Article:
Allen, Sarah, van Hemert, Albert M, de Waal, Margot W M et al. (1 more author) (2025) Exploration of the factor structure of the Physical Symptoms Questionnaire (PSQ-51). Comparisons between Dutch and English samples. General hospital psychiatry. pp. 140-147. ISSN 0163-8343

https://doi.org/10.1016/j.genhosppsych.2025.05.002

General Hospital Psychiatry 95 (2025) 140–147

Contents lists available at ScienceDirect

General Hospital Psychiatry

journal homepage: www.elsevier.com/locate/genhospsych

Exploration of the factor structure of the physical symptoms questionnaire (PSQ-51). Comparisons between Dutch and English samples

Sarah F. Allen [a,b,*], Albert M. van Hemert [c], Margot W.M. de Waal [d], Christina van der Feltz-Cornelis [b,e]

[a] Faculty of Health & Life Sciences, Northumbria University, Newcastle upon Tyne, UK
[b] Department of Health Sciences, University of York, UK
[c] Department of Psychiatry, Leiden University Medical Centre, the Netherlands
[d] Department of Public Health and Primary Care, Leiden University Medical Centre, the Netherlands
[e] Institute of Health Informatics, UCL, London, UK

ARTICLE INFO

Keywords:
Translation
Psychometric analysis
Reliability
Factor analysis
PSQ-51
Physical symptoms

ABSTRACT

The Physical Symptoms Questionnaire (PSQ-51) is a Dutch-language self-report tool listing 51 physical symptoms that may occur in patients with known medical conditions, somatic symptom disorders and the general population. However, the tool is currently only available in Dutch and is yet to be translated or validated into English for utility with English-speaking populations. This study aimed to translate and validate an English version of the tool and determine the PSQ-51's factor structure in both Dutch and UK samples. An English version was translated and then validated through back-translation and refined for clarity. Data from three Dutch samples (general population [$n = 1699$], general practice[$n = 775$], and psychiatric outpatients[$n = 1404$]) and one UK general population sample ($n = 294$) were then analysed to explore the factor structure. An iterative exploratory factor analysis (EFA) on the Dutch psychiatric sample revealed a seven-factor solution including symptom clusters: General Malaise, Autonomic, Musculoskeletal, Gastrointestinal, Loss of Function, Hot Flushes, and Urogenital symptoms. Confirmatory factor analysis (CFA) tested this model across the Dutch and UK samples, with mixed results for fit indices, although good internal reliability was demonstrated. Findings indicate partial cross-cultural consistency in the factor structure. Substantial differences in symptom prevalence between Dutch and UK population samples were observed, possibly due to cultural and situational factors. The PSQ-51 shows promise for assessing somatic symptom burden, for example in multimorbidity or in complex somatic symptom disorders, where it may enhance clinical consultations by identifying symptoms to address clinical complexity. Further research is needed to explore its applicability in diverse populations and refine its factor structure for broader clinical utility.

1. Introduction

Physical symptoms are bodily sensations or mental experiences that are perceived as indicative of a change in health, due to a change in bodily function, illness, injury or disease. [1] Often symptoms can have no obvious underlying cause, and in the absence of appropriate medical treatment, these symptoms can be equally as distressing as those occurring in the context of a clinically diagnosed medical condition, particularly when they persist. [2]

For some physical symptoms such as pain, dizziness and fatigue, fully objective markers are impossible to obtain and proxy measures can be unreliable. [3] Therefore, the use of self-report measures within both diagnostic processes and outcome assessments, is highly warranted. The Physical Symptoms Questionnaire (LKV in Dutch) [4] is a scale which was developed for and implemented in clinical settings in the Netherlands. The Physical Symptoms Questionnaire (PSQ-51) is a Dutch-language self-report measure which lists 51 physical symptoms that may occur in known medical conditions and somatic symptom disorders and can also easily be used to gauge the experience of physical symptoms in the general population. The PSQ-51 is currently only available in Dutch and is yet to be translated and validated into English for use in English-speaking populations. Further, although it was put

* Corresponding author at: Department of Psychology, Faculty of Health and Life Sciences, Northumbria University, Newcastle upon Tyne, UK.
E-mail address: s.f.allen@northumbria.ac.uk (S.F. Allen).

https://doi.org/10.1016/j.genhosppsych.2025.05.002
Received 26 February 2025; Received in revised form 2 May 2025; Accepted 2 May 2025
Available online 10 May 2025

together as listing symptoms relating to a variety of organ systems [4], whether its factor structure is reflective of these systems, or comprises one general factor indicating physical symptom burden in a patient is yet to be determined.

While several English language standardized and validated instruments exist to measure somatic symptoms [5] these vary considerably by length, scaling, dimensionality, reliability, validity, and population relevance. Among 40 self-report symptom scales identified in a recent systematic review, the Patient Health Questionnaire–15 (PHQ-15) [6] and the 12-item Symptom Checklist–90 somatization scale (SCL-90) [7] were identified as the most appropriate measures for large scale studies. [5] Both scales are brief, available in multiple languages, have well-established psychometric properties and include relevant symptoms. [5] However, in comparison to the PSQ-51 these measures are limited by their brevity and are unable to provide a complete list of symptoms that may be clinically relevant for a patient. While they attempt to identify physical symptoms that occur in the context of somatization, their validity to do so has so far been limited. [8] The 30-item Bodily Distress Scale [9] is a more comprehensive measure particularly useful in primary care for diagnosing functional disorders. However, the PSQ-51 due to its broader range of symptoms is more suitable for in-depth symptom exploration in research contexts rather than as a diagnostic tool.

