



# Should we worry about how we measure worry? Insights from an updated version of the Italian Penn State Worry Questionnaire

Gioia Bottesi<sup>a,b,\*</sup>, Andrea Spoto<sup>a,b</sup>

<sup>a</sup> Department of General Psychology, University of Padova, Via Venezia 8, Padova 35131, Italy

<sup>b</sup> Centro di Ateneo dei Servizi Clinici Universitari Psicologici (SCUP), University of Padova, Via Niccolò Tommaseo 47a, Padova 35131, Italy

## ARTICLE INFO

### Keywords:

Worry  
Penn State Worry Questionnaire  
Classical Test Theory  
Item Response Theory  
Community samples

## ABSTRACT

**Background:** Research indicates a rise in self-reported worry, highlighting the need for updated psychometric tools. The Penn State Worry Questionnaire (PSWQ) assesses excessive worry and there is debate over whether only its 11 positively worded items should be used. This study aimed to clarify the factor structure and psychometric properties of the PSWQ and to explore worry features in two diverse Italian community samples from the 2010s and 2020s.

**Methods:** The 2020s sample included 674 participants (44.5 % female; Mage = 29.44 ± 13.20), while the 2010s sample comprised 411 individuals (61.6 % female; Mage = 36.64 ± 13.73). Methods from Classical Test Theory (CTT) were used to compare alternative PSWQ factor structures, assess the best-fitting model's reliability and validity, and evaluate measurement invariance (MI) across sexes in the 2020s sample. Item Response Theory (IRT) was applied to refine and confirm the best-fitting factor structure and to compare item and individual locations across samples.

**Results:** The 11-item one-factor model was the best fit and it showed excellent reliability and concurrent validity. MI across sexes was supported. IRT analyses suggested that items were slightly more difficult for the 2010s sample.

**Conclusions:** The PSWQ-11 is a valid and reliable tool for assessing worry in the Italian community. The findings suggest that societal issues as well as socio-demographic characteristics may contribute shaping differences in worry features across diverse historical contexts.

## Introduction

Worry is a cognitive process characterized by anticipating future threats, manifesting through “a chain of thoughts and images relatively affect laden and relatively uncontrollable; it represents an attempt to engage in mental problem solving on an issue whose outcome is uncertain but contains the possibility of one or more negative outcomes (Borkovec et al., 1983; p. 10)”. Worrying is an inherent facet of the human experience, and cognitive theories propose that worry may play a protective role, enabling individuals to prepare for and cope with upcoming challenges (Behar et al., 2009). However, when worry escalates to disproportionate or uncontrollable levels—reaching a pathological state—it can exert detrimental effects on physical and psychological health. Pathological worry constitutes a central feature of Generalized Anxiety Disorder (GAD) (American Psychiatric Association [APA], 2022), and it manifests as a transdiagnostic construct observable

across various psychopathologies, including anxiety and depressive disorders (e.g. Ehrling & Watkins, 2008; McEvoy et al., 2013; Warren et al., 2021).

Both research and clinical evidence point towards an upward trend in self-reported worry in recent years (e.g. Carleton et al., 2019; Davey et al., 2022). Societal changes, economic pressures, and the widespread impact of digital technology have all been identified as potential contributors to the increased levels of worry and anxiety (Davey, 2018). Additionally, some recent global events may have further significantly augmented normative levels of worry, such as the COVID-19 pandemic, extreme weather events, and geopolitical conflicts (Freeston et al., 2020; Limone et al., 2022; Taylor, 2020). Given these premises, the importance of employing robust psychometric measures adequately capturing this construct and incorporating updated normative data is paramount.

The Penn State Worry Questionnaire (PSWQ) is an extensively used self-report questionnaire designed to measure excessive and

\* Corresponding author.

E-mail address: [gioia.bottesi@unipd.it](mailto:gioia.bottesi@unipd.it) (G. Bottesi).

<https://doi.org/10.1016/j.ijchp.2025.100579>

Received 11 December 2024; Accepted 23 April 2025

Available online 2 May 2025

1697-2600/© 2025 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

uncontrollable worry (Meyer et al., 1990). Due to its established validity and reliability, it is frequently utilized as a screening instrument for pathological worry (Porta-Casteràs et al., 2020). It comprises 11 items framed in a positive manner and 5 items expressed with negative wording, and it was initially conceived as a unifactorial measure of worry. However, investigations into its factor structure through subsequent research have produced inconsistent findings.

Fresco et al. (2002) found that the 11 positively worded items within the PSWQ were associated with a "worry engagement" factor, while the 5 negatively worded items were linked to an "absence of worry" factor. Nevertheless, critics argue against two-factor solution studies, suggesting they may overlook method variance or effects common in measures with both positive and negative items (e.g. Brown, 2003). Hazlett-Stevens et al. (2004) later examined a different alternative model introducing a "method factor" associated with negative wording. They found that a model wherein all items from the PSWQ loaded onto a single worry factor and all negatively worded items loaded also onto a negative wording method factor (i.e., the method-factor model) showed a fit equivalent to the two-factor model. This finding received support in various studies, and the method-factor model was retained as a simpler explanation for assessing worry based on its presence/absence (Hazlett-Stevens et al., 2004; Jeon et al., 2017; Lim et al., 2008; Liu et al., 2022; Zhong et al., 2009). Among them, a recent study involving American undergraduate students, Dutch high school students, and American adults with psychopathology also found that items phrased positively, as opposed to negatively, showed concurrent validity with measures and diagnoses of anxiety and depression symptoms (Liu et al., 2022).

To date, only a few studies on Spanish-speaking samples have also examined an 11-item version of the PSWQ (Carbonell-Bàrtoli & Tume-Zapata, 2022; Padros-Blazquez et al., 2018; Ruiz et al., 2018), which excludes the 5 negatively worded items. These studies all yielded promising results, concluding that this abbreviated version is valid and reliable. However, findings were mainly based on undergraduate samples and/or did not compare the 11-item one-factor model with existing competing models.

