The health of Australian veterans of the 1991 Gulf War: factor analysis of self-reported symptoms

A B Forbes, D P McKenzie, A J Mackinnon, H L Kelsall, A C McFarlane, J F Ikin, D C Glass and M R Sim

Occup. Environ. Med. 2004;61;1014-1020
doi:10.1136/oem.2003.011791

ORIGINAL ARTICLE

The health of Australian veterans of the 1991 Gulf War: factor analysis of self-reported symptoms

A B Forbes, D P McKenzie, A J Mackinnon, H L Kelsall, A C McFarlane, J F Ikin, D C Glass, M R Sim

Occup Environ Med 2004;**61**:1014–1020. doi: 10.1136/oem.2003.011791

See end of article for authors' affiliations

Correspondence to: Associate Professor Andrew Forbes, Department of Epidemiology and Preventive Medicine, Central and Eastern Clinical School, Monash University, The Alfred Hospital, