Scandinavian Journal of Psychology, 2006, 47, 281–291 DOI: 10.1111/j.1467-9450.2006.00518.x

Personality and Social Sciences

A Norwegian adaptation of the Penn State Worry Questionnaire: Factor structure, reliability, validity and norms

STÅLE PALLESEN,[1] INGER HILDE NORDHUS,[1] BERIT CARLSTEDT,[2] JULIAN F. THAYER[3] and TOM BACKER JOHNSEN[1]

[1]*University of Bergen, Norway*
[2]*Swedish National Defence College, Karlstad, Sweden*
[3]*National Institute of Aging, Baltimore, USA*

Pallesen, S., Nordhus, I. H., Carlstedt, B., Thayer, J. F. & Johnsen, T. B. (2006). A Norwegian adaptation of the Penn State Worry Questionnaire: Factor structure, reliability, validity and norms. *Scandinavian Journal of Psychology*, *47*, 281–291

A Norwegian version of the Penn State Worry Questionnaire (PSWQ) was administered to 304 undergraduate students together with the Beck Depression Inventory (BDI), the State-Trait Anxiety Inventory (STAI) and the Maudsley Obsessive Compulsive Inventory (MOCI). The PSWQ was also administered to a community sample comprising 879 subjects, together with the Beck Anxiety Inventory (BAI), the Beck Depression Inventory II (BDI II) and the White Bear Suppression Inventory (WBSI). Structural equation modeling showed that a three-factor solution of the PSWQ gave the best goodness of fit. The Norwegian version of the PSWQ demonstrated adequate psychometric properties in terms of reliability and validity in both samples. Females scored higher than males on PSWQ.

Key words: Penn State Worry Questionnaire, factor structure, validity, reliability, norms.

Ståle Pallesen, Department of Psychosocial Science, University of Bergen. Christiesgt. 12, 5015 Bergen, Norway. Tel: +47 55 58 88 42; fax: +47 55 58 98 77; e-mail: staale.pallesen@psysp.uib.no

INTRODUCTION

Until the 1980s the concept of worry was primarily used within the test-anxiety tradition (Sarason, 1980). In DSM-III-R (American Psychiatric Association, 1987), however, worry was specified as one of the main diagnostic criteria for generalized anxiety disorder (GAD), hence worry was incorporated into the psychiatric nosology.

Worry can be defined as "a chain of thoughts and images, negatively affect-laden and relatively uncontrollable". It is assumed that the worry process represents an attempt to engage in mental problem-solving on an uncertain issue containing the possibility of one or more negative outcomes (Borkovec, Robinson, Pruzinsky & DePree, 1983, p. 10). Later revisions of the worry concept conceive of it as primarily thought based as opposed to imaginal activity. It is hypothesized that worry is initiated to avoid future negative events and immediate somatic anxiety (Borkovec & Inz, 1990). Conceptually, the worry process seems to have some similarity with depressive rumination typically encountered in patients suffering from depression (Starcevic, 1995; Wells, 1994). Both are characterized by negative mood, uncontrollability, repetitiveness and intuitive plausibility. However, depressive rumination is more telegraphic and less consciously mediated compared to worry (Wells, 1994) and whereas worry primarily is related to negative events that can take place in the future, depressive rumination is more centered around negative past events (Matthews, 1990). Some resemblance also seems to be present between worry and obsessions. Both are uncontrollable, intrusive and repetitive but they also seem to be distinct as worry is experienced as less senseless than obsessions. Also, worry is not so greatly resisted as obsessions and the content of worry relates more to normal daily experiences compared to obsessions (Wells, 1994). Other distinctive features are that worry represents a predominance of thought activity whereas obsessions may take a variety of forms and that worry, more than obsessions, is likely to be self-initiated and/or precipitated by common circumstances of everyday living (Turner, Beidel & Stanley, 1992).

Worry has in clinical research been associated with increased frequency of work absenteeism and medical consultations, increased risk of other anxiety disorders, depression, heart disease, diabetes and cancer (Gosselin, Dugas, Ladouceur & Freeston, 2001). Hence, it is important to be able to assess the levels of worry in individuals in clinical as well as in research settings.

The most commonly used self-report measure of worry is the Penn State Worry Questionnaire (PSWQ), consisting of 16 items, each rated on a five-point scale, ranging from 1 to 5. In all, eleven of the items of PSWQ (item no. 2, 4, 5, 6, 7, 9, 12, 13, 14, 15 and 16) are non-reversed, hence high scores on these reflect high levels of worry, while five items are reversed (item no. 1, 3, 8, 10, 11). Thus, high scores on these items reflect absence of or denial of worry. The total score ranges from 16 to 80 (Meyer, Miller, Metzger & Borkovec, 1990). An important aspect of the PSWQ is that the instrument is not related to any specific worry domain or content (Meyer *et al.*, 1990) in contrast to other worry measures (e.g., Worry Domains Questionnaire, WDQ; Tallis, Eysenck & Mathews, 1992).

Originally, the PSWQ was hypothesized to be a unidimensional instrument, thus all items were assumed to reflect a general worry factor (Meyer *et al.*, 1990). This structure was also demonstrated for a French (Ladouceur *et al.*, 1992) version of the PSWQ. Brown, Antony and Barlow (1992) also found a one-factor solution, although it could be argued that a two-factor solution should have been retained according to the Kaiser-Guttman rule (Lohelin, 1998). In several later studies the presence of a two-factor solution has been reported. Stöber (1995), investigating the German version of the PSWQ, found evidence for one general worry factor on which the non-reversed items loaded, and a method factor on which the reversed items loaded. Beck, Stanley and Zebb (1995) reported a similar two-factor structure for the English version of the PSWQ. For the Dutch version of the PSWQ, van Rijsoort, Emmelkamp and Vervaeke (1999) also found evidence for a two-factor solution.

Some studies have examined the factor structure for PSWQ by confirmatory factor analysis. For an Italian version of the PSWQ, Meloni and Gana (2001) found that a model with one trait factor (on which all items of PSWQ loaded), and two methods factors (one on which all non-reversed items loaded, and one on which all the reversed items loaded) gave the best fit with the data. Also, Gana, Martin, Canouet, Trouillet and Meloni (2002), reported that the same structure gave the best fit with the data for a French version of the PSWQ. Fresco, Heimberg, Mennin and Turk (2002) found that the two-factor solution, reported by Stöber (1995) and Beck *et al.* (1995) was superior to the unidimensional solution for the English version of the PSWQ. They also presented a higher order model, in which general worry represented the higher order factor, and where two method factors constituted the lower-order factors (Fresco *et al.*, 2002). However, Brown (2003) raises serious questions about the conclusions based upon these studies. He argues that the acceptability of the two-factor or three-factor models in these studies was solely based upon goodness of fit, without emphasizing the clinical, conceptual and empirical meaningfulness of an "absence of worry" dimension. He argues for retaining a one-factor model based upon substantially meaningful interpretation of the results (Brown, 2003).