### *The current study*

This study was designed to address existing gaps in the literature on the PSWQ. To this end, we focused on the Italian version of the PSWQ, for which the available information is outdated and a rigorous validation is lacking (Meloni & Gana, 2001; Morani et al., 1999). Indeed, Morani et al. (1999) administered the 16 PSWQ items to 388 individuals from the general population, aged 18–86 years. They only reported descriptive statistics (means, standard deviations, and ranges) by age and sex, along with the overall Cronbach's alpha value (0.85). No formal test of factor structure or the main psychometric properties of the measure were conducted. Subsequently, Meloni and Gana (2001) examined the factor structure of the PSWQ on a sample of 142 participants (mainly undergraduates) using both principal component analysis and confirmatory factor analysis. They also assessed external validity by using a measure of self-actualization but did not explore issues about concurrent validity with anxiety and depression symptom measures. The preliminary nature of the study prevented drawing definitive conclusions about the factor structure and the psychometric properties of the questionnaire, despite the authors suggested that the PSWQ could be enhanced by eliminating negatively worded items (Meloni & Gana, 2001).

In our study, we aimed to overcome the limitations identified in previous literature by integrating methods from two psychometric theories: Classical Test Theory (CTT; e.g., Lord & Novik, 1968; Novik, 1966) and Item Response Theory (IRT; e.g., Lord, 1980). Most of the studies on the PSWQ to date have focused on the exclusive application of CTT-related methods, such as exploratory and confirmatory factor analysis (CFA) and classical model fit comparisons. In contrast, only a

few studies have explored the application of IRT, and these have been limited to very specific aims and narrowly defined populations (e.g., Schroeder et al., 2019; Wu et al., 2013). However, as mentioned earlier, recent literature increasingly questions the existence of the two traditionally accepted factors, prompting consideration of IRT's applicability in PSWQ research as a tool for refining the instrument and providing updated norms.

Our first objective was to enhance understanding of the PSWQ's factor structure and psychometric characteristics using a recent, large Italian community sample. Secondly, we aimed to investigate whether distinct community samples recruited across different time periods – likely shaped by distinct societal and economic conditions – may exhibit differences in worry features. To our knowledge, this is the first study examining these issues through the joint application of CTT and IRT methods.

### *Data analysis strategy and study hypotheses*

The first aim was addressed using a sample recruited in 2021 and 2022 (i.e., "2020s sample"). Specifically:

- We initially applied CTT to compare the original 16-item one-factor model (Meyer et al., 1990), the 16-item two-factor model (Fresco et al., 2002), the 16-item method-factor model (Hazlett-Stevens et al., 2004), and the 11-item one-factor model (Carbonell-Bàrtoli & Tume-Zapata, 2022; Padros-Blazquez et al., 2018; Ruiz et al., 2018). The model that best fitted the data was then analyzed using IRT methods to verify the overall model performance and item-specific parameters. Several authors have suggested that enhancing the PSWQ might involve retaining only its 11 positively worded items (see Liu et al., 2022 for details). Thus, we expected the 11-item one-factor model to emerge as the best-fitting model compared to the other solutions and that the model's adequate fit to the data would be confirmed through IRT analysis (H1a).
- We assessed concurrent validity by examining the correlations between the PSWQ and the depression, anxiety, and stress scales of the Depression Anxiety Stress Scales-21 (DASS-21; Bottesi et al., 2015). Consistent with the literature (e.g. Liu et al., 2022; Ruiz et al., 2018), we hypothesized strong positive associations between the PSWQ and DASS-21 scores (H1b).
- We investigated Measurement Invariance (MI) across sexes. MI of the PSWQ across sexes has been supported, despite it has only been examined in a few studies (Carbonell-Bàrtoli & Tume-Zapata, 2022; Ruiz et al., 2018). Thus, we hypothesized that the Italian version of the PSWQ would also be invariant across females and males (H1c).
- We calculated descriptive statistics and percentiles by sex to provide up-to-date Italian normative data. Based on the literature (e.g., Bottesi et al., 2018; Gould & Edelstein, 2010), we expected women to show higher levels of worry than men (H1d).

To address the second aim, we used an additional Italian community sample recruited between 2012 and 2018 ("2010s sample"). We conducted the same IRT analyses to the PSWQ scores from the 2010s sample as those conducted on the 2020s sample and then compared item characteristics and individual location parameters across the two groups. As this is the first study to explore this topic, we did not formulate specific hypotheses. Nonetheless, we anticipated potential differences in worry-related features among samples recruited across distinct historical contexts, based on the premise that societal, cultural, and environmental factors can shape the experience of worry (H2). This expectation was informed by prior research on non-clinical populations in the US and UK, which has documented a rise in reported worry levels over the past two decades (Carleton et al., 2019; Davey et al., 2022).

## Materials and methods

### Participants and procedure

The 2020s sample initially comprised 938 individuals from the Italian community. They were recruited through snowball sampling and invited to complete an online survey focusing on personality traits and psychological features between May 2020 and March 2022. We collected data via a SurveyMonkey link, which was distributed across social networks (e.g., Facebook, Instagram, and Telegram). Participants provided informed consent before completing a socio-demographic schedule and the questionnaires. The survey took approximately 50 min to complete, and no rewards were offered for participation. We retained only participants that did not enter the study during the nationwide lockdown due to the COVID-19 outbreak. Thus, only data collected between November 2021 and March 2022 were used ( $N = 812$ ). Subsequently, we excluded participants that accessed the survey link but did not complete any part of the study—either the demographic questions or the questionnaires ( $N = 128$ ). The final sample consisted of 674 participants with complete data, of whom 44.5 % were female, and ages ranged from 18 to 65 years old ( $M = 29.44$ ;  $SD = 13.20$ ). The mean education level was 13.92 years ( $SD = 2.69$ , range = 5–25). Regarding relationship status, 74 % of participants were single, 22.8 % were married or cohabitating, 2.7 % were separated or divorced, and 0.5 % were widowed. Concerning occupation, 48.2 % were employed, while the remaining were not (they were students, housekeepers, retired, or unemployed).

The 2010s sample was drawn from a dataset of individuals recruited from the Italian community between 2012 and 2018, with some of the data having been previously published in research by Bottesi et al. (2017, 2019). This sample consisted of 411 individuals (61.6 % female) aged between 18 and 65 years ( $M = 36.64$ ,  $SD = 13.73$ ) and with a mean education level of 14.65 years ( $SD = 3.39$ ; range = 5–30). Relationship status was 53.3 % single, 41.4 % married/cohabitating, 4.9 % separated/divorced, and 0.4 % widowed. Regarding occupation, 56 % were employed, while the remaining were not. The 2010s sample included a higher proportion of females ( $\chi^2_{(1)} = 29.69$ ,  $p < .001$ ), was older ( $F_{(1, 1083)} = 73.73$ ,  $p < .001$ ), had a higher level of education ( $F_{(1, 1081)} = 15.43$ ,  $p < .001$ ), was more likely to be in a stable romantic relationship ( $\chi^2_{(3)} = 49.44$ ,  $p < .001$ ), and had a lower rate of unemployment ( $\chi^2_{(1)} = 22.49$ ,  $p < .001$ ) compared to the 2020s sample. Thus, the two samples significantly differed in their socio-demographic composition.

This research was conducted in accordance with the Declaration of Helsinki and it was approved by the Ethics Committee of Psychological Sciences of the University of Padova (Italy).

### Measures

The *Penn State Worry Questionnaire*, a 16-item inventory designed to assess the tendency to worry excessively and uncontrollably. It measures trait worry by asking individuals to rate the extent to which each item is typical of them on a 5-point Likert scale (1 = “Not at all typical of me”, 5 = “Very typical of me”). Internal consistency and test-retest reliability of the original version were good in non-clinical and clinical samples (Meyer et al., 1990).

The *Depression Anxiety Stress Scales-21* (DASS-21; Lovibond & Lovibond, 1995), a 21-item measure evaluating depression, anxiety, and stress symptoms over the previous week on a 4-point Likert scale (0 = “Did not apply to me at all”, 3 = “Applied to me very much, or most of the time”). A total score as well as three subscale scores (i.e., Anxiety, Depression, Stress) can be computed. The total and subscale scores of the Italian version showed an excellent internal consistency in both clinical and non-clinical samples (Bottesi et al., 2015).

### Data analysis

All analyses were conducted using R 4.4.0 statistical software. Specifically, model comparisons and MI were performed with the *lavaan* package, initial data screening and scale reliability were carried out using the *psych* package, and IRT analysis was conducted with the *mir* package. Additionally, Jamovi version 2.3.28 was used, particularly for the *snowIRT* module. The scripts and data are available on OSF: <https://doi.org/10.17605/OSF.IO/YDGPB>.

### CTT methods

Methods from CTT were adopted to analyze data of the 2020s sample. Specifically, they were used to: (i) compare the alternative factor structures of the PSWQ identified in the literature, (ii) assess the reliability and validity of the best-fitting model, (iii) evaluate MI across sexes for the best-fitting model.

We performed model comparisons estimating the fit of different models and comparing the following fit indices:  $\chi^2$ ,  $\chi^2/df$ , Tucker-Lewis Index (TLI), Comparative Fit Index (CFI), Root Mean Square Error of Approximation (RMSEA) and Standardized Root Mean Square Residual (SRMR). Acceptance criteria were based on current literature guidelines:  $RMSEA < 0.08$ ;  $SRMR < 0.08$  TLI and CFI  $> 0.95$  (Hu & Bentler, 1999). Given the data distribution (see paragraph 3.1), all models were fitted using the Diagonal Weighted Least Squares (DWLS) as estimator. The reliability of the resulting scale was estimated by means of McDonald's  $\omega$ , while concurrent validity was analyzed by means of the Spearman  $\rho$  coefficient between the PSWQ and DASS-21 scores.

MI across sexes (300 males vs. 364 females) was tested by imposing increasing restrictions to the model and evaluating the modifications in fit indices. To this aim, a multi-group CFA was performed to assess sex invariance at configural, metric, and scalar levels. Configural invariance examines if the unconstrained model is equal across sexes. Metric invariance implies that the magnitude of loading is similar across the two sexes. Scalar invariance implies similarity on factor loadings and item intercepts between the two groups. We considered  $\Delta CFI \geq -0.005$  and the  $\Delta RMSEA \leq 0.010$  or  $\Delta SRMR \leq 0.005$  for invariance (Chen, 2007). Moreover, a chi-square ( $\chi^2$ ) was utilized to test invariance or partial invariance by comparing the constrained with the unconstrained model.

### IRT methods

IRT methods were used to: (i) refine and further confirm the best fitting solution for the factor structure of PSWQ emerged from CTT, and (ii) compare items and individuals' locations between the 2010s and 2020s samples.

Due to the nature of the PSWQ items, which are based on a 5-point Likert scale, we chose to use the Rating Scale Model (RSM; Andrich, 1978) on both the 2020s and 2010s samples. The RSM assumes that the individual probability of selecting a given score for an item is a logistic function of the person's location parameter and the difficulty of item steps (Andrich, 1978). In the RSM, item category thresholds are constrained to be equal for all items. Naturally, items are still allowed to vary in difficulty (i.e., different item location parameters are permitted; Masters & Wright, 1984). Importantly, each item provides equal weight of information to the scale, sharing the same discrimination parameters (i.e., slopes, Bacci et al., 2014). Model fit was assessed using traditional indices such as RMSEA, TLI, CFI. Additionally, consistent with IRT literature, *infit* and *outfit* indices of items were utilized to analyze items specific behavior. The former is crucial for identifying items where unexpected responses occur among respondents whose estimated ability closely matches the item's difficulty. Outfit, on the other hand, detects unexpected responses to items that may not be aligned with a respondent's ability level. Both indices are expected to approximate 1, with values below 0.5 that indicate underfit of the item, while values above 1.5 suggest overfit. Finally, person reliability was computed. This index refers to the precision and stability of the measure in estimating

individual's ability parameter (e.g., Embretson & Reise, 2000).

## Results

### Preliminary descriptive analyses

As a preliminary step, an exploration of the 2020s sample data was conducted to examine descriptive statistics and distributions for each of the 16 PSWQ items (see Supplementary Materials, Table S1). The Shapiro-Wilk test highlighted that none of the items was normally distributed. For all the items, after inverting the reverse worded ones, all the scale values were observed, and the average score was between 1.85 (item PSWQ\_10) and 3.25 (item PSWQ\_6).

### Models comparison

Four different models for the factor structure of the PSWQ were compared: the 16-item one-factor model (Model 1), the 16-item two-factor model (Model 2), the 16-item method-factor model (Model 3), and the 11-item one-factor model (Model 4). All fit indices are displayed in Table 1.

The inspection of fit indices suggested that the best model was Model 4. All factor loadings were strong and significant, ranging between 0.695 for item PSWQ\_9 and 0.914 for item PSWQ\_5. Table 2 reports the standardized factor loadings.

### IRT analysis

The RSM fit indices of the 11-item one-factor model revealed a mixed fit to the data (RMSEA = 0.100; SRMR = 0.084; TLI = 0.972, CFI = 0.957). Since in the RSM item category thresholds are constrained to be equal for all items, we have only one set of estimates of these parameters whose values were: 0.839 for the first threshold, 1.503 for the second, 3.017 for the third and 3.822 for the last. Item's difficulty parameters are displayed in the third column of Table 2. The lowest value was observed for item PSWQ\_6, which resulted to be the "less difficult" one. On the contrary, the highest value was observed for item PSWQ\_15, which was the "most difficult" one. The person parameters ranged between -3.389 and 4.214. The items tended to be, in general, centered on a level of difficulty rather low.

Moreover, we analyzed *infit* and *outfit* indices for all the items and checked the residuals of the model. With respect to the former, we verified that all these parameters fell within the 0.5–1.5 range, confirming the goodness of all the items in the model as well as the goodness of the model itself. Results are reported in the last two columns of Table 2. Only item PSWQ\_9 reported slightly high values (in any case below the mentioned thresholds). Residual analysis was conducted using Q3 statistics (Chen et al., 2013) to further evaluate model fit. The Q3 values were generally low and centered around zero (Mean = -0.094; Min = -0.277; Max = 0.246), with no evidence of systematic local dependence among item pairs. These results fall well within commonly accepted thresholds (e.g., Christensen et al., 2017),

**Table 1**

Fit indices of the four competing models tested on the 2020s sample.

	$\chi^2$	df	p-value	CFI	TLI	RMSEA	SRMR
<b>Model 1</b>	952.460	104	<0.001	.988	.986	.110	.081
<b>Model 2</b>	437.660	103	<0.001	.995	.995	.069	.051
<b>Model 3</b>	357.939	99	<0.001	.996	.996	.069	.045
<b>Model 4</b>	201.860	44	<0.001	.997	.997	.069	.038

*Note.* Model 1 = the original 16-item one-factor model; Model 2 = the 16-item two-factor model; Model 3 = the 16-item method-factor model; Model 4 = the 11-item one-factor model. CFI = Comparative Fit Index; TLI = Tucker-Lewis Index; RMSEA = Root Mean Square Error of Approximation; SRMR = Standardized Root Mean Square Residual.

**Table 2**

Standardized factor loading and RSM IRT item parameters of the 11 items of PSWQ included in Model 4 (2020s sample).

Item	Std. Loading	IRT Difficulty	<i>infit</i>	<i>outfit</i>
<b>PSWQ_2</b>	.764	-1.515	1.035	1.001
<b>PSWQ_4</b>	.879	-2.055	.647	.632
<b>PSWQ_5</b>	.914	-1.972	.643	.607
<b>PSWQ_6</b>	.781	-2.659	1.010	.978
<b>PSWQ_7</b>	.877	-1.491	.773	.722
<b>PSWQ_9</b>	.695	-1.399	1.371	1.375
<b>PSWQ_12</b>	.756	-2.293	1.058	1.116
<b>PSWQ_13</b>	.834	-1.473	.996	.925
<b>PSWQ_14</b>	.874	-1.257	.775	.723
<b>PSWQ_15</b>	.901	-0.866	.743	.632
<b>PSWQ_16</b>	.736	-2.219	1.046	1.046

supporting the assumption of local independence. Last, we computed person reliability of the scale, and the value was excellent and equal to 0.910.

### Reliability and concurrent validity

The McDonald's  $\omega$  for the 11-item PSWQ was 0.943 (95 % C.I.: 0.936–0.949), indicating a very strong reliability. All items were strongly related to one another as displayed by the values of McDonald's  $\omega$  dropping one item at a time and the item-rest correlations (Table 3). All item-rest correlations were largely above 0.600, and by eliminating each item, the overall value of McDonald's  $\omega$  always worsened, indicating the solid reliability of the unidimensional solution proposed by this model.

With respect to concurrent validity, all correlations between the total score of the 11-item PSWQ and the DASS-21 subscales were significant and moderate-to-strong in size: DASS-21 Depression:  $\rho=0.548$ ,  $p < 0.01$ ; DASS-21 Anxiety:  $\rho=0.510$ ,  $p < 0.01$ ; DASS-21 Stress:  $\rho=0.624$ ,  $p < 0.01$ .

### Measurement invariance

Results on configural, metric and scalar MI across sexes for the 11-item one-factor model are reported in Table 4.

While configural and metric invariance were clearly confirmed, results for scalar invariance were slightly more ambiguous. In particular, the difference between the  $\chi^2$  of the model with the loadings set equal across groups and the model where both loadings and intercepts were set equal across the groups was significant, indicating a worse fit of the second model. Nonetheless, all the other differences between important fit indices (namely  $\Delta CFI$ ,  $\Delta RMSEA$ , and  $\Delta SRMR$ ) were within the tolerance limits, confirming scalar invariance. This fact, together with the well-known reduced reliability of the  $\chi^2$  statistics with large samples (Chen, 2007; Heung & Rensvold, 2002; Kline, 2015; Little, 2013), led us to conclude that scalar invariance between the two sexes seems to be

**Table 3**

McDonald's  $\omega$  if item dropped and item-rest correlation of the PSWQ-11 (2020s sample).

Item	McDonald's $\omega$ if item dropped	Item-rest correlation
<b>PSWQ_2</b>	.939	.701
<b>PSWQ_4</b>	.935	.807
<b>PSWQ_5</b>	.933	.840
<b>PSWQ_6</b>	.939	.700
<b>PSWQ_7</b>	.934	.810
<b>PSWQ_9</b>	.942	.625
<b>PSWQ_12</b>	.939	.691
<b>PSWQ_13</b>	.937	.761
<b>PSWQ_14</b>	.935	.795
<b>PSWQ_15</b>	.935	.803
<b>PSWQ_16</b>	.940	.683



**Table 4**  
Measurement Invariance between males (N=300) and females (N=364) in the 2020s sample.

Model	$\chi^2$	df	CFI	RMSEA	SRMR	Model comparison	$\Delta\chi^2$	$\Delta df$	p	$\Delta CFI$	$\Delta RMSEA$	$\Delta SRMR$
Configural	206.968	88	.998	.063	.041							
Metric	223.306	98	.998	.062	.043	Metric vs Configural	16.338	10	.09	<0.001	-0.001	.002
Scalar	291.419	130	.997	.061	.041	Scalar vs Metric	68.113	32	<0.05	-0.001	-0.001	-0.002

Note. CFI = Comparative Fit Index; RMSEA = Root Mean Square Error of Approximation; SRMR = Standardized Root Mean Square Residual.

confirmed.

Scale descriptives and norms

Table 5 illustrates the main descriptive indices as well as the percentile distribution, separated by sex, of the total score of the 11-item PSWQ. The average score obtained by females was significantly higher than that of males ( $t_{672}=-8.283$ ;  $p<.01$ ; Cohen's  $d=-0.643$ ). Consequently, the tentative screening threshold for males (95th percentile) is 47, while the same threshold for females is fixed at a score of 51.

IRT analysis on the 2010s sample

The RSM IRT analysis indicated that the fit of the 11-item one-factor model on the 2010s sample data was slightly better than the one of the other sample (RMSEA= 0.081; SRMR= 0.073; TLI= 0.976, CFI= 0.963). The range of item difficulty was similar to that observed in the 2020s sample, although both the minimum and maximum values were higher in the 2010s one. Specifically, the lowest value was again observed for item PSWQ\_6, with a value of -2.365, while the highest value was still for item PSWQ\_15, at -0.124. Overall, items appeared to be slightly more "difficult" for the 2010s sample (see Supplementary Materials, Table S2).

Infit and outfit measures were all in the desired range with no exceptions. Some changes in the order of item difficulty occurred. In particular, while the extreme items (namely PSWQ\_6 and PSWQ\_15) remained the same in both samples and the order of the first 5 items was essentially stable between the two samples (PSWQ\_6, PSWQ\_16, PSWQ\_12, PSWQ\_4 and PSWQ\_5), items PSWQ\_2 and PSWQ\_7 in the 2010s sample resulted more difficult than PSWQ\_13, PSWQ\_9, and PSWQ\_14. The opposite trend was observed for the more recent data collection. Person reliability was very good also in this sample (0.881).

Discussion

In this study, we aimed to achieve two objectives: first, to clarify the factor structure and psychometric properties of the PSWQ; and second, to explore whether distinct community samples recruited across different time periods may be characterized by differences in worry features. To this end, we utilized the Italian version of the PSWQ and

**Table 5**  
Descriptive indices, decile distribution and 95th percentile separated by sex of the PSWQ-11 (2020s sample).

	Sex	
	Males	Females
Mean (SD)	26.163 (10.397)	32.960 (10.821)
Min-Max (Range)	11–55 (44)	11–55 (44)
10th Percentile	13.000	19.000
20th Percentile	16.000	24.000
30th Percentile	19.000	26.000
40th Percentile	22.000	29.000
50th Percentile	25.000	32.000
60th Percentile	27.000	36.000
70th Percentile	31.000	39.000
80th Percentile	35.000	43.000
90th Percentile	41.000	48.000
95th Percentile	46.350	51.000

employed methods from both CTT and IRT.

Factor structure and psychometric properties of the PSWQ

The factor structure and psychometric properties of the PSWQ were tested on the 2020s sample which, to our knowledge, represents the most up-to-date normative Italian sample for this questionnaire. Consistent with expectations, the 11-item one-factor model emerged as the best-fitting solution compared to other models reported in the literature, despite the 16-item method-factor model demonstrating similarly good fit indices (H1a). This result contributes to the ongoing debate questioning the validity of the two factors identified by [Fresco et al. \(2002\)](#), namely “worry engagement” and “absence of worry”, and aligns with recent literature advocating for a streamlined version of the PSWQ, retaining only the 11 positively worded items (i.e., the PSWQ-11) ([Carbonell-Bàrtoli & Tume-Zapata, 2022](#); [Padrós-Blázquez et al., 2018](#); [Ruiz et al., 2018](#)). Some studies pointed out the conceptual ambiguity and potential method effects introduced by the negatively worded items ([Hazlett-Stevens et al., 2004](#); [Liu et al., 2022](#); [Meloni & Gana, 2001](#)), thus supporting their removal on both theoretical and methodological grounds. Moreover, from the perspective of parsimony, we believe that the PSWQ-11 offers a cleaner factorial structure, reduces the likelihood of response biases (e.g., acquiescence or confusion due to reverse wording), and facilitates more straightforward interpretation of scores. Furthermore, the practical utility of the 11-item version is particularly important in both research and clinical contexts. Shorter instruments are less burdensome for respondents, improve completion rates, and are especially advantageous in time-constrained settings such as primary care, mental health screenings, or large-scale surveys. The streamlined format also enhances the feasibility of repeated assessments over time, which is critical in longitudinal designs or treatment monitoring ([Porta-Casteràs et al., 2020](#)). IRT analyses further supported the acceptability of the 11-item one-factor model. Additionally, the inspection of item parameters indicated that, overall, the items are centered on a relatively low level of difficulty. This suggests that these items are better suited for discriminate between individuals with lower levels of worry, rather than those in more severe situations. In our view, this could be a desirable feature of this questionnaire, since it is frequently used to detect pathological worry in screening procedures ([Porta-Casteràs et al., 2020](#)). Lastly, the analysis of *infit* and *outfit* indices for all items confirmed the overall acceptability of fit for both the individual items and the model as a whole. The only exception was item PSWQ\_9 (‘As soon as I finish one task, I start to worry about everything else I have to do’), which showed a slight misfit according to the *infit* and *outfit* indices. Future research may consider rephrasing this item, as psychometric standards generally discourage the use of items containing more than one sentence (e.g., [Emberston, 1999](#); [Henson & Douglas, 2005](#)), an issue which might - at least partially - explain this result.

Both McDonald's  $\omega$  and IRT person reliability indices showed excellent values, providing additional support for the strong reliability of this unidimensional model. Concurrent validity was also confirmed, as the total score of the PSWQ-11 exhibited significant, moderate-to-strong correlations with the DASS-21 subscales (H1b). These findings not only mirror previous studies demonstrating strong associations with anxiety and depression measures ([Liu et al., 2022](#); [Ruiz et al., 2018](#)) but are also in line with literature suggesting the transdiagnostic nature of worry ([Ehring & Watkins, 2008](#); [McEvoy et al., 2013](#); [Warren et al.,](#)

2021).

Consistent with our hypothesis and previous research (Carbonell-Bàrtoli & Tume-Zapata, 2022; Ruiz et al., 2018), we found that the 11-item one-factor model is invariant across sexes. This ensures that mean differences in the PSWQ-11 total scores are unbiased between males and females, meaning that observed differences accurately reflect true variability in worry levels between groups. Thus, the use of the total score for making valid inferences across sexes, at least in the Italian version of the scale, is supported. When comparing the mean scores of females and males, we observed that females scored significantly higher than males, corroborating both our expectations and the prevailing consensus in the literature, which indicates that women tend to worry more than men (Bottesi et al., 2018; Gould & Edelstein, 2010) (H1d).

### Worry features in the 2010s and 2020s

Beyond enabling us to refine and confirm the best-fitting factor structure of the PSWQ as identified through CTT, IRT also provided valuable insights into the construct of worry and its features over the past decade by examining item characteristics and individual location parameters across our two community samples.

First, the RSM IRT analysis revealed that the 11-item one-factor model had a good fit also in the 2010s sample, which demonstrates that the PSWQ-11 is a reliable tool for measuring worry across different decades. With respect to item difficulty, in both the 2010s and the 2020s samples the lowest value was observed for item PSWQ\_6 ('When I am under pressure, I worry a lot'), indicating that this item is more likely to receive higher scores and is thus less clinically relevant. Conversely, the highest value was observed for item PSWQ\_15 ('I worry all the time'), making it the most clinically relevant, as it tends to receive lower scores in general population. This finding aligns with established cognitive models of worry (see Behar et al., 2009; Borkovec et al., 1983), which posit that worry is a common experience (i.e., it is quite likely that most people report worrying under pressure) that becomes pathological when it is perceived as excessive and persistent (i.e., it is unlikely that most people report being constantly worrying).

Interestingly, the items appeared to be slightly more difficult for the 2010s sample than for the 2020s one. Moreover, we noted some shifts in the order of item difficulty between the two samples: specifically, in the 2010s sample, PSWQ\_2 ('My worries overwhelm me') and PSWQ\_7 ('I am always worrying about something') were more difficult than PSWQ\_13 ('I notice that I have been worrying about things'), PSWQ\_9 ('As soon as I finish one task, I start to worry about everything else I have to do'), and PSWQ\_14 ('Once I start worrying, I cannot stop'). The opposite trend was observed in the 2020s sample. Overall, these results may point to an increase in worry within the general population over time. Indeed, the fact that the items are generally less difficult in the 2020s tentatively implies that people today may be more prone to experiencing worry. Additionally, specific characteristics of worry that were considered more clinically relevant in the 2010s sample, particularly those related to its intensity and pervasiveness, may be viewed as easier to experience—or even more normative—in the current decade (H2). These findings may align with previous research indicating increasing levels of worry over the past two decades, a trend that has been linked to global crises, rapid technological advancement, and the constant influx of information (Carleton et al., 2019; Davey et al., 2022). However, the observed differences between the 2010s and 2020s samples may also be attributable to demographic variations in sample composition, rather than—or in addition to—genuine societal changes. Notably, the 2020s sample was younger, and included fewer employed individuals compared to the 2010s sample, and previous research has demonstrated that certain sociodemographic characteristics—such as younger age and unemployment—are consistently associated with higher levels of worry across different historical contexts (e.g., Hassen & Adane, 2021; Schwartz & Melech, 2000).

### Limitations

There are several limitations in the present study. First, due to its cross-sectional design, we were unable to address either the test-retest reliability or the longitudinal MI of the PSWQ-11 structure in the 2020s sample. Future research should examine this issue to gain a more precise and comprehensive understanding of the questionnaire's factor stability over time. Moreover, this sample was recruited using snowball sampling through social media platforms—an approach that may limit the generalizability of the findings due to potential selection biases (e.g., a higher likelihood of including younger or more technologically savvy individuals). Second, we recognize that comparing item characteristics and individual location parameters between two distinct samples does not allow for definitive conclusions regarding the changes in worry over time. This further underscores the need for longitudinal studies. Third, while our samples covered a broad and representative age range (18–65 years), their composition did not allow for separate analyses by age group. Moreover, we only included community individuals. Future research should address these limitations to assess the known-group validity of the PSWQ-11 and to establish clearer cut-off scores. The normative data we reported, expressed as percentiles, do not equate to clinical significance. Indeed, a score at the 95th percentile in a community sample may indicate subclinical levels of pathological worry rather than a pathological condition. While such information is not diagnostic and requires validation in clinical populations, it remains valuable for preliminary screening and identifying individuals who may benefit from further assessment.

### Conclusions

Overall, the results of our study contribute to the existing literature on the measurement of worry through the PSWQ, highlighting that this shortened 11-item version may serve as a valid and reliable screening tool for assessing this construct. Furthermore, the comparison between our 2010s and 2020s samples in terms of IRT-based items and individuals' parameters suggested that contextual and demographic factors might influence worry features. As a whole, we believe that our findings outline the importance of using robust psychometric measures that incorporate updated normative data to help clinicians and researchers interpret questionnaire scores more effectively. This issue assumes crucial relevance in clinical practice, where the use of psychometrically sound assessment tools is critical not only during the screening and assessment phases, but also in evaluating clinically significant changes after treatment (APA, 2020).

### Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

### Acknowledgments

None.

### Supplementary materials

Supplementary material associated with this article can be found, in the online version, at [doi:10.1016/j.ijchp.2025.100579](https://doi.org/10.1016/j.ijchp.2025.100579).

### References

- American Psychiatric Association. (2022). Diagnostic and statistical manual of mental disorders, 5th ed., Text Revision. Washington, DC.
- American Psychological Association. (2020). APA guidelines for psychological assessment and evaluation. Washington, DC.

- Andrich, D. (1978). Application of a psychometric rating model to ordered categories which are scored with successive integers. *Applied Psychological Measurement*, 2(4), 581–594. <https://doi.org/10.1177/0146621678002004>
- Bacci, S., Bartolucci, F., & Gnaldi, M. (2014). A class of multidimensional latent class IRT models for ordinal polytomous item responses. *Communications in Statistics: Theory and Methods*, 43(4), 787–800. <https://doi.org/10.1080/03610926.2013.827718>
- Behar, E., DiMarco, I. D., Hekler, E. B., Mohlman, J., & Staples, A. M. (2009). Current theoretical models of generalized anxiety disorder (GAD): Conceptual review and treatment implications. *Journal of Anxiety Disorders*, 23, 1011–1023. <https://doi.org/10.1016/j.janxdis.2009.07.006>
- Borkovec, T. D., Robinson, E., Pruzinsky, T., & DePree, J. A. (1983). Preliminary exploration of worry: Some characteristics and processes. *Behavior Therapy*, 21, 9–16. [https://doi.org/10.1016/0005-7967\(83\)90121-3](https://doi.org/10.1016/0005-7967(83)90121-3)
- Bottesi, G., Ghisi, M., Altoè, G., Conforti, E., Melli, G., & Sica, C. (2015). The Italian version of the depression anxiety stress scales-21: Factor structure and psychometric properties on community and clinical samples. *Comprehensive Psychiatry*, 60, 170–181. <https://doi.org/10.1016/j.comppsy.2015.04.005>
- Bottesi, G., Ghisi, M., Sica, C., & Freeston, M. H. (2017). Intolerance of uncertainty, not just rigid experiences, and compulsive checking: Test of a moderated mediation model on a non-clinical sample. *Comprehensive Psychiatry*, 73, 111–119. <https://doi.org/10.1016/j.comppsy.2016.11.014>
- Bottesi, G., Martignon, A., Cerea, S., & Ghisi, M. (2018). Worry and associated cognitive features in Italian university students: Does gender make a difference? *Personality and Individual Differences*, 126, 38–43. <https://doi.org/10.1016/j.paid.2018.01.016>
- Bottesi, G., Noventa, S., Freeston, M. H., & Ghisi, M. (2019). Seeking certainty about Intolerance of Uncertainty: addressing old and new issues through the Intolerance of Uncertainty Scale-Revised. *PLoS ONE*, 14(2), Article e0211929. <https://doi.org/10.1371/journal.pone.0211929>
- Brown, T. A. (2003). Confirmatory factor analysis of the Penn State worry questionnaire: Multiple factors or method effects? *Behavior Research and Therapy*, 41(12), 1411–1426. [https://doi.org/10.1016/s0005-7967\(03\)00059-7](https://doi.org/10.1016/s0005-7967(03)00059-7)
- Carbonell-Bàrtoli, E. L., & Tume-Zapata, V. N. (2022). Cuestionario de Preocupación de Pensilvania (PSWQ 11): Evidencias Psicométricas en una muestra de adultos de lima metropolitana. *PsiqueMag*, 11(2), 31–41. <https://doi.org/10.18050/psiquemag.v11i2.2108>
- Carleton, R. N., Desgagné, G., Krakauer, R., & Hong, R. Y. (2019). Increasing intolerance of uncertainty over time: The potential influence of increasing connectivity. *Cognitive Behaviour Therapy*, 48(2), 121–136. <https://doi.org/10.1080/16506073.2018.1476580>
- Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling*, 14(3), 464–504. <https://doi.org/10.1080/10705510701301834>
- Chen, W. H., Yao, J., & Thissen, D. (2013). A general method for assessing the fit of item response models. *Psychometrika*, 78(1), 26–55.
- Christensen, K. B., Makransky, G., & Horton, M. (2017). Critical values for Yen's Q3: Identification of local dependence in the Rasch model using residual correlations. *Applied Psychological Measurement*, 41(3), 178–194. <https://doi.org/10.1177/0146621616677520>
- Davey, G. C. I. (2018). *The anxiety epidemic*. Robinson.
- Davey, G. C., Meeten, F., & Field, A. P. (2022). What's worrying our students? Increasing worry levels over two decades and a new measure of student worry frequency and domains. *Cognitive Therapy and Research*, 1–14. <https://doi.org/10.1007/s10608-021-0270-0>
- Ehring, T., & Watkins, E. R. (2008). Repetitive negative thinking as a transdiagnostic process. *International Journal of Cognitive Therapy*, 1(3), 192–205. <https://doi.org/10.1521/ijct.2008.1.3.192>
- Embretson, S. E. (1999). Generating items during testing: Psychometric issues and models. *Psychometrika*, 64(4), 407–433. <https://doi.org/10.1007/BF02294564>
- Embretson, S. E., & Reise, S. P. (2000). *Item response theory for psychologists*. Lawrence Erlbaum Associates.
- Freeston, M., Tiplady, A., Mawn, L., Bottesi, G., & Thwaites, S. (2020). Towards a model of uncertainty distress in the context of Coronavirus (COVID-19). *Cognitive Behaviour Therapy*, 13, e31. <https://doi.org/10.1017/S1754470X2000029X>
- Fresco, D. M., Heimberg, R. G., Mennin, D. S., & Turk, C. L. (2002). Confirmatory factor analysis of the Penn State worry questionnaire. *Behavior Research and Therapy*, 40(3), 313–323. [https://doi.org/10.1016/S0005-7967\(00\)00113-3](https://doi.org/10.1016/S0005-7967(00)00113-3)
- Gould, C. E., & Edelstein, B. A. (2010). Worry, emotion control, and anxiety control in older and young adults. *Journal of Anxiety Disorders*, 24, 759–766. <https://doi.org/10.1016/j.janxdis.2010.05.009>
- Hassen, S., & Adane, M. (2021). Facemask-wearing behavior to prevent COVID-19 and associated factors among public and private bank workers in Ethiopia. *PLoS One*, 16(12), Article e0259659. <https://doi.org/10.1371/journal.pone.0259659>
- Hazlett-Stevens, H., Ullman, J. B., & Craske, M. G. (2004). Factor structure of the Penn State worry questionnaire: Examination of a method factor. *Assessment*, 11(4), 361–370. <https://doi.org/10.1177/1073191104269872>
- Henson, R., & Douglas, J. (2005). Test construction for cognitive diagnosis. *Applied Psychological Measurement*, 29(4), 262–277. <https://doi.org/10.1177/0146621605275722>