While some physical symptoms may occur in isolation, often they cluster together, particularly in the context of certain illnesses, or concerning different body parts or organ systems. Previous studies employing factor analysis methods have yet to reach a consensus on the structure of symptom clusters. For example, the Subjective Health Complaints Inventory (SHC) [10] comprises musculoskeletal pain; allergies; gastrointestinal problems; pseudoneurology and flu; and the Cohen-Hobermann Inventory of Physical Symptoms (CHIPS) [11] has an eight-factor solution. [12] The factor structure of the PSQ-51 has not yet been explored and could provide important information regarding a patient's symptomatic experience of their illness in addition to the information regarding symptom burden currently provided by the scale. Given how the list of symptoms was formulated it is hypothesised that the factor structure will reflect the organ systems to which the symptoms relate.

1.1. Aims and objectives

The primary aim of the current study is to examine the factor structure of the PSQ-51 tool. A secondary aim is to translate and validate an English version of the tool. The factor structure will initially be explored in Dutch samples, and then tested in a UK sample utilising the newly translated English language version. The current study aims to address the following objectives:

1. Translate and examine the psychometric properties of an English-language version of the PSQ-51
2. Determine the factor structure of the original Dutch language PSQ-51 tool and the fit of the model across Dutch general practice, psychiatric and general population samples.
3. Examine if the English-language version of the PSQ-51 shares the same factor structure as the original Dutch version.

2. Methods

The PSQ-51 lists 51 commonly occurring somatic non-gender specific symptoms (see Table 1 for English translated items, as described below). The list was based on 47 somatic symptoms that are mentioned in the DSM-III as criteria for a psychiatric classification, mostly for somatoform disorders and some for anxiety and depressive disorders [13]. Additionally, 8 common symptoms were added from a registration study in internal medicine outpatients (being easily fatigued after little exertion, tense muscles, muscle aches or stiffness, dry mouth, heat intolerance, intolerance for cold, headache and frequent urinating). There are 11 general/ neurological items, 10 autonomic items, 8 musculoskeletal/pain items, 13 gastrointestinal items, 5 urological/ genital items and 4 items about feeling hot/cold [14]. Four gender-specific items were not included as per previous studies [15] to rule out bias by gender. The tool asks respondents how they have been bothered by these symptoms during the past week. Response choices are on a Likert scale from 0 to 3: 0 = "I was not bothered by this", 1 = "I was sometimes bothered by this", 2 = "I was regularly bothered by this", and 3 = "I was often bothered by this". Since the instrument was originally conceived as a checklist, a symptom is rated as absent (0) for response

Table 1
Prevalence of the 51 physical symptoms assessed by the Dutch and English PSQ-51 (%).

Symptom	NL Pop (n = 1699)	NL GP (n = 775)	NL Psy (n = 1404)	UK Pop (n = 294)
1. General fatigue or listlessness	21.1	28.6	74.2	76.4
2. Being easily fatigued after little exertion	13.0	21.2	61.8	62.3
3. Shortness of breath without exercise	4.6	6.4	17.9	24.5
4. Palpitations	2.9	4.9	17.5	31.1
5. Chest pain or discomfort	3.0	4.5	13.6	25.5
6. Feeling dizzy	6.1	8.4	26.4	37.7
7. Fainting	0.1	0.3	1.4	4.7
8. Difficulty falling or staying asleep	12.2	15.6	43.9	67.0
9. Sleepiness	14.8	17.1	33.0	82.1
10. Forgetfulness	8.3	9.7	41.5	66
11. Numbness or Tingling sensations	6.3	10.3	19.3	26.4
12. Trembling or shaking	2.9	4.3	20.9	23.6
13. Muscle weakness or paralysis	1.9	2.8	11.4	10.4
14. Tense muscles	17.1	21.1	44.7	35.8
15. Muscle aches or stiffness	16.4	22.7	39.5	61.3
16. Difficulty walking	6.2	10.8	16.6	14.3
17. Loss of voice	0.8	1.8	4.5	6.6
18. Deafness	4.0	5.8	5.2	4.7
19. Double vision. Blurred vision	2.4	2.5	12.6	16.0
20. Blindness	0.0	0.1	0.8	1.4
21. Seizures	0.0	0.1	0.7	1.7
22. Nausea	2.2	5.0	16.5	27.4
23. Vomiting	0.4	1.5	3.0	2.9
24. Dry mouth	4.5	8.3	26.1	27.4
25. Difficulty swallowing	1.6	2.6	7.8	10.4
26. Choking frequently	1.2	2.1	4.6	2.8
27. Poor tolerance of certain foods	4.1	7.2	11.4	16.0
28. Lack of appetite (last month)	3.1	4.9	24.7	22.9
29. Weight loss	2.2	3.1	16.2	4.7
30. Heartburn	6.2	8.5	15.0	26.7
31. Stomach pain	5.3	9.5	19.2	31.1
32. Abdominal distress	7.5	11.6	24.7	43.8
33. Diarrhoea	3.8	6.0	12.5	20.8
34. Constipation	2.7	4.6	12.6	30.2
35. Flatulence	12.4	15.7	21.9	38.7
36. Sweatiness	9.8	15.5	29.5	29.2
37. Hot flushes	8.8	13.9	25.9	34.0
38. Heat intolerance	7.4	10.9	25.0	22.6
39. Cold chills	3.0	4.6	16.0	22.6
40. Intolerance for cold	8.3	10.7	24.0	33.3
41. Headache	14.4	18.2	39.0	51.9
42. Pain in the joints	14.9	21.2	28.5	42.5
43. Pain in arms or legs	13.9	20.5	28.1	28.6
44. Back pain	19.3	25.9	38.1	49.1
45. Other pain	8.5	15.1	23.8	18.9
46. Frequent urinating	10.8	16.1	27.2	26.4
47. Difficulty urinating	1.2	2.0	4.1	4.7
48. Pain urinating	0.5	1.1	1.7	5.7
49. Burning feeling genitals or anus	1.8	3.4	4.1	6.6
50. Painful intercourse	2.2	2.4	3.5	5.7
51. Lack of sexual interest	8.9	13.9	23.1	34.9

choices 0 or 1 and present (1) for 2 and 3. Total scores range between 0 and 51. Cronbach's alpha has previously been reported as 0.88 [15].