When it comes to the internal consistency of the PSWQ, it has consistently been shown to be high. In studies with the original (English) version of the PSWQ alpha values of 0.91–0.95 (Meyer *et al.*, 1990), 0.94 (Davey, 1993), 0.80–0.89 (Beck *et al.*, 1995), and 0.86–0.95 (Brown *et al.*, 1992) have been reported. High Cronbach alphas have also been reported for French ($\alpha = 0.90$; Gana *et al.*, 2002), Dutch ($\alpha = 0.88$; van Rijsoort *et al.*, 1999), German ($\alpha = 0.86–0.89$; Stöber, 1995, 1998) and Icelandic ($\alpha = 0.92$; Jonsdottir & Smari, 2000) versions of the PSWQ.

Test-retest correlation of the PSWQ has across intervals from 2–10 weeks ranged from 0.74 to 0.93 (Molina & Borkovec, 1994). In a group of elderly patients with generalized anxiety disorder where the interval between the test and retest was between 5 and 20 weeks, the test-retest correlation was however considerably lower ($r = 0.54$; Stanley, Novy, Bourland, Beck & Averill, 2001).

The PSWQ has been administered to samples with different anxiety disorders, and the general conclusion from these studies is that GAD patients seem to score higher on the PSWQ compared to patients suffering from other anxiety disorders, thus indicating satisfactory criterion validity (Molina & Borkovec, 1994; Fresco, Mennin, Heimberg & Turk, 2003). Furthermore, the PSWQ has demonstrated sensitivity to change during therapy (Nordhus & Pallesen, 2003).

The PSWQ has also been validated with older adults (Beck *et al.*, 1995), but has mostly been administered to unselected samples, primarily consisting of students (Molina & Bokovec, 1994). In student populations the scores of the PSWQ showed high correlations with scores of other worry scales ($r = 0.59–0.68$; Davey, 1993, $r = 0.68$; Stöber, 1998), and with scores of trait measures of anxiety ($r = 0.64–0.74$; Davey, 1993; Meyer *et al.*, 1990) and somewhat lower with state measures of anxiety ($r = 0.49$) and measures of depression ($r = 0.36$; Meyer *et al.*, 1990). Similar patterns of correlations between the PSWQ and other instruments have been found for patient populations, although the overall magnitude of the correlations have been somewhat lower compared to student populations (Beck *et al.*, 1995; Brown *et al.*, 1992; Stanley *et al.*, 2001).

Normative values for the PSWQ have also been established. Gillis, Haaga and Ford (1995) reported that the mean score for a representative sample of US adults between the ages of 18 and 65 was 42.2 ($SD = 11.5$). No gender, race or income effects were found, but younger subjects (below 45 years) scored higher ($M = 43.5$, $SD = 7.3$) than the older (45 years and above) subjects ($M = 38.9$, $SD = 9.0$). Women in other studies, however sometimes score higher than men (Jonsdottir & Smari, 2000; Meloni & Gana, 2001; Meyer *et al.*, 1990).

Although the PSWQ in general has been found to have good psychometric properties, some questions still remain unresolved. This pertains particularly to its factor structure, but also to the issue of gender differences. Furthermore, the psychometric properties of the PSWQ have only to a limited extent been investigated in community samples. Additionally, the PSWQ has not yet been investigated in Scandinavia. Based on this we decided to conduct a study, investigating the psychometric properties (factor structure, internal consistency, test-retest reliability and validity) of a Norwegian version of the PSWQ in a Norwegian student and in a Norwegian community sample. We also aimed at obtaining Norwegian norms for the PSWQ.

METHOD

Translation

The translation of the English version of the PSWQ was done by two of the authors. Agreement about the best translation for each single item was then reached. Each item was then translated back by two bilingual professionals. All persons involved examined the back translations for the psychological essence of the items. At this stage, one of the sixteen Norwegian items was modified, translated back into English, and again rechecked (see Appendix).

Subjects

Student sample. A total of 304 undergraduate students at the University of Bergen and from the Royal Norwegian Naval Academy participated in the study. The total sample consisted of 135 females (mean age 22.4, *SD* = 4.87) and 169 males (mean age 23.6, *SD* = 3.31).

Community sample. A total of 879 subjects, representing all counties of Norway, constituted the community sample. This sample consisted of 514 females (mean age 43.5, *SD* = 13.08), and 365 males (mean age 45.4, *SD* = 14.61).

Instruments

Student sample. In addition to the PSWQ, the following measures were used in the student sample:

(1) *Beck Depression Inventory* (BDI, Beck, Ward, Mendelson, Mock & Erbaugh, 1961). The BDI is a 21-item self-report instrument that measures common symptoms of depression, along a four-point scale ranging from 0 to 3. The BDI has demonstrated good psychometric properties (Beck, Steer & Garbin, 1988).
(2) *State-Trait-Anxiety Inventory* (STAI; Spielberger, Gorsuch, Lushene, Vagg & Jacobs, 1983). The STAI consists of 40 items, of which 20 measures trait anxiety and the other 20 measure state anxiety. Each item is rated on a four-point scale, ranging from 1 to 4. The two subscales have shown adequate validity (Spielberger *et al.*, 1983).
(3) *Maudsley Obsessive Compulsive Inventory* (MOCI; Hodgson & Rachman, 1977). The MOCI has 30 items and consists of four subscales, of which Doubting and Checking primarily are considered as measures of obsessions, whereas the subscales Washing and Repetition mainly reflect compulsions. Each item is rated on a two-point scale (true or false). Additionally, a single, total score for the MOCI is normally calculated (Hogdson & Rachman, 1977).

Community sample. In addition to the PSWQ the following measures were used in the community sample:

(1) *Beck Anxiety Inventory* (BAI; Beck & Steer, 1990). The BAI is a 21-item self-report questionnaire measuring common symptoms of clinical anxiety. Each symptom is rated on a four-point scale ranging from 0 to 3. The BAI has shown good psychometric properties (Beck and Steer, 1990).
(2) *Beck Depression Inventory II* (BDI-II; Beck, Steer & Brown, 1996). The BDI-II is a 21-item self report questionnaire of depressive symptoms. Each symptom is rated on a four-point scale ranging from 0 to 3. Beck *et al.* (1996) report good psychometric properties for The BDI-II.
(3) *White Bear Suppression Inventory* (WBSI; Wegner & Zanakos, 1994). The WBSI is a 15-item self-report instrument that assesses the respondents' general tendency to suppress thoughts. Each item is rated on a five-point scale, ranging from 1 to 5.

Procedure

Student sample. The students were recruited at lectures, and the ones that agreed to participate were asked to complete the PSWQ, BDI, STAI and MOCI, in addition to stating their gender and age. The response rate was about 95%. In order to investigate the test-retest reliability, the PSWQ was readministered to a subsample of 105 males (mean age 23.4, *SD* = 2.58) and 65 females (mean age 23.5, *SD* = 6.11) from the original sample three weeks after the first administration of the questionnaires, yielding an attrition rate of 44.4%.

Community sample. The community sample was drawn randomly from a survey population, consisting of a register of phone numbers covering approximately 98% of the population. In all, 4,738 phone calls were made. Using the last birthday technique, 2003 subjects stratified by the number of inhabitants of the counties of Norway, agreed to participate in a survey by Opinion Research Institute. Of those, 1,196 agreed to participate in the present study, hence they received the survey questionnaires (PSWQ, BDI-II, BAI and WBSI) by mail. The number of those actually completing the survey and returning it was 879, yielding a total response rate of 18.6%.

Statistical analysis

The statistical analyses were performed by SPSS version 10.1 (SPSS Inc, 2000), STREAMS version 2.52 (Gustafsson & Stahl, 2001) and AMOS version 4.01 (Arbuckle, 1999). The analyses of the dimensionality of the PSWQ were based on confirmatory factor analyses investigating: (1) the fit of the original one-factor model (Meyer *et al.*, 1990); (2) the two-factor model suggested by Stöber (1995), Beck *et al.* (1995) and Fresco *et al.* (2002) where we would allow the two latent variables non-reversed and reversed items to be correlated; (3) a model with one general worry factor and a non-reversed and reversed items factor. In the latter model the non-reversed and reversed items factors were regarded as orthogonal (Meloni & Gana, 2001; Gana *et al.*, 2002). Since these models are non-nested, that is, not subsets of each others, we chose fit indexes such as the root mean square error of approximation (RMSEA) and the Akaike information criteria (AIC; Akaike, 1987) to order the models from best to worse fitting. RMSEA is zero for a perfect fit, while RMSEA for a well-fitting model should have a value below 0.05, while a value of about 0.08 or less would indicate a reasonable error of approximation. Models with values greater than 0.1 should not be retained (Browne & Cudeck, 1993). AIC is a modification of the standard goodness-of fit χ^2 statistic that includes a penalty for complexity. Given two or more non-nested models, the one with the lowest AIC is preferred (Akaike, 1987). In order to investigate whether the structure of the retained model (the model with best fit) was equal in different samples and subgroups a two-group model was fitted, imposing constraints over all parameters (means, variances and relations) and compared to a model with no constraints. If this difference was significant, we would locate more closely in what part of the model the group differences were located by successively relaxing the model with constraints over all parameters. This approach would also be employed for potential gender and age differences (based upon a median split procedure) in the community sample. Based on the regression weights of the model with the best fit, factor scores for each latent variable were computed.

Table 1. *Standardized parameter estimates from the final three-factor model of PSWQ*

Item no.	General worry factor	Non-reversed items	Reversed items	Error	General worry factor	Non-reversed items	Reversed items	Error
	Student sample (N = 304)				Community sample (N = 879)			
1[a]	−0.48		0.30	0.82	−0.27		0.44	0.86
2	0.71	0.24	0.66	0.69	0.30	0.66		
3[a]	−0.61		0.35	0.71	−0.47		0.57	0.68
4	0.76	0.24	0.61	0.78	0.03*	0.63		
5	0.86	0.07*	0.51	0.87	0.06*	0.49		
6	0.67	0.15	0.73	0.69	0.11	0.72		
7	0.68	0.42	0.60	0.68	0.48	0.55		
8[a]	−0.56		0.34	0.75	−0.53		0.53	0.66
9	0.55	0.24	0.80	0.62	0.38	0.69		
10[a]	−0.55		0.47	0.69	−0.53		0.46	0.71
11[a]	−0.38		0.41	0.83	−0.32		0.58	0.75
12	0.64	0.32	0.71	0.63	0.36	0.68		
13	0.59	0.00*	0.81	0.66	0.20	0.72		
14	0.63	0.32	0.71	0.71	0.29	0.64		
15	0.54	0.76	0.37	0.66	0.56	0.48		
16	0.52	0.30	0.80	0.59	0.32	0.74		
	Community sample							
	Males (N = 362)				Females (N = 505)			
1[a]	−0.24		0.46	0.86	−0.25		0.43	0.86
2	0.70	0.22		0.68	0.69	0.27		0.67
3[a]	−0.43		0.59	0.69	−0.46		0.58	0.68
4	0.75	0.07*		0.66	0.78	−0.03*		0.63
5	0.86	0.04*		0.51	0.87	0.03*		0.49
6	0.68	0.11*		0.73	0.68	0.09*		0.72
7	0.70	0.34		0.63	0.71	0.47		0.53
8[a]	−0.50		0.62	0.61	−0.50		0.47	0.73
9	0.66	0.23		0.72	0.62	0.38		0.68
10[a]	−0.54		0.49	0.69	−0.49		0.46	0.74
11[a]	−0.27		0.68	0.68	−0.33		0.50	0.80
12	0.71	0.41		0.57	0.60	0.32		0.73
13	0.72	0.13		0.68	0.62	0.22		0.75
14	0.73	0.28		0.63	0.69	0.28		0.67
15	0.67	0.65		0.36	0.69	0.52		0.50
16	0.65	0.22		0.73	0.59	0.32		0.74

[a] Reversed items.
* Non-significant.

Internal consistency was estimated with Cronbach's alpha. Test-retest correlation and corrected item-total correlations were calculated by Pearson's product moment correlation. Pearson's product moment correlations were also used to calculate the relationships between the PSWQ and the other instruments. T-tests for dependent samples were used in order to investigate whether differences between correlations were statistically significant.

RESULTS

Factor structure

Student sample. A confirmatory factor analysis showed that the one-factor model had relatively poor fit with the data (RMSEA = 0.086; AIC = 433). The two-factor model had somewhat better fit (RMSEA = 0.066; AIC = 366). The correlation between the non-reversed and the reversed item factors was −0.76. The model with one trait factor and the two orthogonal factors represented the best fit with the data (RMSEA = 0.048; AIC = 277). The factor loadings for the three-factor model are shown in Table 1. The general worry factor explained 52.2% of the variance, the non-reversed items factor explained 20.3% of the variance and the reversed items factor explained a total of 8.2% of the variance.