- Heung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling*, 9(2), 233–255. [https://doi.org/10.1207/S15328007SEM0902\\_5](https://doi.org/10.1207/S15328007SEM0902_5)
- Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6(1), 1–55. <https://doi.org/10.1080/10705519909540118>
- Jeon, J. W., Kim, D., Kim, E., & Roh, S. (2017). Exploratory and confirmatory factor analysis of the Korean version of the Penn State worry questionnaire. *Anxiety and Mood*, 13(2), 86–92. <https://www.koreascience.or.kr/article/JAKO201734158359254.page>
- Kline, R. B. (2015). *Principles and practice of structural equation modeling* (4th ed.). Guilford Press.
- Lim, Y. J., Kim, Y. H., Lee, E. H., & Kwon, S. M. (2008). The Penn State worry Questionnaire: Psychometric properties of the Korean version. *Depression and Anxiety*, 25(10), E97–E103. <https://doi.org/10.1002/da.20356>
- Limone, P., Toto, G. A., & Messina, G. (2022). Impact of the COVID-19 pandemic and the Russia-Ukraine war on stress and anxiety in students: A systematic review. *Frontiers in Psychiatry*, 13, Article 1081013. <https://doi.org/10.3389/fpsyt.2022.1081013>
- Little, T. D. (2013). *Longitudinal structural equation modeling*. Guilford Press.
- Liu, K., Nijme, J. S., & Warren, S. L. (2022). Factor structure, measurement invariance, and concurrent validity of the Penn State Worry Questionnaire across development, psychopathology, and culture. *Assessment*, 29(5), 909–924. <https://doi.org/10.1177/1073191121993223>
- Lord, F. M. (1980). *Applications of item response theory to practical testing problems*. L. Erlbaum Associates.
- Lord, F. M., & Novick, M. R. (1968). *Statistical theories of mental test scores*. Menlo Park: Addison-Wesley.
- Lovibond, S. H., & Lovibond, P. F. (1995). *Manual for the depression anxiety stress scales* (2nd ed.). Psychology Foundation.
- Masters, G. N., & Wright, B. D. (1984). The essential process in a family of measurement models. *Psychometrika*, 49, 529–544. <https://doi.org/10.1007/BF02302590>
- McEvoy, P. M., Watson, H., Watkins, E. R., & Nathan, P. (2013). The relationship between worry, rumination, and comorbidity: Evidence for repetitive negative thinking as a transdiagnostic construct. *Journal of Affective Disorders*, 151(1), 313–320. <https://doi.org/10.1016/j.jad.2013.06.014>
- Meloni, F., & Gana, K. (2001). Wording effects in the Italian version of the Penn State Worry questionnaire. *Clinical Psychology and Psychotherapy*, 8(4), 282–287. <https://doi.org/10.1002/cpp.294>
- Meyer, T. J., Miller, M. L., Metzger, R. L., & Borkovec, T. D. (1990). Development and validation of the Penn State worry questionnaire. *Behavior Research and Therapy*, 28, 487–495. [https://doi.org/10.1016/0005-7967\(90\)90135-6](https://doi.org/10.1016/0005-7967(90)90135-6)
- Morani, S., Pricci, D., & Sanavio, E. (1999). Penn State worry questionnaire e worry domains questionnaire. Presentazione delle versioni italiane ed analisi della fedeltà. *Psicoterapia Cognitiva e Comportamentale*, 5(3), 13–34.
- Novick, M. R. (1966). The axioms and principal results of classical test theory. *Mathematical Psychology*, 3(1), 1–18. [https://doi.org/10.1016/0022-2496\(66\)90002-2](https://doi.org/10.1016/0022-2496(66)90002-2)
- Padros-Blazquez, F., Gonzalez-Betanzos, F., Martinez-Medina, M. P., & Wagner, F. (2018). Psychometric characteristics of the original and brief version of the Penn State Worry Questionnaire (PSWQ) in Mexican samples. *Actas Españolas de Psiquiatria*, 46(4), 117–124.
- Porta-Casteràs, D., Fullana, M. A., Tinoco, D., Martínez-Zalacáin, I., Pujol, J., Palao, D. J., & Cardoner, N. (2020). Prefrontal-amygdala connectivity in trait anxiety and generalized anxiety disorder: Testing the boundaries between healthy and pathological worries. *Journal of Affective Disorders*, 267, 211–219. <https://doi.org/10.1016/j.jad.2020.02.029>
- Ruiz, F. J., Monroy-Cifuentes, A., & Suárez-Falcón, J. C. (2018). Penn State worry questionnaire-11 validity in Colombia and factorial equivalence across gender and nonclinical and clinical samples. *Anales de Psicología*, 34(3), 451–457. <https://doi.org/10.6018/analesps.34.3.300281>
- Schroder, H. S., Clark, D. A., & Moser, J. S. (2019). Screening for problematic worry in adults with a single item from the Penn State worry questionnaire. *Assessment*, 26(2), 336–346. <https://doi.org/10.1177/1073191117694453>
- Schwartz, S. H., & Melech, G. (2000). National differences in micro and macro worry: Social, economic, and cultural explanations. In E. Diener, & E. M. Suh (Eds.), *Culture and subjective well-being* (pp. 219–256). Cambridge, MA: The MIT Press.
- Taylor, S. (2020). Anxiety disorders, climate change, and the challenges ahead: Introduction to the special issue. *Journal of Anxiety Disorders*, 76, Article 102313. <https://doi.org/10.1016/j.janxdis.2020.102313>
- Warren, S. L., Heller, W., & Miller, G. A. (2021). The structure of executive dysfunction in depression and anxiety. *Journal of Affective Disorders*, 279, 208–216. <https://doi.org/10.1016/j.jad.2020.09.132>
- Wu, S. M., Schuler, T. A., Edwards, M. C., Yang, H. C., & Brothers, B. M. (2013). Factor analytic and item response theory evaluation of the Penn State worry Questionnaire in women with cancer. *Quality of Life Research*, 22(6), 1441–1449. <https://doi.org/10.1007/s11136-012-0253-0>
- Zhong, J., Wang, C., Li, J., & Liu, J. (2009). Penn State worry questionnaire: Structure and psychometric properties of the Chinese version. *Journal of Zhejiang University Science B*, 10(3), Article 211. <https://doi.org/10.1631/jzus.B0820189>. Article.