2.1. English translation

The Dutch scale [4] was deemed suitable for translation as the scale uses simple language, and the topic (i.e. self-reported physical symptoms) is an easily understandable and transferable construct which avoids cultural bias. [16] Therefore, the scale was translated using the recommended forward-backward translation with two bilingual translators. [16] Firstly, the 51 items were independently translated from Dutch into English by a bilingual translator. The item translations were then discussed between the bilingual translator and two native English speakers to ensure the translation was easily understood. The translated version was then back-translated into Dutch by a second independent translator. The translated version was then compared to the original Dutch version. Most items were translated identically or very closely to the original version when back-translated. These items were automatically included in the English version as it was deemed the translated version was likely equivalent in meaning. Both interpreters were consulted to discuss minor discrepancies, and a consensus was met in all cases. The translated questionnaire was examined by the study team to ensure accurate translation and interpretability with no major suggestions. Minor remarks were given to improve the clarity of the wording, e.g., the word 'insomnia' was changed to 'difficulty falling or staying asleep' and 'sexual disinterest' was changed to 'lack of sexual interest' as this was easier to understand. In addition, a few language errors were detected and corrected. Discussions between native English speakers and native Dutch speakers ensured the meaning of each item was consistent.

2.2. Data samples

The factor structure of the original Dutch-language PSQ-51 was assessed with data obtained from two distinct sources: the Somatisation Study at the University of Leiden (SOUL study) [17] and the Leiden Routine Outcome Monitoring Study (Leiden ROM study) [18,19]. The former contributed data from both the general population and patients seeking consultation with their general practitioner, while the latter contributed data for patients who were referred to secondary mental health care for mood, anxiety, or somatoform (MAS) disorders.

The SOUL study is a two-stage prevalence and intervention study on somatoform disorders in general practice. Dutch native patients aged 25–69 years were recruited from patients registered with 16 family physicians in urban areas in the Western part of The Netherlands. From April 2000 until September 2002 a total of 1830 consulting and 4579 registered patients were screened with the PSQ-51 as one of the main instruments. Given the nearly 100 % coverage of general practices in The Netherlands, the group of registered patients is considered a representative sample of the general population. Complete PSQ-data were available for 775 consulting (65.4 % female, mean age = 45.9 SD10.9) and 1698 registered (58.2 % female, mean age = 43.2 SD11.0) patients. All patients provided written informed consent and the study was approved by the Medical Ethics Committee of Leiden University Medical Centre.

The Leiden ROM study is a natural cohort study among adult patients who were referred for treatment of a mood-, anxiety-, or somatoform (MAS) disorder to the Dutch Regional Mental Health Provider (RMHP) Rivierduinen (RD) or the psychiatric outpatient department of the Leiden University Medical Centre (LUMC) between January 2004 and December 2006. As part of the intake procedure, patients were assessed with a standard Routine Outcome Monitoring (ROM) consisting of a battery of psychometric instruments. The PSQ was included in the ROM procedure between March 2005 and July 2006. Patients with insufficient proficiency in the Dutch language or those unable to complete computerized and written questionnaires were deemed ineligible for ROM. On average, 80 % of the referred patients were assessed with ROM. Complete PSQ data were available for 1404 patients (64.2 % female, mean age = 38.3 SD12.7). The ROM procedure was approved by the Medical Ethics Committee of Leiden University Medical Centre and informed consent was not required.

The final English version of the PSQ-51 was tested for validity and reliability in a sample of the general population through an online cross-sectional questionnaire-based study conducted via the survey platform Qualtrics. Participants were recruited between December 2019 and April 2021 through social media advertisements [20] inviting individuals to take part in a study evaluating a questionnaire designed to assess physical symptom burden. As part of the study, participants also completed the SS8, PHQ-9, and GAD-7 to support the assessment of validity, however, these data are not the focus of the current paper.A total of 356 participants started the survey (clicked the URL), however 62 participants did not complete the survey and data were removed in line with the ethically approved withdrawal procedure. The final sample included 294 adults (75.9 % female, mean age = 40.33 SD = 13.71, 88.4 % white). The study protocol was approved by the University of York, Department of Health Sciences Ethics and Governance committee. Informed consent was obtained from all participants at the beginning of the study.

2.3. Statistical analyses

The Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy was used to assess the suitability of the data for factor analysis. Values closer to 1 indicate greater suitability with values above 0.60 generally considered acceptable. An iterative exploratory factor analysis was conducted using principal component factor analysis with oblimin oblique normalised rotation on the Dutch ROM sample. This was followed by an iterative exploration of factor loadings before selecting a final balanced item set. The ROM sample was utilized for the exploratory factor analysis, because the average ratings on the PSQ-51 were significantly higher than in the other Dutch data samples, enabling the inclusion of less common items, such as fainting, blindness, and seizures, which were endorsed by at least 10 individuals. Items were entered with the 0–3 Likert scale. The exploratory analyses were conducted with Stata 17.0 [21].

Confirmatory factor analysis was subsequently implemented via structural equation modelling to assess model fit in the Dutch general practice and general population samples, followed by the UK general population sample. Confirmatory factor analyses were conducted in R version 4.0.5 - Lavaan 0.6–9 both in the Dutch and the English samples to ensure reproducibility. Modification indices were calculated to identify pairs of items, where co-variances could be added to improve the fit of the model. For within cluster pairs with modification indices larger than 40 co-variances were added to the model.

To test the goodness of fit we examined the following indices: Chi-squared test of significance, Root Mean Square Error of Approximation (RMSEA), Standardized Root Mean Square Residual (SRMR), The Comparative Fit Index (CFI), and Tucker-Lewis Index. The Chi-Square value is the traditional measure of assessing overall model fit and a non-significant result is thought to indicate good model fit [22]. Values for the RMSEA and SRMR range from 0.0 to 1.0 with smaller values indicating better fit (cut-offs indicate good fit values of less than 0.06/7 for RMSEA and < 0.05 for SRMR; [23,24] It must, however, be noted that SRMR will be lower for larger sample sizes and models with a higher number of parameters. Values for CFI and TLI also range between 0.0 and 1.0 but with values closer to 1.0 indicating better fit (Cut-offs of CFI >0.95 and TFI >0.8 are recommended for determining a good fitting model; [23,25].

The internal consistency of each of the factors and the total scale were calculated using Cronbach's alpha in each of the four samples.

3. Results

Response rates were 42 % for the Dutch general population sample [17] and 59 % for the Dutch consulting population sample, [26] with an underrepresentation of males in both samples. For the Leiden ROM data [19]the response rate was 80 % on average, as the questionnaires were part of routine care. For the UK-sample a response rate could not be calculated as it was a self-selected sample. In the Dutch general population (SOUL-registered) subjects endorsed 3.4 (SD 4.6) symptoms on average, in the general practice consulting population (SOUL-consulting), these were 4.9 (SD 5.5) symptoms and in the psychiatric sample (Leiden ROM) 10.6 (SD 7.8) symptoms. In the English general population sample, the average number of symptoms experienced was 13.9 (SD 8.5) The percentage of the sample that endorsed a particular symptom is presented in Table 1.

Table 2
Factor loadings on the final balanced item set of seven factors (factor loadings <0.20 were suppressed) in the Dutch psychiatric (ROM)sample.

Symptom	Factor 1: General malaise 1 & 2		Factor 2: Autonomic	Factor 3: Musculoskeletal	Factor 4: gastrointestinal	Factor 5: Loss of function	Factor 6: Hot Flushes	Factor 7: Urogenital
1. General fatigue or listlessness	0.83	–		–	–	–	–	–
2. Being easily fatigued	0.71	–		–	–	–	–	–
3. Shortness of breath without exercise	–	–	0.48	–	–	–	–	–
4. Palpitations	–	–	0.53	–	–	–	–	–
5. Chest pain or discomfort		–	0.51	–	–	–	–	–
6. Feeling dizzy	–	–	0.45	–	–	–	–	–
7. Fainting	–	–	–	–	–	0.30	–	–
8. Difficulty falling or staying asleep	–	0.38	–	–	–	–	–	–
9. Sleepiness	0.34	–	–	–	–	–	–	–
10. Forgetfulness	0.44	–	–	–	–	–	–	–
11. Numbness or Tingling sensations	–	–	0.45	–	–	–	–	–
12. Trembling or shaking	–	–	0.47	–	–	–	–	–
13. Muscle weakness or paralysis	–	–	–	0.35	–	–	–	–
14. Tense muscles	–	–	–	0.42	–	–	–	–
15. Muscle aches or stiffness	–	–	–	0.70	–	–	–	–
16. Difficulty walking	–	–	–	0.53	–	–	–	–
17. Loss of voice	–	–	–	–	–	0.34	–	–
18. Deafness	–	–	–	–	–	0.25	–	–
19. Double vision, blurred vision	–	–	0.30	–	–	–	–	–
20. Blindness	–	–	–	–	–	0.38	–	–
21. Seizures	–	–	–	–	–	0.34	–	–
22. Nausea	–	–	–	–	0.42	–	–	–
23. Vomiting	–	–	–	–	0.25*	–	–	–
24. Dry mouth	–	–	–	–	–	–	0.25*	–
25. Difficulty swallowing	–	–	0.29*	–	–	–	–	–
26. Choking frequently	–	–	–	–	–	0.32	–	–
27. Poor tolerance of certain foods	–	–	–	–	0.43	–	–	–
28. Lack of appetite (last month)	–	0.73	–	–	–	–	–	–
29. Weight loss	–	0.62	–	–	–	–	–	–
30. Heartburn	–	–	–	–	0.34	–	–	–
31. Stomach pain	–	–	–	–	0.67	–	–	–
32. Abdominal distress	–	–	–	–	0.66	–	–	–
33. Diarrhoea	–	–	–	–	0.34	–	–	–
34. Constipation	–	–	–	–	0.32	–	–	–
35. Flatulence	–	–	–	–	0.40	–	–	–
36. Sweatiness	–	–	–	–	–	–	0.88	–
37. Hot flushes	–	–	–	–	–	–	0.92	–
38. Heat intolerance	–	–	–	–	–	–	0.57	–
39. Cold Chills	–	–	0.27*	–	–			
40. Intolerance for cold	–	–	–	0.20*	–	–	–	–
41. Headache	–	–	–	0.27*	–	–	–	–
42. Pain in the joints	–	–	–	0.86	–	–	–	–
43. Pain in arms or legs	–	–	–	0.86	–	–	–	–
44. Back pain	–	–	–	0.66	–	–	–	–
45. Other pain	–	–	–	0.58	–	–	–	–
46. Frequent urinating	–	–	–	–	–	–	0.24*	–
47. Difficulty urinating	–	–	–	–	–		–	0.53
48. Pain urinating	–	–	–	–	–		–	0.64
49. Burning feeling genitals or anus	–	–	–	–	–		–	0.46
50. Painful intercourse	–	–	–	–	–		–	0.33
51. Lack of sexual interest	0.20*	–	–	–	–	–	–	–

[a] indicates items which were not included in the final model.

3.1. Exploratory factor analysis

The Kaiser-Meyer-Olkin measure verified the sampling adequacy for the analysis (KMO = 0.929 (marvellous according to Hutcheson and Sofroniou, 1999) and Bartlett's test of Sphericity was a significant result ($p < .001$), therefore confirming the suitability of the data for PCA.

Kaiser's criterion [27] which states factors with eigenvalues above 1.00 should be extracted resulted in a one-factor solution, which could represent a general tendency to experience bodily discomforts. As a one factor solution was deemed uninformative an iterative approach was followed to examine how items moved across the factors for different solutions ranging from a six-factor structure up to a twelve-factor structure (see supplementary material). A seven-factor structure was extracted based on i) the consistency of item loadings across factors throughout the iterative process, ii) an adequate number of items per factor (i.e., three or more) and iii) the practical meaningfulness of the factors in terms of the included items (i.e. similarity of symptoms). As such, two factors were amalgamated to form the 'General Malaise' factor based on conceptual item similarity, and one factor was split into two: 'Loss of Function' and 'Urogenital' due to the distinctiveness of the items. Subsequently, a final balanced item set of seven factors was selected.

There was limited evidence of overlap between factors, however these items were included on the factor for which they demonstrated the highest loading if this made conceptual sense. Factor loadings for each item on the factors is shown in Table 2.

Eight items were not included in the final factor structure as they either i) lacked stability across solutions, ii) did not demonstrate loadings >0.25 on any factors, or iii) did not make conceptual sense in terms of organ systems ('vomiting'; 'dry mouth'; 'difficulty swallowing' 'cold chills', 'intolerance to cold, 'headache'; 'frequent urination' and 'lack of sexual interest;'). The final seven factors are identified and labelled as shown in Table 3.

3.2. Confirmatory factor analysis

Confirmatory factor analysis was implemented via structural equation modelling, to test how well the seven-factor structure model fit in each sample. The results of the CFA are demonstrated in Table 4. Model 1 indicates the original 7-factor structure as determined by the EFA. Model 2 indicates a revised model in which, based on the examination of modification indices, within cluster inter-item co-variances were added to the model.

3.3. Reliability

Cronbach's alpha was shown to be excellent (>0.90) for the total scale in all four samples and are shown in Table 5.

4. Discussion

The primary aim of the current study was to explore and validate the factor structure of the PSQ-51. Firstly, the factor structure of the original Dutch language tool was determined using iterative exploratory factor analysis. The fit of the model was then tested across Dutch psychiatric, general practice and general population samples. The secondary aim of the study was to translate an English-language version of the PSQ-51 and examine if this version shares the same factor structure as the original Dutch version.

The EFA was conducted in the Dutch psychiatric sample. An iterative EFA process was implemented to assess how symptoms moved around on subscales. The Kaiser's criterion suggested a one-factor solution, which could represent a general tendency to experience bodily discomforts. As this was not very informative, we further explored potential factors and settled on a 7-factor structure, which best fit the data from a conceptual point of view, although the loading of items throughout the iterative process was not static. This brings into question whether the experience of physical symptoms is indeed dimensional in a consistent way. Nevertheless, seven factors were proposed as follows; 'General malaise' contained 7 items associated with fatigue, issues with vitality (e.g., 'loss of appetite') and cognition (e.g., forgetfulness), this was an amalgamation of two separate but conceptually related factors. 'Autonomic symptoms' contained 7 items associated with cardiac function and the autonomic system (e.g., 'feeling dizzy', 'palpitations'). The 'musculoskeletal' factor comprised 7 items associated with muscle pain and weakness. The 'gastrointestinal' factor contained 7 items and 6 items related to 'loss of function' (e.g., 'loss of voice', 'deafness'). The final factors were 'hot flushes' and 'urogenital' related symptoms.

It was hypothesised that the factor structure would reflect the organ systems to which the symptoms relate, and this was partially supported by the emergence of the musculoskeletal, gastrointestinal, urogenital, and autonomic factors. The loss of function and general malaise factors included a more generic selection of items, not necessarily aligned with a singular biological system, however this was also expected given the unspecified nature of symptoms such as fatigue. The factor structure proposed here is similar to, the factor structure of the 33-item Cohen-Hoberman inventory of physical symptoms (CHIPS; [11,12] and the subjective health complaints inventory [10] as both include factors relating to 'musculoskeletal' pain and 'gastrointestinal' symptoms. The CHIPS structure also includes 'sympathetic symptoms' which contain similar items to our 'autonomic' factor (e.g., chest pain), and 'metabolic symptoms' similar to the 'general malaise' factor (e.g., loss of appetite).

Our findings also partially align with previous research implementing factor analysis on physical symptom checklists. For example, Kroenke [28] identified cardiopulmonary, gastrointestinal, and general pain/fatigue clusters, while cardiopulmonary, musculoskeletal, and gastrointestinal clusters were found by Fink et al. [29] across clinical settings. Further, Gara and colleagues [30]studying primary care patients reporting medically unexplained symptoms, grouped 41 symptoms into eight organ-based clusters: three of which; gastrointestinal,

Table 3
Items included on each symptom factor in the final model.

General malaise	Autonomic	Musculo-skeletal	Gastro-intestinal	Loss of function	Hot Flushes	Urogenital	Excluded
• General fatigue or listlessness • Being easily fatigued • Difficulty falling or staying asleep • Sleepiness • Forgetfulness • Lack of appetite • Weight loss	• Shortness of breath without exercise • Palpitations • Chest pain or discomfort • Feeling dizzy • Numbness or Tingling sensations • Trembling or shaking • Double/blurred vision	• Pain in the joints • Pain in arms or legs • Back pain • Other pain • Tense muscles • Muscle aches or stiffness • Difficulty walking	• Heartburn • Stomach pain • Abdominal distress • Diarrhoea • Constipation • Flatulence • Poor tolerance of certain foods	• Fainting • Loss of voice • Deafness • Blindness • Seizures • Choking frequently	• Sweatiness • Hot flushes • Heat intolerance	• Difficulty urinating • Pain urinating • Burning feeling genitals or anus • Painful intercourse	• Vomiting • Dry mouth • Difficulty swallowing • Cold chills • Intolerance for cold • Headache • Frequent urinating • Lack of sexual interest

Table 4
Goodness of fit indices for each model across the four samples.

Model	Fit indices	NL Psy (n = 1404)	NL GP (n = 775)	NL Pop (n = 1699)	UK Pop (n = 294)
Model 1 (Original)	X^2 *p* value	<0.001	<0.001	0.996	<0.001
	RMSEA	0.055	0.057	0.048	0.068
	[90 % CI]	[0.053,0.057]	[0.055,0.059]	[0.046,0.049]	[0.064,0.072]
	SRMR	0.055	0.056	0.044	0.087
	CFI	0.835	0.792	0.843	0.800
	TLI	0.822	0.776	0.831	0.783
Model 2 (Revised)*	X^2 p value	1.0	0.98	1.0	<0.001
	RMSEA	0.042	0.047	0.039	0.065
	[90 % CI]	[0.040,0.043]	[0.044,0.049]	[0.038,0.041]	[0.061,0.069]
	SRMR	0.043	0.050	0.038	0.080
	CFI	0.904	0.861	0.894	0.818
	TLI	0.895	0.848	0.884	0.800

* Revised for modification indices in which the model is estimated if the following between item co-variances were added: 28 ~ ~29; 14 ~ ~15; 08 ~ ~09; 04 ~ ~05; 31 ~ ~32; 42 ~ ~43; 01 ~ ~02; 01 ~ ~09; 32 ~ ~35; 27 ~ ~30.

Table 5
Reliability coefficients (Cronbach's alpha) for each factor across the four samples.

Factors	Number of Items	``NL Psy (n = 1404)	NL GP (n = 775)	NL Pop (n = 1699)	UK Pop (n = 294)
General malaise	7	0.66	0.73	0.73	0.82
Autonomic	7	0.80	0.70	0.69	0.82
Musculo-skeletal	7	0.87	0.83	0.85	0.83
Gastrointestinal	7	0.79	0.80	0.77	0.77
Loss of function	7	0.67	0.49	0.51	0.84
Hot flushes	3	0.80	0.83	0.80	0.68
Urogenital	4	0.66	0.52	0.50	0.65
Total	42	0.92	0.91	0.91	0.93

musculoskeletal, genitourinary closely align with our musculoskeletal, gastrointestinal, and urogenital clusters. They also found a pseudo-neurological (e.g., paralysis, blindness) cluster which may conceptually overlap with our 'Loss of function' cluster, in addition to female reproductive, cardiorespiratory, headache/other pain, and skin symptoms [30]. While there are commonalities across studies, particularly in musculoskeletal and gastrointestinal clusters, variations are likely due to differences in checklist length, symptoms included, and characteristics of study samples.

The PSQ-51 was translated and back-translated by two bilingual translators as per the recommendations [16]. The resulting English version is easily understandable and an accurate and realistic translation of the original Dutch version. Surprisingly, however, the frequency of symptoms was considerably higher in the UK sample compared to the Dutch samples. While the higher symptom ratings may just be a peculiarity of the UK sample, it must be acknowledged that the data collection for the UK sample took place partially during the height of the COVID-19 pandemic, in which health anxiety peaked [31]. It also could be indicative of poorer health status in the UK, or could be speculatively attributed to cultural differences in illness perceptions, health inequalities [32]), or levels of distress between the two countries. For example, the prevalence of the Distressed personality (Type D), is higher in UK general population samples (e.g. 45.6 %; [33] compared to Dutch general population samples [34]. Further, In the Dutch SOUL sample [14] it has been shown previously that symptoms of the PSQ are better accounted for by mental distress than by physical illnesses.

Our secondary aim was to test the fit of the model across the samples. The Chi-squared statistics indicated that the original model was a good fit to the data in the Dutch general population sample (i.e., non-significant) but not the other three samples, and the revised model, was found to be a good fit in the Dutch but not the UK sample. However, it must be noted that the Chi-squared lacks power in small samples and therefore may not be a useful indicator of model fit in the smaller UK sample. Furthermore, the test assumes multivariate normality, which is atypical of symptom experience data, and deviations from normality can result in model rejections even when the model is properly specified. [35]

This is further supported by some model fit indices assessed in the current analyses. The RMSEA fit indices for the 7-factor structure for the original and revised models were found to be good (<0.06) in all three Dutch samples, and acceptable (<0.07) in the UK sample. The RMSEA demonstrated that the fit was better for the revised model over the original model across all samples. The SRMR values were generally good for the three Dutch samples, and acceptable (<0.05) for the revised models. The SRMR value for the UK sample was, however, unacceptable for both the original and revised models. The CFI values for both models in all three samples did not reach the threshold of acceptance (<0.95), although the best fit was indicated by the Dutch psychiatric sample in the revised model (>0.90). This may suggest that the model may be misspecified [23]. However, the CFI statistic assumes that all latent variables are uncorrelated which is not the case given that physical symptoms regularly co-occur, particularly in specific illness populations. The TLI fit indices suggested better model fit with all models meeting the threshold of acceptable fit (0.80).

The seven-factor structure showed generally good internal consistency across all samples. The total scale was deemed excellent; however, this is not surprising given the large number of items. Each symptom subscale was deemed acceptable to good except for 'Hot Flushes' and 'Loss of function' in the two Dutch samples. It must be acknowledged that the self-administered nature of the PSQ fundamentally brings with it the usual issues surrounding the use of self-report measures for health data [36]; however, as the experience of symptoms is subjective, self-report is often the simplest way to access this data.

Taken together, the results of the CFA are mixed but promising. The RMSEA and TLI suggest decent model fit for all samples, however the Chi-squared statistic and SRMR indicate good model fit for the Dutch samples but not for the UK sample. Given that RMSEA has been regarded as one of the most informative model fit indices (Hooper et al., 2005), we could tentatively suggest that the model is generally well specified, however, the stability of the factor structure, particularly in the UK sample should be interpreted cautiously. This uncertainty is not unsurprising given several statistical and conceptual issues associated with the assessment of physical symptoms as a multi-dimensional construct. Firstly, symptom data is often right-skewed, which has implications for

the specification of model fit as demonstrated. All symptoms are not necessarily demonstrable of ill health so may be endorsed by healthy participants, may not always co-occur in the same patterns and can be experienced differently from person to person [37]. A questionnaire asking about the experience of symptoms can be interpreted very differently by individuals due to a variety of intrinsic and extrinsic factors, most notably by levels of distress and individual illness representations [38]. This is highlighted by the variety of symptoms included on the PSQ-51 and the stark difference in number of symptoms reported, between samples. Nevertheless, the UK sample demonstrated the best Cronbach's alpha.

It should be noted that response rates varied from 42 % to 80 % in the Dutch samples, with a likely underrepresentation of younger male participants. For the UK sample a response rate was not available but given the gender distribution, there must have been a considerable underrepresentation of males. It cannot be excluded that some of the results could be attributed to self-selection, as subjects who experience more symptoms may be more likely to participate. This is especially relevant for outcomes of symptom frequency, but perhaps less for the factor structure or consistency.

In summary, the PSQ-51 is a simple, inexpensive, and practical tool for the quick assessment of physical symptoms in both English and Dutch-speaking populations. While the SSS-8 [39] and PHQ-15 [6] are valued for their brevity and robust psychometric properties across diverse populations, making them particularly useful in time-limited clinical settings, the PSQ-51 provides greater depth and breadth in symptom assessment. The PSQ-51 therefore may be particularly useful in contexts requiring comprehensive symptom profiling, such as in-depth clinical consultations or research studies exploring distinct symptom experiences. The scale can be used as a checklist of symptoms and potentially as a one factor approach indicative of a general tendency to experience bodily discomfort [14]. Additionally, out of 51 items, 42 symptoms can be categorised into seven relatively distinct factors based on conceptually similar symptoms, however, this factor solution is not entirely stable, and the current study suggests we may not be able to consistently differentiate between symptom clusters in different populations.

The reporting of specific symptoms in isolation does not necessarily indicate a specific pathology or condition [37] yet consideration of physical symptoms as distinct symptom clusters is not straightforward. The PSQ-51 may still be useful in clinical consultation as it may identify symptoms that are not the direct focus of the consultation and could be used to screen for elevated somatic symptom burden and measure illness severity or response to treatment, which would be particularly useful in complex conditions and multimorbidity. Further, it may be particularly useful in certain research contexts.

CRediT authorship contribution statement

Sarah F. Allen: Writing – review & editing, Writing – original draft, Project administration, Formal analysis, Data curation, Conceptualization. **Albert M. van Hemert:** Writing – review & editing, Writing – original draft, Project administration, Methodology, Formal analysis, Data curation. **Margot W.M. de Waal:** Writing – review & editing, Project administration, Methodology, Data curation. **Christina van der Feltz-Cornelis:** Writing – review & editing, Writing – original draft, Supervision, Methodology, Conceptualization.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.genhosppsych.2025.05.002.

Data availability

Data will be made available on request.