Communality sample. A confirmatory factor analysis showed relatively poor fit between the one-factor model and the data

(RMSEA = 0.106; AIC = 1214). The two factor model fitted better with the data (RMSEA = 0.062; AIC = 540). The correlation between the two factors was –0.61. The model with one trait factor and the two orthogonal factors represented the best fit with the data (RMSEA = 0.040; AIC = 337). The factor loadings for the final three-factor model are shown in Table 1. In all, the general worry factor explained 54.1% of the variance, the non-reversed items factor explained 16.8% of the variance and the reversed items factor explained a total of 14.6% of the variance.

Multi-model samples

A two-group model (community population and student population) with full constraints over all parameters (means, variances and relations) had good fit with the data (χ^2 = 631.6, df = 240, $p < 0.01$, RMSEA = 0.037). The test-statistics for a model with no constraints of equality over groups was however χ^2 = 358.1, df = 176, $p < 0.01$, RMSEA = 0.030), which implies that the difference test is significant ($\Delta\chi^2$ = 273.5, df = 64) and that there were differences in the model between the community and student sample. Comparing with a model with full constraints over all parameters, but removing the constraints of latent variable means gave χ^2 = 611.7, df = 237, $p < 0.01$, RMSEA = 0.037, thus the difference test ($\Delta\chi^2$ = 19.9, df = 3) was significant, indicating difference in the mean of the latent variables between the community sample and the student sample. The community sample scored 0.23 standard units (*z*-scores) lower on the latent mean for worry, 0.14 standard units higher for the latent mean for the non-reversed items factor and 0.35 standard units lower for the latent mean for reversed items factor compared to the student sample. Removing constraints on means from the model altogether gave χ^2 = 498.0, df = 224, $p < 0.01$, RMSEA = 0.032, thus this difference test ($\Delta\chi^2$ = 113.7, df = 13) was significant, indicating that there were differences between the community and student sample regarding the means of the manifest variables. Only for one item (PSWQ2) the difference was more than 0.3 standard units. The mean score for the community sample was 0.36 lower on this item compared to the score for the student sample. Then, removing constraints with respect to error variances for the manifest variables yielded χ^2 = 464.1, df = 208, $p < 0.01$, RMSEA = 0.032, thus this difference test ($\Delta\chi^2$ = 33.9, df = 16) was significant, indicating differences between the community and student population regarding error variances. For items 1 and 11 the explained variance was more than 5% lower in the community sample compared to the student sample. Further, removing the constraints related to the variance of the latent variables gave χ^2 = 432.4, df = 205, $p < 0.01$, RMSEA = 0.031, thus this difference test ($\Delta\chi^2$ = 31.7, df = 3) was also significant, indicating differences between the community and student population in terms of variance of the latent variables. For the student sample the worry, non-reversed items and the reversed items factor explained 56.2%, 18.7% and 7.4% of the variance, respectively, while the corresponding figures for the community sample were 50.1%, 19.9% and 14.9%. Finally, we tested the homogeneity of the regressions of the manifest variables on the latent variables between the community and student sample. The test statistics of this model was χ^2 = 358.1, df = 176, $p < 0.01$, RMSEA = 0.030, which of course was the same as the completely unconstrained model. The difference test ($\Delta\chi^2$ = 74.3, df = 29) was significant, demonstrating differences between the community and student sample regarding the homogeneity of the regressions of the manifest variables on the latent variables as well (see Table 1).

A two-group model (men vs. females of the community sample) with full constraints over all parameters had good fit with the data (χ^2 = 664.0, df = 240, $p < 0.01$, RMSEA = 0.045). The test-statistics for a model with no constraints of equality over groups was, however, χ^2 = 326.2, df = 176, $p <$ 0.01, RMSEA = 0.031), which implies that the difference test is significant ($\Delta\chi^2$ = 337.8, df = 64) and that there were differences in the model between the females and males. Compared to a model with full constraints over all parameters, but removing the constraints of latent variable means gave χ^2 = 584.5, df = 237, $p < 0.01$, RMSEA = 0.041, thus the difference test ($\Delta\chi^2$ = 80.5, df = 3) was significant, indicating difference in the mean of the latent variables between the females and men. Men scored 0.64 standard units (*z*-scores) lower on the latent mean for worry, 0.40 standard units higher for the latent mean for the non-reversed items and 0.12 standard units lower for the latent mean for the reversed items compared to the females. Removing constraints of means from the model altogether gave χ^2 = 564.7, df = 224, $p < 0.01$, RMSEA = 0.042, thus this difference test ($\Delta\chi^2$ = 18.8, df = 13) was not significant, indicating no differences between the genders regarding the means of the manifest variables. Then, removing constraints with respect to error variances for the manifest variables yielded χ^2 = 407.2, df = 208, $p < 0.01$, RMSEA = 0.033, thus this difference test ($\Delta\chi^2$ = 157.5, df = 16) was significant, indicating differences between the genders regarding error variances. For items 2, 4, 5, 6, 9, 12, 13, 14 and 15 the explained variance was more than 5% lower for females compared to males. For the other items the differences in explained variance was less than 5%.

Further, removing the constraints related to the variance of the latent variables gave χ^2 = 382.1, df = 205, $p < 0.01$, RMSEA = 0.032, thus this difference test ($\Delta\chi^2$ = 27.0, df = 3) was also significant, indicating differences between the genders in terms of variance of the latent variables. For the females the worry, non-reversed items and reversed items factor explained 54.1%, 20.0% and 11.0% of the variance, respectively, while the corresponding figures for men were 52.9%, 12.4% and 20.6%. Finally we tested the homogeneity of the regressions of the manifest variables on the latent variables between the community and student sample. The test statistics of this model was identical to the model with

Table 2. *Norms of PSWQ for different subgroups*

	Student sample (N = 304)	
Men	(n = 166)	36.9 (8.6)
Women	(n = 134)	48.3 (12.5)

	Community sample (N = 874)							
	18–29 yrs		30–44 yrs		45–59 yrs		60+	
Men	(n = 102)	38.9 (14.0)	(n = 127)	37.0 (11.8)	(n = 100)	37.3 (13.7)	(n = 97)	35.0 (8.5)
Women	(n = 97)	46.4 (11.8)	(n = 122)	43.8 (13.3)	(n = 97)	41.3 (11.8)	(n = 123)	41.3 (11.6)

no constraints, χ^2 = 326.2, df = 176, p < 0.01, RMSEA = 0.031. The difference test ($\Delta\chi^2$ = 56.0, df = 29) was not significant demonstrating equality between men and females regarding the homogeneity of the regressions of the manifest variables on the latent variables as well (see Table 1).

A two-group model (young, ≤42 years vs. older, ≥43 years) with full constraints over all parameters had good fit with the data (χ^2 = 464.5, df = 240, p < 0.01, RMSEA = 0.033). The test-statistics for a model with no constraints of equality over groups was however χ^2 = 419.2, df = 176, p < 0.01, RMSEA = 0.044), which implies that the overall difference test was not significant ($\Delta\chi^2$ = 45.3, df = 64) and that overall there was equality of the models for young (≤42 years) compared to the older (≥43 years) subjects.

Norms

The mean score on the PSWQ for the student sample was 41.9 (SD = 11.9, range = 20–76). The weighted mean (weighted according to the population distribution of age and gender) for the community sample was 37.1 (SD = 11.7, range = 16–74). Table 2 shows the means and standard deviations of the PSWQ for different subgroups.

Reliability analyses

Cronbach's alpha for the Norwegian version of the PSWQ was 0.92 for the student sample and 0.92 for the community sample. The three week test-retest reliability of the PSWQ based on a subsample (n = 166) of the student sample was 0.84 (p < 0.01). Table 3 shows the means and standard deviations for the items of the PSWQ together with corrected item-total correlations.

Table 3. *Means, standard deviations and corrected item-total correlations of items in the PSWQ*

	Student sample (N = 304)			Community sample (N = 874)		
Item no.	Mean	*SD*	Item-total correlation	Mean	*SD*	Item-total correlation
1[a]	3.47	1.10	−0.49	3.51	1.21	−0.36
2	2.17	1.10	0.70	1.80	1.02	0.69
3[a]	3.47	1.12	−0.61	3.41	1.25	−0.55
4	2.45	1.17	0.73	2.40	1.22	0.69
5	2.52	1.29	0.78	2.46	1.32	0.78
6	2.91	1.17	0.65	2.60	1.24	0.64
7	2.03	1.22	0.72	1.87	1.18	0.73
8[a]	3.31	1.09	−0.54	3.50	1.26	−0.59
9	2.03	1.00	0.59	1.90	1.13	0.67
10[a]	4.12	1.08	−0.56	3.98	1.15	−0.57
11[a]	2.65	1.08	−0.44	2.98	1.34	−0.41
12	1.80	1.04	0.66	1.77	1.12	0.66
13	2.89	1.19	0.53	2.55	1.24	0.64
14	2.14	1.06	0.67	2.07	1.15	0.71
15	1.52	0.87	0.67	1.53	0.95	0.73
16	2.45	1.06	0.56	2.21	1.17	0.62

[a] Reversed item.

Convergent/discriminative validity

Student sample. The PSWQ correlated higher (t = 6.90, df = 291, p < 0.01) with the trait version of the STAI (r = 0.76) compared to the state version (r = 0.56). The correlation with the BDI was moderately high (r = 0.56). The PSWQ correlated higher with the predominantly obsessive subscale Doubting (r = 0.40) as compared to the predominantly compulsive subscales Washing (r = 0.19, t = 3.35, df = 294, p < 0.01) and Slowness/Repetition (r = –0.14, t = 8.18, df = 294, p < 0.01) of the MOCI. Analogously, the PSWQ also correlated higher with the other predominantly obsessive subscale of the MOCI, Checking (r = 0.38) as compared to the subscales Washing (r = 0.19, t = 3.53, df = 295, p < 0.01) and Slowness/Repetition (r = –0.14, t = 7.34, df = 295, p < 0.01). The non-reversed items actor had moderate positive correlations with the different clinical scales (BDI, STAI, MOCI), whereas the reversed items factor had very low correlations with other scales (see Table 4).

Table 4. *Pearson product moment correlation between PSWQ,PSWQ factor scores, Trait and State version of STAI, BDI, and MOCI with its four subscales: Doubting, Checking, Washing and Slowness/Repetition for the student sample (*N = *304)*

	General worry factor	Non-reversed items	Reversed items	STAI-T	STAI-S	BDI	Doubting	Checking	Washing	Slowness/ Repetition	MOCI
PSWQ	0.96**	0.31**	−0.25**	0.76**	0.56**	0.56**	0.40**	0.38**	0.19**	−0.14*	0.42**
General worry factor		0.13*	−0.13*	0.70**	0.51**	0.52**	0.36**	0.34**	0.16**	−0.13*	0.37**
Non-reversed items			0.11	0.41**	0.30**	0.37**	0.26**	0.22**	0.15**	−0.08	0.27**
Reversed items				−0.13*	−0.15*	−0.05	0.10	−0.07	−0.04	0.04	−0.04
STAI-T					0.71**	0.73**	0.43**	0.39**	0.23**	−0.14*	0.45**
STAI-S						0.53**	0.32**	0.32**	0.19**	−0.04	0.36**
BDI							0.42**	0.33**	0.23**	−0.06	0.45**
Doubting								0.37**	0.30**	0.08	0.72**
Checking									0.49**	0.10	0.80**
Washing										0.29**	0.77**
Slowness/Repetition											0.29**

* $p < 0.05$; ** $p < 0.01$.

Table 5. *Pearson product moment correlation between PSWQ, factor scores of PSWQ, BDI-II, BAI and WBSI for the community sample (*N = *879)*

	General worry factor	Non-reversed items	Reversed items	BDI-II	BAI	WBSI
PSWQ	0.95**	0.36**	−0.36**	0.57**	0.53**	0.55**
General worry factor		−0.20**	−0.09*	0.55**	0.52**	0.56**
Non-reversed items			0.11*	0.30**	0.27**	0.19**
Reversed items				−0.09*	−0.08*	−0.07*
BDI-II					0.62**	0.51**
BAI						0.48**

* $p < 0.05$; ** $p < 0.01$.

Community sample. The PSWQ correlated moderately with the BDI-II ($r = 0.57$), The BAI ($r = 0.53$) and the WBSI ($r = 0.55$). As for study 1, the non-reversed items factor showed moderate positive correlations to different clinical scales (BDI II, BAI, WBSI) whereas the reversed items factors had low and inconsistent correlations with the other scales (see Table 5).

DISCUSSION

Factor structure

The original one-factor model of the PSWQ, presented by Meyer *et al.* (1990) had relatively poor fit with the data in the present study. Studies using exploratory factor analyses instead suggest a two-factor solution (Beck *et al.*, 1995; Stöber, 1995), while some studies, as the present one, found best fit for a three-factor solution, with one general worry factor and a non-reversed items and a reversed items factor (Gana *et al.*, 2002; Meloni & Gana, 2001). However, the three-factor solution was not invariant across all multi-group comparison. Our results regarding the factor structure run counter with the results from studies by Meyer *et al.* (1990) and Brown *et al.* (1992), both reporting a one-factor solution. As both these studies were conducted in the USA, the contrast with the results of the present study may reflect that the structure of the PSWQ is different in different cultures. Alternatively, the structural divergence can be attributed to the type of factor analysis, as the two former studies used exploratory, whereas the present study was conducted with confirmatory factor analyses.

When it comes to the three-factor solution found in the present study, an essential question to ask is whether the non-reversed items and the reversed items factors should be interpreted as meaningful and substantial or if they should be regarded as method artifacts. Other studies have shown that when scales contain non-reversed and reversed worded items, factor analyses produce factors based upon these categories of items. Thus, basically, such factors could be regarded as method effects (Marsh, 1996). Others have, however, argued that such factors may have substantial meaning. Studies have for example shown that response style, "a behavior consistency operating across measures of several conceptually distinct content traits" (Bentler, Jackson & Messick, 1971, p. 188) provides one conceptual basis for interpreting systematic effects associated with item wording

(Horan, DiStefano & Motl, 2003). When it comes to the results from the present study the correlation matrices (Tables 4 and 5) show that the reversed items factor had low and inconsistent correlations with other measures. Thus, it may be argued that it is reasonable to interpret the reversed items factor as a method factor. The non-reversed items factor, in contrast, seems to be systematically related to the scores of the other instruments. Consequently, this factor may be reflecting a meaningful psychological construct. Response style (Horan *et al.*, 2003) and negative affect (Beck *et al.*, 1995) may be likely candidates for concepts expressed by this factor. Inspection of the specific items, revealed that items no. 7 and 15 loaded particularly high while items 4–6 had low loadings on the non-reversed items factor. As both item 7 and 15 are concerned with worrying taking place "always" and, "all the time" and as items 4, 5 and 6 seem to be more related to situational specific worrying, it could also be argued that the non-reversed items factor reflects constant worrying, in contrast to worrying related to specific situations/things.

Multi-model samples

The analysis showed that the three-factor model was not identical in the student sample compared to the community sample. The latent mean for worry was lower in the community sample compared to the student sample. There were also differences regarding the mean of the manifest variables, variances and the regressions of the manifest variables on the latent variables. The community and student sample did differ in age, gender and probably also on several other variables, thus it is difficult to point to specific factors related to the lack of invariance.

In the present study we found that women scored higher than males on the latent worry variable. Other studies with the PSWQ have found that females report higher levels of worry compared to men (Jonsdottir & Samri, 2000; Meloni & Gana, 2001; Meyer *et al.*, 1990), while other studies report no such differences (Stöber, 1995; van Rijsoort *et al.*, 1999). It is not clear why females score higher than males on worry. It has, however, been speculated that women engage in a ruminative style of coping when mildly distressed, while men, on the other hand, engage in strategies of distraction. An alternative explanation is that men under-report worry due to social desirability (Robichaud, Dugas & Conway, 2003). There were also gender differences regarding variances. For females the reversed items factor explained more variance than the non-reversed items factor, while the reverse was true for men.

Our multi-model comparison showed that the three-factor solution was invariant across age groups (≤42 years and ≥43 years). This finding runs counter with the findings of Gillis *et al.* (1995) who found that younger subjects reported higher levels of worry compared to older subjects. It should be pointed out, however, that that latter conclusion was based on an observed variable analysis, whereas in the present study the latent variable methodology was employed.

Norms

The mean score of the PSWQ in the student sample corresponded well with the scores in a German (Stöber, 1995) and Icelandic student sample (Jonsdottir & Smari, 2000), but was somewhat lower than the mean of equivalent US samples (Molina & Borkovec, 1994). However, as the factor structure may vary across study and nations, such comparisons do probably not represent meaningful enterprises. This becomes obvious as cultural differences on test scores may be related to other factors than substantial differences across cultures. Such factors could be lack of functional equivalence as subjects from different cultures may differ in willingness and motivation to report attitudes, feelings and behavior. Different scores across cultures may also reflect lack of conceptual equivalence which is the meaning individuals attach to specific stimuli such as test items on a questionnaire. Cross-cultural differences in scores on a questionnaire could also be attributed to lack of linguistic (translation) equivalence, as some words and sentences may not convey the exact same meaning across cultures. Finally, there could be metric or scalar differences across cultures explaining different scores and factor structures of questionnaires. Consequently, such differences may arise from the fact that different ethnic or cultural groups may perceive the scale or the response alternatives in different ways (Lonner & Ibrahim, 1996, pp. 302–3).

Reliability

The Cronbach alpha was high (0.92) in both samples, indicating high internal consistency. This finding is reported quite consistently in the literature (Molina & Borkovec, 1994). The high three-week test-retest correlation ($r = 0.84$) demonstrated in a subsample of the students is a further reflection of satisfactory reliability of the Norwegian version of the PSWQ and is also in line with results from the majority of earlier studies (Molina & Borkovec, 1994).

Validity

As for the convergent/discriminative validity in the student sample, the PSWQ correlated higher with the trait as compared to the state version of the STAI, strengthening the conception of the PSWQ as a trait measure. The PSWQ showed a moderate correlation ($r = 0.56$) with the BDI. As some resemblance between worry and depressive rumination exists, this correlation would be expected to be reasonably high. Thus, the results can be interpreted as one indication of discriminative validity. These results are also in line with former studies conducted with the PSWQ (Davey, 1992; Meyer *et al.*, 1990). As expected, the PSWQ correlated higher with the two obsessive as compared to the two compulsive subscales of the MOCI. This was also expected as both worry and obsessions primarily are cognitive processes, whereas compulsions have a much stronger behavioral component. In addition, similar findings have been reported elsewhere (Brown, Moras, Zinbarg & Barlow, 1993). In the student sample, the PSWQ correlated about 0.55 with all the instruments. For the community sample, the PSWQ correlated at approximately 0.55 with all the other instruments (BDI II, BAI and WBSI). The strength of the correlation of the PSWQ with the BDI II (0.57) is similar to the correlation between the PSWQ and the BDI in the student sample (0.56) and is in the expected range. As the BAI is a state measure of anxiety, the correlation with the PSWQ in the community sample (0.53) is similar to the correlation between the PSWQ and the state version of STAI (0.56) but lower than the correlation between the PSWQ and the trait version of STAI (0.76) in the student sample. This result could be interpreted as the PSWQ being a trait measure. As the WBSI is a measure of the effort and tendency to suppress certain thoughts it would be expected to correlate moderately with the PSWQ, an expectation supported by the data. In both the student and community sample the general worry factor had correlations with the other instruments very similar to the correlations between the PSWQ and the other instruments, suggesting that it really measures worry. The discussion regarding the non-reversed and the reversed items factors have been presented in the earlier.

Limitations of the present study

Related to the community sample in the present study some cautions should be noted, as the response rate only was about 19%, hence, it should probably be regarded as a convenience sample. The low response rate weakened the normative power of the study. Still, compared to Gillis *et al.* (1995) who recruited subjects at two shopping malls, the survey population was probably more representative of the general population in our study. It should also be noted that the sample sizes of Gillis *et al.* (1995; $n = 244$) and van Rijsoort *et al.* (1999; $n = 161$) were relatively small compared to the sample size of this study ($n = 879$).

It should also be noted that the scales used in the convergent/discriminant analyses were not subjected to a latent variable analysis. As some of these scales (STAI and MOCI) also contain reversed items, future validation studies of the PSWQ could also subject such scales to latent variable analysis in order to interpret the non-reversed and the reversed factors in a more stringent way.

Conclusion

For the PSWQ unidimensionality should be expected. However, we found evidence for a three-factor solution. It is suggested that two of the factors probably have a substantial basis, whereas the third is interpreted as a methodological artifact. The Norwegian version of the PSWQ had high test-retest reliability and acceptable convergent and discriminative validity. Females seem to have higher levels of worry than men. The PSWQ might function better without reversed items.

REFERENCES

Akaike, H. (1987). Factor analysis and AIC. *Psychometrika*, *52*, 317–322.

American Psychiatric Association. (1987). *Diagnostic and Statistical Manual of Mental Disorders* (3rd rev. edn). Washington, DC: APA.

Arbuckle, J. L. (1999). *AMOS (Version 4.01)*. [Computer software]. Chigaco, IL: SmallWaters Corporation.

Beck, A. T. & Steer, R. A. (1990). *Manual for the Beck Anxiety Inventory*. San Antonio, TX: The Psychological Corporation.

Beck, A. T., Steer, R. A. & Brown, G. K. (1996). *Beck Depression Inventory II. Manual*. San Antonio, TX: The Psychological Corporation.

Beck, A. T., Steer, R. A. & Garbin, M. G. (1988). Psychometric properties of the Beck Depression Inventory: Twenty-five years of evaluation. *Clinical Psychology Review*, *8*, 77–100.

Beck, A. T., Ward, C. H., Mendelson, M., Mock, J. & Erbaugh, J. (1961). An inventory for measuring depression. *Archives of General Psychiatry*, *4*, 561–571.

Beck, J. G., Stanley, M. A. & Zebb, B. J. (1995). Psychometric properties of the Penn State Worry Questionnaire in older adults. *Journal of Clinical Geropsychology*, *1*, 33–41.

Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin*, *107*, 238–246.

Bentler, P. M., Jackson, D. N. & Messick, S. (1971). Identification of content and style – 2-dimsensional interpretation of acquisence. *Psychological Bulletin*, *76*, 186–204.

Bollen, K. A. (1989). *Structural equations with latent variables*. New York: John Wiley.

Borkovec, T. D. & Inz, J. (1990). The nature of worry in generalized anxiety disorder: A predominance of thought activity. *Behavior Research and Therapy*, *28*, 153–158.

Borkovec, T. D., Robinson, E., Pruzinsky, T. & DePree, J. A. (1983). Preliminary exploration of worry: Some characteristics and processes. *Behaviour Research and Therapy*, *21*, 9–16.

Brown, T. A. (2003). Confirmatory factor analysis of the Penn State Worry Questionnaire: Multiple factors or method effects? *Behaviour Research and Therapy*, *41*, 1411–1426.

Brown, T. A., Antony, M. M. & Barlow, D. H. (1992). Psychometric properties of the Penn State Worry Questionnaire in a clinical anxiety disorder sample. *Behaviour Research and Therapy*, *30*, 33–37.

Brown, T. A., Moras, K., Zinbarg, R. E. & Barlow, D. H. (1993). Diagnostic and symptom distinguishability of generalized anxiety disorder and obsessive-compulsive disorder. *Behavior Therapy*, *24*, 227–240.

Browne, M. W. & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation* models (pp. 136–162). Newbury Park, CA: Sage Publications.

Davey, G. C. L. (1993). A comparison of three worry questionnaires. *Behaviour Research and Therapy*, *31*, 51–56.

Fresco, D. M., Heimberg, R. G., Mennin, D. S. & Turk, C. L. (2002). Confirmatory factor analysis of the Penn State Worry Questionnaire. *Behaviour Research and Therapy*, *40*, 313–323.

Fresco, D. M., Heimberg, R. G., Mennin, D. S. & Turk, C. L. (2003). Using the Penn State Worry Questionnaire to identify individuals with generalized anxiety disorder: A receiver operating characteristic analysis. *Journal of Behaviour Therapy and Experimental Psychiatry*, *34*, 283–291.

Gana, K., Martin, B., Canouet, M. D., Trouillet, R. & Meloni, F. (2002). Factorial structure of a French version of the Penn State Worry Questionnaire. *European Journal of Psychological Assessment*, *18*, 158–164.

Gillis, M. M., Haaga, D. A. F. & Ford, G. T. (1995). Normative values for the Beck Anxiety Inventory, Fear Questionnaire, Penn State Worry Questionnaire, and Social Phobia and Anxiety Inventory. *Psychological Assessment*, *7*, 450–455.

Gosselin, P., Dugas, M. J., Ladouceur, R. & Freeston, M. H. (2001). Evaluation of worry: validation of a French Translation of the Penn State Worry Questionnaire. *Encephale-Revue de Psychiatrie Clinique Biologique et Therapeutique*, *27*, 475–484.

Gustafsson, J. E. & Stahl, P. A. (2001). *STREAMS (Version 2.52)* [Computer software]. Göthenburg: Multivariate Ware.

Hodgson, R. J. & Rachman, S. (1977). Obsessive-compulsive complaints. *Behaviour Research and Therapy*, *15*, 389–395.

Horan, P. M., DiStefano, C. & Motl, R. W. (2003). Wording effects in self-esteem scales: Methodological artifact or response style? *Structural Equation Modeling*, *10*, 435–455.

Hu, L. & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, *6*, 1–55.

Jonsdottir, S. D. & Smari, J. (2000). Measuring obsessions without worry: Convergent and discriminant validity of the Revised Padua Inventory in an Icelandic student population. *Scandinavian Journal of Behaviour Therapy*, *29*, 49–56.

Ladouceur, R., Freeston, M. H., Dumont, J., Letarte, H., Rheaume, J., Thibodeau, N. & Gagnon, F. (1992). *The Penn State Worry Questionnaire: Psychometric properties of a French Translation*. Paper presented at the Annual Convention of the Canadian Psychological Association, Quebec City, Canada, June.

Lohelin, J. C. (1998). *Latent variable models. An introduction to factor, path, and structural analysis* (3rd edn). Hahway, NJ: Lawrence Erlbaum.

Lonner, W. J. & Ibrahim, F. A. (1996). Appraisal and assessment in cross-cultural counseling. In P. B. Pedersen & J. G. Draguns (Eds.), *Counseling across cultures* (4th edn, pp. 292–322). Thousand Oaks, CA: Sage Publications.

