Journal of Psychosomatic Research 79 (2015) 378–383

Contents lists available at ScienceDirect

Journal of Psychosomatic Research

Somatic symptom reports in the general population: Application of a bi-factor model to the analysis of change

Jarry T. Porsius [a,1], Astrid L. Martens [a,b,*,1], Pauline Slottje [a], Liesbeth Claassen [a,c], Joke C. Korevaar [d], Danielle R.M. Timmermans [a,c], Roel Vermeulen [b], Tjabe Smid [a,e]

[a] Department of Public and Occupational Health, EMGO Institute for Health and Care Research, VU University Medical Center, Amsterdam, The Netherlands
[b] Institute for Risk Assessment Sciences (IRAS), Division Environmental Epidemiology, Utrecht University, Yalelaan 2, 3584 CM Utrecht, The Netherlands
[c] National Institute for Public Health and the Environment (RIVM), Bilthoven, The Netherlands
[d] Netherlands Institute for Health Services Research (NIVEL), The Netherlands
[e] KLM Health Services, Schiphol, The Netherlands

ARTICLE INFO

Article history:
Received 29 May 2015
Received in revised form 14 September 2015
Accepted 18 September 2015

Keywords:
Bi-factor model
Latent structure
Measurement invariance
Self-report
Somatic symptoms
Symptom patterns

ABSTRACT

Objective: To investigate the latent structure of somatic symptom reports in the general population with a bi-factor model and apply the structure to the analysis of change in reported symptoms after the emergence of an uncertain environmental health risk.
Methods: Somatic symptoms were assessed in two general population environmental health cohorts (AMIGO, n = 14,829 & POWER, n = 951) using the somatization scale of the four-dimensional symptom questionnaire (4DSQ-S). Exploratory bi-factor analysis was used to determine the factor structure in the AMIGO cohort. Multi-group and longitudinal models were applied to assess measurement invariance. For a subsample of residents living close to a newly introduced power line (n = 224), we compared a uni- and multidimensional method for the analysis of change in reported symptoms after the power line was put into operation.
Results: We found a good fit (RMSEA = 0.03, CFI = 0.98) for a bi-factor model with one general and three symptom specific factors (musculoskeletal, gastrointestinal, cardiopulmonary). The latent structure was found to be invariant between cohorts and over time. A significant increase ($p < .05$) was found only for musculoskeletal and gastrointestinal symptoms after the power line was put into operation.
Conclusions: In our study we found that a bi-factor structure of somatic symptoms reports was equivalent between cohorts and over time. Our findings suggest that taking this structure into account can lead to a more informative interpretation of a change in symptom reports compared to a unidimensional approach.

Introduction

The experience of non-specific somatic symptoms such as headaches or back pain has negative effects on daily functioning in a considerable proportion of the general population, and is a major cause of health care utilization [1–3]. These experiences are typically assessed with self-report questionnaires [4] and are frequently used in varying research disciplines such as psychosomatic medicine [e.g. 5] or environmental health [e.g. 6,7]. In most studies the total symptom score is analyzed and/or the individual symptoms separately. Neither approach reflects the multifactorial origin of reporting somatic symptoms [8,9].

Self-report symptom questionnaires such as the PHQ-15 [10] or the SCL-90 SOM [11] were designed to measure the experience of distressing somatic symptoms. A high score (clinical cut-off scores are generally provided) is interpreted as an indication of somatization. Although these questionnaires were designed to measure one underlying construct (i.e. somatization), there is evidence for the latent structure to be multi- rather than unidimensional [12–15]. This is due to the existence of specific symptom patterns, such as symptoms pertaining to musculoskeletal or gastrointestinal complaints. A wide range of influences can lead to higher scores on symptoms from a specific symptom group (e.g. infections, diseases, and psychosocial distress) while scores on other domains are less affected. It is therefore plausible that additional variance in reported symptoms is explained by symptom specific factors. The bi-factor model separates the general variance of scores on all symptoms (i.e. general factor, representing a general tendency to report symptoms), from the unique variance of scores relating to specific

Abbreviations: 4DSQ, Four-dimensional symptom questionnaire; 4DSQ-S, Somatization scale of the 4DSQ; RMSEA, Root mean square error of approximation; CFI, Comparative fit index; MI, Measurement invariance; WLSMV, Weighted least squares means and variance adjusted.

* Corresponding author at: Institute for Risk Assessment Sciences, Utrecht University, PO Box 80178, Yalelaan 2, 3508 TD Utrecht, The Netherlands.
E-mail address: a.l.martens@uu.nl (A.L. Martens).
[1] These authors contributed equally to this work.

http://dx.doi.org/10.1016/j.jpsychores.2015.09.006
0022-3999/

symptom groups (i.e. specific factors). This model allows studying both components of somatic symptom reporting simultaneously.

So far, only a few studies [16–18] have applied a bi-factor model to data gathered with symptom questionnaires. These studies showed that specific factors explain unique variance over and above the common variance in symptom reporting explained by a general factor. In addition the bi-factor model has been shown to provide a better fit than alternative factor models. However, the evidence gathered so far is limited and mainly based on two cross-sectional clinical samples using two different symptom questionnaires. There may be differences in the underlying structure between populations and symptom questionnaires which could impact application to health effect studies.

In order to compare symptom scores on underlying constructs between different populations and over time, measurement invariance (MI) must be established [19]. MI refers to the underlying factor structure being equivalent across samples and over time. Changes in the underlying factor structure complicate the interpretation of differences in symptom scores. When the structure is not invariant a score difference could reflect a change in the score on the underlying latent construct, or reflect a change in the construct itself. If MI can be established, there may be useful practical applications of the bi-factor model to intervention studies using somatic symptom reports as an outcome. One could assess the effect of an intervention or exposure on general symptom reporting (i.e. over and above reporting symptoms from specific symptom groups), as well as on symptom specific factors (i.e. over and above general symptom reporting). A potential benefit of a bi-factor model is the greater conceptual clarity provided by a separation between general and specific variance [20].

The aim of the present study is threefold. First, we aim to test the structural validity of a bi-factor model for the somatization scale of the 4DSQ [4DSQ–S, 21] in a large general population sample. Structural validity of this subscale has not been investigated before. Second, we assess MI of the resulting latent structure by comparing the structure between two different general population samples, as well as across time in one sample. Third and last, we apply a bi-factor structure to analyze change in symptom reports after the emergence of an uncertain environmental health risk. In previous work we found a larger increase in overall reported somatic symptoms after a new power line was put into operation for residents living close by, compared to a control group of residents living farther away [22]. We extend those findings by evaluating the change in reported somatic symptoms in line with the underlying latent structure of the 4DSQ-S.

Methods

Participants

Participants were members of the adult general population in the Netherlands enrolled in two different cohorts. The first cohort (AMIGO) was set up to study environmental and occupational determinants of diseases and symptoms [see 23 for a full description]. The AMIGO cohort at baseline consisted of 14,829 subjects of which 50.2% men. The mean age of the AMIGO participants was 51 years (SD = 9). The second cohort (POWER) was set up to study health responses to the introduction of a new high-voltage power line [see 24 for a full description]. At baseline the POWER cohort consisted of 951 subjects of which 46% men. Mean age of the participants was 52 years (SD = 13). The longitudinal models to assess measurement invariance were based on a total of 1241 subjects. This number is higher than the number of participants at baseline, because new subjects were enrolled at T2 [22]. For the analysis of change we focused on the group of residents within 300 m of the new high voltage power line (n = 224), as we established in previous work that only this group experienced more symptoms after the line was put into operation [22]. The overall response rate to the baseline questionnaires was similar in both cohorts (AMIGO: 16%, POWER: 19%).

Procedures

In both cohorts invitations were sent through postal mail. Both studies were presented to participants as longitudinal environmental health studies, which consisted of filling out questionnaires by one adult per household about health and the environment. To reduce the chance of response bias, there was no mentioning of power lines in the POWER cohort invitation letter.

The AMIGO cohort subjects (31–65 years old) were recruited using a national information network of general practitioners established at the Netherlands Institute for Health Services Research (NIVEL), called NIVEL Primary Care Database. Participants were invited between April 2011 and July 2012. For the POWER cohort one member older than 18 of each household within 500 m of the planned construction of a new power line (n = 2379) was invited to participate, as well as a random stratified sample of households within 500–2000 m (n = 2382). Data was collected before the power line was put into operation, starting in June 2012 (T1), 5 months later (T2), and after the power line was put into operation, 12 months (T3) and 18 months (T4) after the baseline measurement (T1). The study protocols of both studies were approved by the Medical Ethics Committee of the research boards of the involved institutes, and all participants participated voluntarily with informed consent.

Measures

Somatic symptoms

In both cohorts the somatization scale of the 4DSQ [21] was used to measure self-reported somatic symptoms. The 4DSQ consists of 4 scales measuring distress, depression, anxiety and somatization, but only the somatization scale was administered in our study samples. The somatization scale (4DSQ-S) consists of 16 non-specific somatic symptoms (e.g. headaches, low back pain, and dizziness) commonly reported in general practices (see Fig. 1 for a list of all symptoms). For each symptom, participants indicated whether they were bothered by it during the previous week on a 5-point scale (ranging from no, through to constantly). The scores were trichotomized before analysis (no = 0; sometimes = 1, regularly/often/constantly = 2) [21].

Statistical analyses

To answer the first research question regarding the underlying latent structure of the 4DSQ-S we conducted a categorical exploratory bi-factor analysis on the AMIGO baseline data with Bi-Geomin rotation [25] and WLSMV estimation. Two (1 general, 1 specific factor) up to six (1 general, 5 specific factors) factor solutions were considered and one bi-factor specification was selected for a confirmatory analysis, based on the theoretical interpretation of the models as well as the statistical fit. We assigned items to a factor only if the factor loading for that item on that factor was greater than 0.30. The variances of the common factors were identified by fixing the loading of the first item to one. Root mean square error of approximation (RMSEA) and comparative fit index (CFI) were used to assess model fit. For RMSEA, models with values ≤0.06 had acceptable fit and for CFI values ≥0.95 had acceptable fit [26].

To answer the second research question regarding MI of the 4DSQ-S we fitted a multi-group model where we increasingly constrained more parameters to be equal across the baseline AMIGO and POWER cohort samples to assess invariance [19,27]. The following models were tested consecutively: configural invariance (factor loadings freely estimated and thresholds constrained), loading invariance (factor loadings and thresholds constrained), and residual invariance (factor loadings, thresholds and residual variance constrained). We compared the models using the criteria suggested by Chen et al. [28] to establish MI: a decrease in CFI of ≥0.01, and an increase in RMSEA of ≥0.015 were

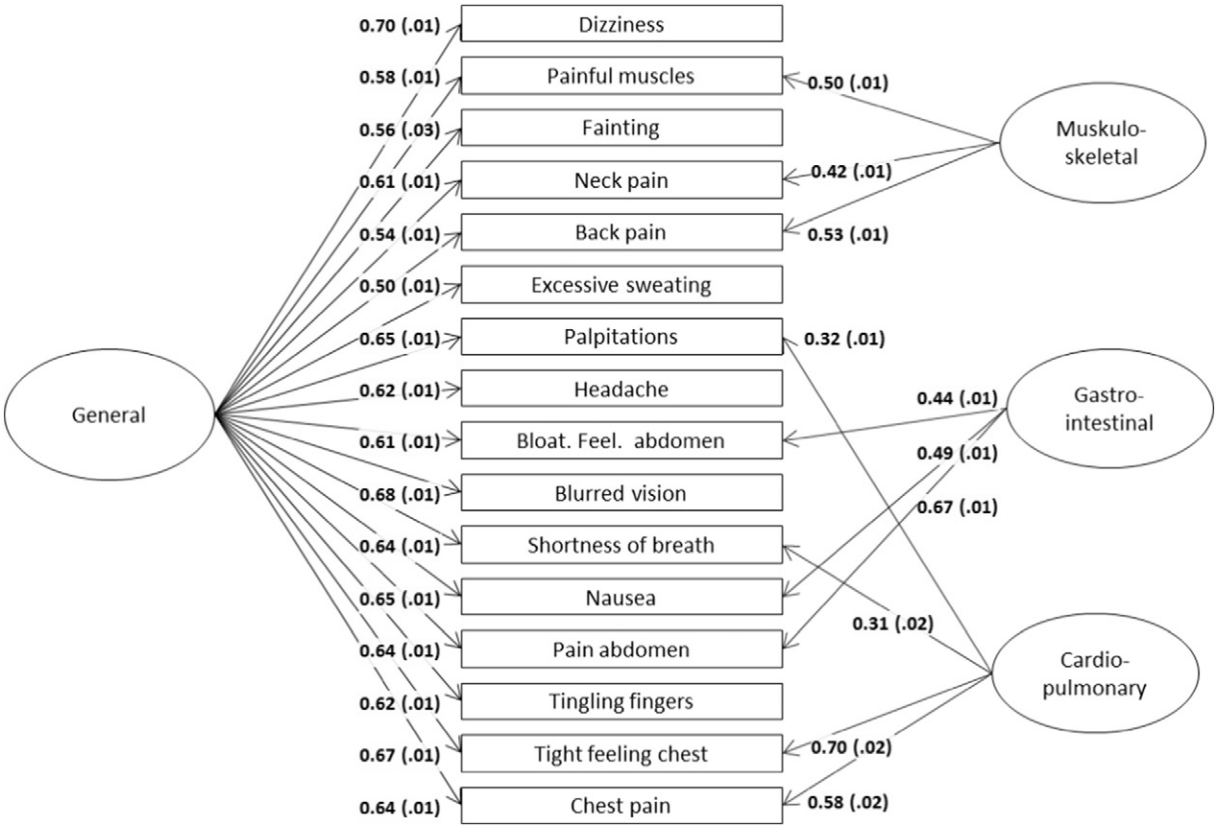

Fig. 1. AMIGO confirmatory baseline bi-factor model with standardized factor loadings and standard errors.

taken as an unacceptable decrease in model fit (i.e. indicating absence of MI).

To test whether the measurement model was invariant across time within the full POWER cohort we used similar procedures [see 29]. In addition to cross-sectional orthogonal constraints between the general and specific factors and constraints necessary for identification, we added orthogonal constraints to the longitudinal relations between latent factors [30]. Residuals of the same indicators over time were allowed to correlate. Comparison of the models and assessment of model fit was similar to the multi-group models.

To answer the third and final research question regarding the analysis of change in reported somatic symptoms, we applied linear mixed models with an unstructured residual covariance structure to the data gathered from participants in the POWER cohort living within 0–300 m of the new power line (n = 224). Sum scores were calculated for the specific symptom factors identified in the bi-factor analysis of the 4DSQ-S, as well as the total sum score as indication of the general factor. Time was entered as categorical predictor (T2, T3, and T4) with T1 as reference category. We compared two different ways to analyze change in reported somatic symptoms. First, the standard approach where we assessed the effects of time (i.e. change from baseline) on the specific and total sum scores. Second, a bi-factor approach, where we adjusted the effects of time on specific symptom scores for the total sum score (minus overlapping items). This analysis indicates to what extent a change in reports from specific symptom groups was confounded by a general symptom reporting pattern. For the mixed model with the total sum score as outcome, we adjusted the effect of time for scores on the specific symptom groups. This analysis indicates to what extent a change in overall symptom reporting was confounded by specific symptom patterns.

SPSS version 20 was used for the mixed models analyses; Mplus version 7.2 was used for all other analyses.

Results

Latent structure of the 4DSQ-S

Information about all symptoms was missing for 330 (2%) subjects in the AMIGO cohort, and for 11 (1%) in the baseline POWER cohort. These subjects were not included in further analyses. The exploratory bi-factor analyses indicated that a model with one general and three specific factors provided a good fit (RMSEA = 0.027, CFI = 0.992) to the AMGIO baseline data, and was most readily interpretable. We named the specific factors musculoskeletal (muscle pain, neck pain, back pain), gastrointestinal (bloated feeling in abdomen, nausea, pain in stomach) and cardiopulmonary (palpitations, shortness of breath, tight feeling in chest, pain in chest). Fit indices of the confirmatory bi-factor model in the baseline AMIGO (RMSEA = 0.032, CFI = 0.984) and POWER cohort (RMSEA = 0.029, CFI = 0.985) indicated that the selected bi-factor model fitted well to the data from both cohorts. Fig. 1 presents the confirmatory AMIGO baseline model with standardized factor loadings for the 4DSQ-S items on the general and symptom specific factors.

Measurement invariance of the bi-factor model

Table 1 shows the fit statistics for the measurement invariance models. For both the multi-group and longitudinal models, the fit indices indicated a good fit to the data (i.e. RMSEA ≤ 0.06 and CFI ≥ 0.95). If the RMSEA and CFI values do not deteriorate in the more constrained models (i.e. the loading and residual invariance models), this is indicative of measurement invariance. As can be seen in Table 1, there was no decrease in fit for the more restricted models in the multi-group and longitudinal comparisons. The differences are well below the

Table 1
Model fit indices for the measurement invariance models.

Model	n	Chi-square	d.f.	*p*-Value	CFI	RMSEA	RMSEA 90%-CI
Multi-group baseline (AMIGO & POWER)							
Configural invariance[a]	15,439	1636.383	200	<0.001	0.983	0.031	0.029–0.032
Loading invariance[b]	15,439	1421.149	222	<0.001	0.986	0.026	0.025–0.028
Residual invariance[c]	15,439	1253.197	238	<0.001	0.988	0.024	0.022–0.025
Longitudinal[d] (POWER)							
Configural invariance[a]	1241	1920.475	1589	<0.001	0.987	0.013	0.011–0.015
Loading invariance[b]	1241	1971.465	1652	<0.001	0.987	0.012	0.010–0.015
Residual invariance[c]	1241	1960.503	1697	<0.001	0.989	0.011	0.009–0.013

[a] Factor loadings freely estimated and thresholds constrained.
[b] Factor loadings and thresholds constrained.
[c] Factor loadings, thresholds and residual variance constrained.
[d] The item assessing 'fainting' was removed from the longitudinal models due to counts of 0 in the higher categories (sometimes = 1, regularly/often/constantly = 2) at some time-points.

suggested cut-off points for establishing MI (a decrease in CFI of ≥0.01, and an increase in RMSEA of ≥0.015 [28]).

Change in somatic symptom patterns

Table 2 displays the parameter estimates for the change from baseline for the general and symptom specific scores in participants living close to the newly introduced power line. The mean scores at baseline (T1) were 4.02 (SD = 3.81) for the total sum score of somatic complaints, 1.53 (SD = 1.55) for musculoskeletal, 0.63 (SD = 1.01) for gastrointestinal, and 0.42 (SD = 0.87) for cardiopulmonary complaints. After the new power line was put into operation we found an increase in overall symptom reports (on the total sum score) from baseline (previously reported in [22]). When these estimates were adjusted for scores on the symptom specific factors, in line with a bi-factor model, we no longer found a significant change from baseline in the total score. This suggests that the change we found using a sum score of the total scale was mainly due to change in symptom specific factors. This was confirmed in the mixed models for the symptom specific factors. When we inspected the estimates of the specific symptom scores, a significant increase from baseline was seen for musculoskeletal symptoms at T2, T3 and T4, and for gastrointestinal symptoms at T3. When adjusting for a general symptom reporting pattern, only the change at T4 for musculoskeletal and T3 for gastrointestinal symptoms remained significant.

Discussion

This study applied a multidimensional approach to the analysis of somatic symptom reports in a general population. We found that:

1. A bi-factor model with one general and three specific factors provided a good fit to symptom data from two large general population samples, providing further evidence for a multidimensional latent structure of somatic symptom reports in the general population.
2. The bi-factor structure was stable when measurement invariance was evaluated across two large general population samples, and over time in one sample.
3. Application of the bi-factor structure to a general population sample showed that the longitudinal course of symptom reports differed for the different symptom patterns after emergence of an uncertain environmental health risk.

Previous studies used unidimensional methods to analyze effects of interventions or environmental exposures on symptoms see [31] for an overview. The application of bi-factor models to somatic symptom reports has so far been rare. In two cross-sectional samples a good fit was found for a bi-factor model applied to two different somatic symptom questionnaires; the MMPI-2-RF-RC1 [17] and the frequently used PHQ-15 [16]. Both studies found a bi-factor structure with a general factor and a number of specific symptom factors (specific factors found in [16]: pain, gastroenterological, cardio-pulmonary, fatigue; and in [17]:

Table 2
Longitudinal development of general and specific somatic complaints in participants living close (0–300 m, n = 224) to a newly introduced power line.

	Beta estimates (95% CI)		
	T2[a]	T3[a]	T4[a]
Somatic complaints			
Total sum score (0–32)	.33 (−.26, .92)	.80 (.23, 1.38)**	.85 (.21, 1.48)**
Adjusted for specific symptoms[b]	−.04 (−.29, .22)	−.09 (−.33, .15)	−.07 (−.31, .18)
Musculoskeletal complaints			
Specific sum score (0–6)	.24 (.02, .46)*	.30 (.03, .57)*	.51 (.25, .76)**
Adjusted for general symptoms[c]	.20 (−.02, .42)	.19 (−.07, .44)	.44 (.19, .68)**
Gastrointestinal complaints			
Specific sum score (0–6)	.10 (−.10, .30)	.29 (.11, .47)**	.11 (−.07, .28)
Adjusted for general symptoms[c]	.08 (−.10, .25)	.22 (.05, .38)*	.01 (−.15, .17)
Cardiopulmonary complaints			
Specific sum score (0–8)	−.01 (−.16, .13)	.07 (−.07, .21)	.09 (−.10, .28)
Adjusted for general symptoms[c]	−.05 (−.19, .09)	.00 (−.14, .14)	.02 (−.16, .19)

* $p < .05$.
** $p < .01$.
[a] T1 (10 months *before* power line was put into operation) is reference category. T2 = 5 months *before* the line was put into operation, T3 = 2 months *after* the line was put into operation, T4 = 7 months *after* the line was put into operation.
[b] Estimates adjusted for the sum scores of the symptom specific factors (i.e. musculoskeletal, gastrointestinal and cardiopulmonary).
[c] Estimates adjusted for the total sum score minus overlapping items.

gastrointestinal, head pain, neurological). Although there are some differences, the overall factor structure in these studies is similar to our findings. Differences may be explained by differences in specific symptoms included in the used questionnaires (e.g. no fatigue symptoms in 4DSQ-S).

Our study found that symptom patterns can be affected differently by the emergence of an uncertain environmental health risk. Using a total sum score we previously reported an effect of the introduction of a power line on somatic symptom reports [22]. In the present study we found that the introduction of a new power line was uniquely associated with reporting more musculoskeletal and gastrointestinal somatic complaints when accounting for the general factor. This finding illustrates the relevance of acknowledging the underlying bi-factor structure when studying the mechanisms and determinants of a change in symptom reports. A total sum score of somatic complaints does not reflect just one source of variation which blurs the interpretation of a change if one does not take into account the other known sources (i.e. symptom specific factors). This could particularly be problematic when the changes over time in these sources of variation are in opposite directions. As a result one may develop inappropriate theories to explain the research findings, or implement ineffective intervention strategies [32].

More research into determinants of change in symptom scores on the general and symptom specific factors is needed. Both the general and specific factors may reflect influences of diseases, environmental factors and psychosomatic mechanisms. Findings of Witthoft and colleagues [16] suggest that the general as well as certain specific factors are associated with functional somatic syndromes (e.g. irritable bowel syndrome). They hypothesize that where symptom specific factors might reflect temporary (environmental) influences, the general factor could reflect a disposition relevant for perpetuation of symptom experiences. This might explain the absence of an effect of time on the general factor in the POWER cohort. Any (temporary) intervention or illness may be more likely to affect symptom specific factors, which, if combined with a higher score on the general factor, could be perpetuated and lead to chronic health problems.

There are some limitations to the interpretation of our findings. First, we used a regression based method to study the longitudinal course for each symptom pattern. The assumption in all regression based methods that a construct is measured without error is untenable. Hence, it may be better to investigate change on the different factors within a structural equation model (e.g. latent growth curve model). However, this method requires a large sample size due to the computational complexity of the model. Our experimental group (the subgroup of the POWER cohort living within 300 m from the new power line ($n = 224$)) was too small to estimate such a model. The current method has the advantage of reduced complexity, but cannot perfectly separate general and specific variance within each item. This is disadvantageous when assessing effects on the general factor over and above effects on specific symptom patterns. The overlap in items leads to overcorrection when including all specific symptom patterns as a covariate, and therefore to less power to demonstrate the unique effect on the general factor.

Second, we used a single symptom questionnaire to address our research questions (i.e. 4DSQ-S). Although this questionnaire represents the major symptom specific groups identified in a review of somatic symptom questionnaires [4], one might find a different latent structure when other symptoms are probed. To improve the analysis of symptoms in health effect studies, it is important to use a symptom questionnaire which covers all potentially relevant symptom groups (see [33]).

Third and last, we did not investigate whether alternative models such as a hierarchical or correlated factor model would lead to different conclusions regarding change over time. Hierarchical or correlated group factor models are the most likely candidates as these are also multidimensional [34]. A disadvantage of the correlated factor model is that it does not explicitly represent a general tendency to report symptoms (general factor). The hierarchical model specifies that there is no direct relationship between the items and the general construct, instead this relationship is mediated by the specific factors [35]. We prefer the bi-factor model for our study because of the conceptual differences between the two models. The bi-factor model specifies the specific factors as orthogonal from the general factor. Because of this representation it is possible to study whether symptom scores are affected differently over time for the general and the specific factors.

Strengths of our study are the large sample size to assess the latent structure of symptom reports as well as the extensive assessment of measurement invariance, and its application to the analysis of change with an ecologically valid example. The importance of measurement invariance as a prerequisite to interpret scores on a questionnaire has been widely acknowledged [e.g. 19,29,36], but rarely addressed for the use of symptom questionnaires. By establishing invariance of the bi-factor model for the 4DSQ-S in a general population we found support to study symptom scores based on the underlying constructs over time and to compare scores between different general population samples.

In conclusion, we demonstrated the potential use of applying the bi-factor model in an analysis of change in reported symptoms, using the example of an emerging uncertain environmental health risk. Our findings have implications for the analysis and interpretation of symptom checklists in psychosomatic and (environmental) health research. Application of the bi-factor structure can lead to a more informative interpretation of changes in somatic symptom reporting, as it allows to separately evaluate the impact of an intervention or change in the environment on the longitudinal course of each symptom pattern. Future health effect studies are needed to compare different methods to approach the multidimensional nature of symptom reports, as well as to improve insight in determinants of specific symptom patterns.

Conflicts of interest and source of funding

The research reported in this article was funded by The Netherlands Organization for Health Research and Development (ZonMw) within the Electromagnetic Fields and Health Research program, grant numbers 85600006 and 85200001. The authors declare no conflicts of interest.

Acknowledgments

We would like to thank Conor Dolan for statistical advice and comments on an earlier version of the paper.

References

[1] D. Bruusgaard, H. Tschudi-Madsen, C. Ihlebæk, Y. Kamaleri, B. Natvig, Symptom load and functional status: results from the Ullensaker population study, BMC Public Health 12 (2012) 1085, http://dx.doi.org/10.1186/1471-2458-12-1085.
[2] A.J. Barsky, E.J. Orav, D.W. Bates, Somatization increases medical utilization and costs independent of psychiatric and medical comorbidity, Arch. Gen. Psychiatry 62 (2005) 903–910, http://dx.doi.org/10.1001/archpsyc.62.8.903.
[3] K.J. Petrie, K. Faasse, F. Crichton, A. Grey, How common are symptoms? Evidence from a New Zealand national telephone survey, BMJ Open 4 (2014), e005374http://dx.doi.org/10.1136/bmjopen-2014-005374.
[4] W.L. Zijlema, R.P. Stolk, B. Löwe, W. Rief, P.D. White, J.G.M. Rosmalen, How to assess common somatic symptoms in large-scale studies: a systematic review of questionnaires, J. Psychosom. Res. 74 (2013) 459–468, http://dx.doi.org/10.1016/j.jpsychores.2013.03.093.
[5] R.J. Waldinger, M.S. Schulz, A.J. Barsky, D.K. Ahern, Mapping the road from childhood trauma to adult somatization: the role of attachment, Psychosom. Med. 68 (2006) 129–135.
[6] E.J. Bromet, J.M. Havenaar, S.F. Gluzman, N.L. Tintle, Psychological aftermath of the lviv air show disaster: a prospective controlled study, Acta Psychiatr. Scand. 112 (2005) 194–200, http://dx.doi.org/10.1111/j.1600-0447.2005.00566.x.
[7] A.C. Huizink, P. Slottje, A.B. Witteveen, J.A. Bijlsma, J.W.R. Twisk, N. Smidt, et al., Long term health complaints following the Amsterdam air disaster in police officers and fire-fighters, Occup. Environ. Med. 63 (2006) 657–662, http://dx.doi.org/10.1136/oem.2005.024687.
[8] H. Leventhal, Symptom reporting: a focus on process, Illn. Behav. Springer 1986, pp. 219–237.