References

[1] Rhodes VA, Watson PM. Symptom distress—The concept: Past and present. In: Semin Oncol. Elsevier; 1987.
[2] Walker EA, Unützer J, Katon WJ. Understanding and caring for the distressed patient with multiple medically unexplained symptoms. J Am Board Fam Pract 1998;11(5):347–56.
[3] Cowen R, et al. Assessing pain objectively: the use of physiological markers. Anaesthesia 2015;70(7):828–47.
[4] Van Hemert A, De Waal M, Van Rood Y. Meetinstrumenten bij somatoforme stoornissen. Tijdschr Psychiatr 2004;46(10):693–6.
[5] Zijlema WL, et al. How to assess common somatic symptoms in large-scale studies: a systematic review of questionnaires. J Psychosom Res 2013;74(6):459–68.
[6] Kroenke K, Spitzer RL, Williams JB. The PHQ-15: validity of a new measure for evaluating the severity of somatic symptoms. Psychosom Med 2002;64(2):258–66.
[7] Derogatis L. SCL-90-R Administration, Scoring and Procedures Manual, Minneapolis, MN. National Computer Systems. Inc.[Google Scholar]; 1994.
[8] de Vroege L, et al. Validation of the PHQ-15 for somatoform disorder in the occupational health care setting. J Occup Rehabil 2012;22(1):51–8.
[9] Budtz-Lilly A, et al. A new questionnaire to identify bodily distress in primary care: the 'BDS checklist'. J Psychosom Res 2015;78(6):536–45.
[10] Eriksen HR, Ihlebæk C, Ursin H. A scoring system for subjective health complaints (SHC). Scand J Public Health 1999;27(1):63–72.
[11] Cohen S, Hoberman HM. Positive events and social supports as buffers of life change stress 1. J Appl Soc Psychol 1983;13(2):99–125.
[12] Allen SF, Wetherell MA, Smith MA. The Cohen–Hoberman inventory of physical symptoms: Factor structure, and preliminary tests of reliability and validity in the general population. Psychol Health 2017;32(5):567–87.
[13] APA. Diagnostic and Statistical Manual of Mental Disorders. 2000.
[14] de Waal MW, et al. The reporting of specific physical symptoms for mental distress in general practice. J Psychosom Res 2005;59(2):89–95.
[15] de Waal MW, et al. The role of comorbidity in the detection of psychiatric disorders with checklists for mental and physical symptoms in primary care. Soc Psychiatry Psychiatr Epidemiol 2009;44(1):78–85.
[16] Hilton A, Skrutkowski M. Translating instruments into other languages: development and testing processes. Cancer Nurs 2002;25(1):1–7.
[17] Arnold IA, et al. Medically unexplained physical symptoms in primary care: a controlled study on the effectiveness of cognitive-behavioral treatment by the family physician. Psychosomatics 2009;50(5):515–24.
[18] de Beurs E, et al. Routine outcome monitoring in the Netherlands: practical experiences with a web-based strategy for the assessment of treatment outcome in clinical practice. Clin Psychol Psychother 2011;18(1):1–12.
[19] van Noorden, M.S., et al., Gender differences in clinical characteristics in a naturalistic sample of depressive outpatients: the Leiden routine outcome monitoring study. J Affect Disord, 2010. 125(1–3): p. 116–123.
[20] Branley D, Covey J, Hardey M. Online Surveys: Investigating Social Media Use and Online Risk. SAGE Publications, Ltd; 2014.
[21] StataCorp L. Stata statistical software: release 17 college station. TX StataCorp LP 2021;5:231–9.
[22] Barrett P. Structural equation modelling: adjudging model fit. Personal Individ Differ 2007;42(5):815–24.
[23] Hu LT, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. Struct Equ Model Multidiscip J 1999;6(1):1–55.
[24] Steiger JH. Understanding the limitations of global fit assessment in structural equation modeling. Personal Individ Differ 2007;42(5):893–8.
[25] Hooper D, Coughlan J, Mullen M. Evaluating model fit: a synthesis of the structural equation modelling literature. In: 7th European Conference on Research Methodology for Business and Management Studies; 2008.
[26] De Waal MWM, et al. Somatoform disorders in general practice: prevalence, functional impairment and comorbidity with anxiety and depressive disorders. Br J Psychiatry 2004;184(6):470–6.
[27] Kaiser HF. The varimax criterion for analytic rotation in factor analysis. Psychometrika 1958;23(3):187–200.
[28] Kroenke K, et al. A symptom checklist to screen for somatoform disorders in primary care. Psychosomatics 1998;39(3):263–72.
[29] Fink P, et al. Symptoms and syndromes of bodily distress: an exploratory study of 978 internal medical, neurological, and primary care patients. Psychosom Med 2007;69(1):30–9.
[30] Gara MAS, Escobar JI, Holman A, Waitzkin H. A hierarchical classes analysis (HICLAS) of primary care patients with medically unexplained somatic symptoms. Psychiatry Res 1998;19(1):77–86.
[31] Heinen A, et al. Understanding health anxiety in the COVID-19 pandemic. Int J Soc Psychiatry 2022;68(8):1756–63.
[32] Doorslaer EV, Koolman X. Explaining the differences in income-related health inequalities across European countries. Health Econ 2004;13(7):609–28.
[33] Allen SF, Wetherell MA, Smith MA. A one-year prospective investigation of type D personality and self-reported physical health. Psychol Health 2019;34(7):773–95.
[34] Denollet J. DS14: standard assessment of negative affectivity, social inhibition, and type D personality. Psychosom Med 2005;67(1):89–97.
[35] McIntosh CN. Rethinking fit assessment in structural equation modelling: a commentary and elaboration on Barrett (2007). Personal Individ Differ 2007;42(5):859–67.
[36] Justice AC, et al. Sensitivity, specificity, reliability, and clinical validity of provider-reported symptoms: a comparison with self-reported symptoms. JAIDS J Acquir Immune Deficien Syndr 1999;21(2):126–33.

[37] Petersen S, et al. Illness and symptom perception: a theoretical approach towards an integrative measurement model. Clin Psychol Rev 2011;31(3):428–39.
[38] Petrie K, Weinman J. Why illness perceptions matter. Clin Med 2006;6(6):536.
[39] Gierk B, et al. The somatic symptom scale–8 (SSS-8): a brief measure of somatic symptom burden. JAMA Intern Med 2014;174(3):399–407.