–576. https://doi.org/10.1017/s1352465813000325

McEvoy, P. M., Hyett, M. P., Ehring, T., Johnson, S. L., Samtani, S., Anderson, R., & Moulds, M. L. (2018). Transdiagnostic assessment of repetitive negative thinking and responses to positive affect: Structure and predictive utility for depression, anxiety, and mania symptoms. *Journal of Affective Disorders*, *232*, 375–384. https://doi.org/10.1016/j.jad.2018.02.072

McEvoy, P. M., Hyett, M. P., Ehring, T., Johnson, S. L., Samtani, S., Anderson, R., & Moulds, M. L. (2021). Corrigendum to Transdiagnostic assessment of repetitive negative thinking and responses to positive affect: Structure and predictive utility for depression, anxiety, and mania symptoms.' Journal of Affective Disorders, 232, 375-384. Doi: 10.1016/j.jad.2018.02.072. *Journal of Affective Disorders*, *286*, Article 371. https://doi.org/10.1016/j.jad.2021.03.025

McEvoy, P. M., & Mahoney, A. E. (2013). Intolerance of uncertainty and negative metacognitive beliefs as transdiagnostic mediators of repetitive negative thinking in a clinical sample with anxiety disorders. *Journal of Anxiety Disorders*, *27*(2), 216–224. https://doi.org/10.1016/j.janxdis.2013.01.006

Mundfrom, D. J., Shaw, D. G., & Ke, T. L. (2005). Minimum sample size recommendations for conducting factor analyses. *International Journal of Testing*, *5*(2), 159–168. https://doi.org/10.1207/s15327574ijt0502_4

Nolen-Hoeksema, S. (1991). Responses to depression and their effects on the duration of depressive episodes. *Journal of Abnormal Psychology*, *100*(4), 569–582. https://doi.org/10.1037/0021-843x.100.4.569

Novick, M. R., & Lewis, C. (1967). Coefficient alpha and the reliability of composite measurements. *Psychometrika*, *32*(1), 1–13. https://doi.org/10.1007/bf02289400

Reinholt, N., Hvenegaard, M., Christensen, A. B., Eskildsen, A., Hjorthøj, C., Poulsen, S., Arendt, M., Rosenberg, N., Gryesten, J., Aharoni, R., Alrø, A., Christensen, C., & Arnfred, S. M. (2022). Transdiagnostic versus diagnosis-specific group cognitive behavioral therapy for anxiety disorders and depression: A randomized controlled trial. *Psychotherapy and Psychosomatics*, *91*(1), 36–49. https://doi.org/10.1159/000516380

Satorra, A., & Bentler, P. M. (1994). Corrections to test statistics and standard errors in covariance structure analysis. In A. von Eye, & C. C. Clogg (Eds.), *Latent variables analysis: Applications for developmental research* (pp. 399–419). Sage Publications.

Shi, D., Lee, T., & Maydeu-Olivares, A. (2019). Understanding the model size effect on SEM fit indices. *Educational and Psychological Measurement*, *79*(2), 310–334. https://doi.org/10.1177/0013164418783530

Topp, C. W., Østergaard, S. D., Søndergaard, S., & Bech, P. (2015). The WHO-5 well-Being Index: A systematic review of the literature. *Psychotherapy and Psychosomatics*, *84*(3), 167–176. https://doi.org/10.1159/000376585

Topper, M., Emmelkamp, P. M., Watkins, E., & Ehring, T. (2017). Prevention of anxiety disorders and depression by targeting excessive worry and rumination in adolescents and young adults: A randomized controlled trial. *Behaviour Research and Therapy*, *90*, 123–136. https://doi.org/10.1016/j.brat.2016.12.015

Valencia, P. D. (2020). Initial psychometric evaluation of the Perseverative Thinking Questionnaire in Peruvian undergraduates. *LIBERABIT. Revista Peruana de Psicología*, *26*(2), Article e404. https://doi.org/10.24265/liberabit.2020.v26n2.05

Watkins, E. R. (2008). Constructive and unconstructive repetitive thought. *Psychological Bulletin*, *134*(2), 163–206. https://doi.org/10.1037/0033-2909.134.2.163

Watkins, E. R., & Roberts, H. (2020). Reflecting on rumination: Consequences, causes, mechanisms and treatment of rumination. *Behaviour Research and Therapy*, *127*, Article 103573. https://doi.org/10.1016/j.brat.2020.103573

Weber, F., & Exner, C. (2013). Metacognitive beliefs and rumination: A longitudinal study. *Cognitive Therapy and Research*, *37*, 1257–1261. https://doi.org/10.1007/s10608-013-9555-y

History
Received March 14, 2023
Revision received September 13, 2023
Accepted November 6, 2023
Published online December 5, 2023

Section: Clinical Psychology

Acknowledgments
We thank all the patients and therapists who took part in this study. We acknowledge the work of Stephen Austin in back-translating the PTQ and the administrative support that has been provided by the clinic managers in the participating clinics.

Conflict of Interest
None of the authors have competing interests to declare.

Publication Ethics
All study procedures complied with the World Medical Association Declaration of Helsinki. The Ethics Committee Region Zealand (Registration Number: 3084871-SJ-582) and the Danish Data Protection Agency Region Zealand (Registration Number: REG-104-2016) approved the study. All patients gave written informed consent.

Open Data
The mother study that this study used data from was registered on the ClinicalTrials.gov website (ID No. NCT02954731). The data that support the findings of this study are available on request from the last author, Sidse Marie Arnfred. The data are not publicly available due to privacy restrictions imposed by the Danish Data Security Agency.

Funding
The mother study was funded by grants from TrygFonden (Grant ID 114241); JaschaFonden (Grant ID 5577 and 6215); Region Zealand Research Foundation (Grants RSSF2017-000667, RSSF2016-000342, and RSSF2015-000342); Region Zealand Mental Health Services, Mental Health Services Capital Region and Central Region Research foundation (no grant ID). The funding sources had no role in the study design, data collection, data analysis, or publication decisions.

ORCID
Stine Bjerrum Moeller
https://orcid.org/0000-0002-5019-8065
Ida-Marie T. P. Arendt
https://orcid.org/0000-0002-0889-4526
Thomas Ehring
https://orcid.org/0000-0001-9502-6868
Nina Reinholt
https://orcid.org/0000-0002-5117-9376
Anne Bryde Christensen
https://orcid.org/0000-0002-3087-8311

Stine Bjerrum Moeller
Department of Psychology
University of Southern Denmark
Campusvej 55
DK-5230 Odense
Denmark
stinebm@health.sdu.dk