[9] J.W. Pennebaker, Psychological factors influencing the reporting of physical symptoms, Sci. Self-Rep. Implic. Res. Pract. (2000) 299–315.
[10] K. Kroenke, R.L. Spitzer, J.B.W. Williams, The PHQ-15: validity of a new measure for evaluating the severity of somatic symptoms, Psychosom. Med. 64 (2002) 258–266.
[11] L.R. Derogatis, K. Rickels, A.F. Rock, The SCL-90 and the MMPI: a step in the validation of a new self-report scale, Br. J. Psychiatry 128 (1976) 280–289, http://dx.doi.org/10.1192/bjp.128.3.280.
[12] I.a.N.J. Deary, A taxonomy of medically unexplained symptoms, J. Psychosom. Res. 47 (1999) 51–59.
[13] A.C.H. Schat, E.K. Kelloway, S. Desmarais, The physical health questionnaire (PHQ): construct validation of a self-report scale of somatic symptoms, J. Occup. Health Psychol. 10 (2005) 363–381, http://dx.doi.org/10.1037/1076-8998.10.4.363.
[14] P. Fink, T. Toft, M.S. Hansen, E. Ørnbøl, F. Olesen, Symptoms and syndromes of bodily distress: an exploratory study of 978 internal medical, neurological, and primary care patients, Psychosom. Med. 69 (2007) 30–39, http://dx.doi.org/10.1097/PSY.0b013e31802e46eb.
[15] A. Budtz-Lilly, P. Fink, E. Ørnbøl, M. Vestergaard, G. Moth, K.S. Christensen, et al., A new questionnaire to identify bodily distress in primary care: the "BDS checklist.", J. Psychosom. Res. 78 (2015) 536–545, http://dx.doi.org/10.1016/j.jpsychores.2015.03.006.
[16] M. Witthöft, W. Hiller, N. Loch, F. Jasper, The latent structure of medically unexplained symptoms and its relation to functional somatic syndromes, Int. J. Behav. Med. 20 (2013) 172–183, http://dx.doi.org/10.1007/s12529-012-9237-2.
[17] M.L. Thomas, D.E.C. Locke, Psychometric properties of the MMPI-2-RF somatic complaints (RC1) scale, Psychol. Assess. 22 (2010) 492–503, http://dx.doi.org/10.1037/a0019229.
[18] M.L. Thomas, Rewards of bridging the divide between measurement and clinical theory: demonstration of a bifactor model for the brief symptom inventory, Psychol. Assess. 24 (2012) 101–113, http://dx.doi.org/10.1037/a0024712.
[19] R. Van de Schoot, P. Lugtig, J. Hox, A checklist for testing measurement invariance, Eur. J. Dev. Psychol. 9 (2012) 486–492, http://dx.doi.org/10.1080/17405629.2012.686740.
[20] F.F. Chen, A. Hayes, C.S. Carver, J.P. Laurenceau, Z. Zhang, Modeling general and specific variance in multifaceted constructs: a comparison of the bifactor model to other approaches, J. Pers. 80 (2012) 219–251, http://dx.doi.org/10.1111/j.1467-6494.2011.00739.x.
[21] B. Terluin, H.W.J. van Marwijk, H.J. Adèr, H.C.W. de Vet, B.W.J.H. Penninx, M.L.M. Hermens, et al., The four-dimensional symptom questionnaire (4DSQ): a validation study of a multidimensional self-report questionnaire to assess distress, depression, anxiety and somatization, BMC Psychiatry 6 (2006) 34, http://dx.doi.org/10.1186/1471-244X-6-34.
[22] J.T. Porsius, L. Claassen, T. Smid, F. Woudenberg, K.J. Petrie, D.R.M. Timmermans, Symptom reporting after the introduction of a new high-voltage power line: a prospective field study, Environ. Res. 138 (2015) 112–117, http://dx.doi.org/10.1016/j.envres.2015.02.009.
[23] P. Slottje, C.J. Yzermans, J.C. Korevaar, M. Hooiveld, R.C.H. Vermeulen, The population-based occupational and environmental health prospective cohort study (AMIGO) in the Netherlands, BMJ Open 4 (2014)http://dx.doi.org/10.1136/bmjopen-2014-005858 (e005858–e005858).
[24] J.T. Porsius, L. Claassen, T. Smid, F. Woudenberg, D.R.M. Timmermans, Health responses to a new high-voltage power line route: design of a quasi-experimental prospective field study in the Netherlands, BMC Public Health 14 (2014) 237, http://dx.doi.org/10.1186/1471-2458-14-237.
[25] R.I. Jennrich, P.M. Bentler, Exploratory bi-factor analysis: the oblique case, Psychometrika 77 (2012) 442–454, http://dx.doi.org/10.1007/s11336-012-9269-1.
[26] L. Hu, P.M. Bentler, Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives, Struct. Equ. Model A Multidiscip. J. 6 (1999) 1–55, http://dx.doi.org/10.1080/10705519909540118.
[27] R.E. Millsap, J. Yun-Tein, Assessing factorial invariance in ordered-categorical measures, Multivar. Behav. Res. 39 (2004) 479–515.
[28] F. Chen, P.J. Curran, K.A. Bollen, J. Kirby, P. Paxton, An empirical evaluation of the use of fixed cutoff points in RMSEA test statistic in structural equation models, Sociol. Methods Res. 36 (2008) 462–494, http://dx.doi.org/10.1177/0049124108314720.
[29] K.F. Widaman, E. Ferrer, R.D. Conger, Factorial invariance within longitudinal structural equation models: measuring the same construct across time, Child Dev. Perspect. 4 (2010) 10–18, http://dx.doi.org/10.1111/j.1750-8606.2009.00110.x.
[30] T.D. Little, Longitudinal Structural Equation Modeling, Guilford Press, 2013.
[31] B. Van Den Berg, L. Grievink, J. Yzermans, E. Lebret, Medically unexplained physical symptoms in the aftermath of disasters, Epidemiol. Rev. 27 (2005) 92–106, http://dx.doi.org/10.1093/epirev/mxi001.
[32] J.G. Hull, D.A. Lehn, J.C. Tedlie, A general approach to testing multifaceted personality constructs, J. Pers. Soc. Psychol. 61 (1991) 932–945, http://dx.doi.org/10.1037/0022-3514.61.6.932.
[33] K. Beirens, J.R.J. Fontaine, Development of the Ghent multi-dimensional somatic complaints scale, Assessment 17 (2010) 70–80.
[34] F.F. Chen, S. West, K. Sousa, A comparison of bifactor and second-order models of quality of life, Multivar. Behav. Res. 41 (2006) 189–225, http://dx.doi.org/10.1207/s15327906mbr4102_5.
[35] S.P. Reise, T.M. Moore, M.G. Haviland, Bifactor models and rotations: exploring the extent to which multidimensional data yield univocal scale scores, J. Pers. Assess. 92 (2010) 544–559, http://dx.doi.org/10.1080/00223891.2010.496477.
[36] R.J. Vandenberg, C.E. Lance, A review and synthesis of the measurement invariance literature: suggestions, practices, and recommendations for organizational research, Organ. Res. Methods 3 (2000) 4–70, http://dx.doi.org/10.1177/109442810031002.