Marsh, H. W. (1996). Positive and negative global self-esteem: A substantively meaningful distinction or artifactors? *Journal of Personality and Social Psychology*, *70*, 810–819.

Matthews, A. (1990). Why worry? The cognitive function of anxiety. *Behaviour Research and Therapy*, *28*, 455–468.

Meloni, F. & Gana, K. (2001). Wording effects in the Italian version of the Penn State Worry Questionnaire. *Clinical Psychology and Psychotherapy*, *8*, 282–287.

Meyer, T. J., Miller, M. L., Metzger, R. L. & Borkovec, T. D. (1990). Development and validation of the Penn State Worry Questionnaire. *Behaviour Research and Therapy*, *28*, 487–495.

Molina, S. & Borkovec, T. D. (1994). The Penn State Worry Questionnaire: psychometric properties and associated characteristics. In G. Davey & F. Tallis (Eds.), *Worrying: Perspectives on theory, assessment and treatment* (pp. 265–297). Chichester: John Wiley & Sons.

Nordhus, I. H. & Pallesen, S. (2003). Psychological treatment of late-life anxiety: An empirical review. *Journal of Consulting and Clinical Psychology*, *71*, 643–651.

Robichaud, M., Dugas, M. J. & Conway, M. (2003). Gender differences in worry and associated cognitive-behavioral variables. *Anxiety Disorders*, *17*, 501–516.

Sarason, I. G. (Ed.). (1980). *Test anxiety: Theory, research, and applications*. Hillsdale, NJ: Lawrence Erlbaum.

Spielberger, C. D., Gorsuch, R. L., Lushene, R., Vagg, P. R. & Jacobs, G. A. (1983). *Manual for the State-Trait Anxiety Inventory (Form Y) ("Self-Evaluation Questionnaire")*. Palo Alto, CA: Consulting Psychological Press.

Stanley, M. A., Novy, D. M., Bourland, S. L., Beck, J. G. & Averill, P. M. (2001). Assessing older adults with generalized anxiety: A replication and extension. *Behaviour Research and Therapy*, *39*, 221–235.

SPSS, Inc. (2000). *SPSS (Version 10.1)* [Computer software]. Chicago, IL: SPSS.

Starcevic, V. (1995). Pathological worry in major depression – a preliminary report. *Behaviour Research and Therapy*, *33*, 55–56.

Stöber, J. (1995). Besorgnis: Ein Vergleich dreier Inventare zur Erfassung allgemeiner Sorgen. *Zeitschrift für Differentielle und Diagnostische Psychologie*, *16*, 50–63.

Stöber, J. (1998). Reliability and validity of two widely-used worry questionnaires: Self-report and self-peer convergence. *Personality and Individual Differences*, *24*, 887–890.

Tallis, F., Eysenck, M. W. & Mathews, A. (1992). A questionnaire for the measurement of nonpathological worry. *Personality and Individual Differences*, *13*, 161–168.

Turner, S. M., Beidel, D. C. & Stanley, M. A. (1992). Are obsessional thoughts and worry different cognitive phenomena? *Critical Psychology Review*, *12*, 252–270.

van Rijsoort, S., Emmelkamp, P. & Vervaeke, G. (1999). The Penn State Worry Questionnaire and the Worry Domains Questionnaire: structure, reliability and validity. *Clinical Psychology and Psychotherapy*, *6*, 297–307.

Wegner, D. M. & Zanakos, S. (1994). The White Bear Suppression Inventory. *Journal of Personality*, *62*, 615–640.

Wells, A. (1994). Attention and the control of worry. In G. Davey & F. Tallis (Eds.), *Worrying: Perspectives on theory, assessment and treatment* (pp. 91–114). Chichester: John Wiley & Sons.

Received 12 July 2004, accepted 1 March 2005

APPENDIX

The Norwegian version of the Penn State Worry Questionnaire

Instruksjon: Sett inn et tall fra 1 til 5 (se skalen nedenfor) som best beskriver hvor typisk eller karakteristisk hvert utsagn er for deg. Tallet settes på linjen utenfor hvert av utsagnene.

1	2	3	4	5
Ikke typisk		**Noe typisk**		**Meget typisk**

____ 1. Hvis jeg ikke har nok tid til alt jeg skal gjøre, bekymrer jeg meg ikke for det. (R)
____ 2. Jeg blir overveldet av mine bekymringer.
____ 3. Jeg pleier ikke å bekymre meg over ting. (R)
____ 4. Mange situasjoner får meg til å bli bekymret.
____ 5. Jeg vet jeg ikke burde bekymre meg for ting, men jeg klarer ikke å la være.
____ 6. Når jeg er under press, bekymrer jeg meg mye.
____ 7. Jeg bekymrer meg alltid for et eller annet.
____ 8. Jeg synes det er lett å stenge ute bekymringsfulle tanker. (R)
____ 9. Så fort jeg er ferdig med en oppgave, begynner jeg å bekymre meg for alt annet jeg må gjøre.
____ 10. Jeg bekymrer meg aldri over noe. (R)
____ 11. Når det ikke er noe mer jeg kan gjøre med en sak, bekymrer jeg meg ikke lenger for den. (R)
____ 12. Jeg har vært en bekymret person hele mitt liv.
____ 13. Jeg legger merke til at jeg har bekymret meg over ting.
____ 14. Når jeg først begynner å bekymre meg, kan jeg ikke holde opp.
____ 15. Jeg bekymrer meg hele tiden.
____ 16. Jeg bekymrer meg over alt jeg skal gjøre helt til det er unnagjort.

(R) indikerer snudde/omvendte ledd.

The original version of the Penn State Worry Questionnaire

Instruction: Enter the number that best describes how typical or characteristic each item is of you, putting the number next to each item.

1	2	3	4	5
Not at all typical		**Somewhat typical**		**Very typical**

____ 1. If I don't have enough time to do everything, I don't worry about it. (R)
____ 2. My worries overwhelm me.
____ 3. I don't tend to worry about things. (R)
____ 4. Many situations make me worry.
____ 5. I know I shouldn't worry about things, but I just can't help it.
____ 6. When I'm under pressure, I worry a lot.
____ 7. I am always worrying about something.
____ 8. I find it easy to dismiss worrisome thoughts. (R)
____ 9. As soon as I finish one task, I start to worry about everything else I have to do.
____ 10. I never worry about anything. (R)
____ 11. When there is nothing more I can do about a concern, I don't worry about it anymore. (R)
____ 12. I've been a worrier all my life.
____ 13. I notice that I have been worrying about things.
____ 14. When I first start worrying, I can't stop.
____ 15. I worry all the time.
____ 16. I worry about projects until they are all done.

(R) indicates a reverse-scored item.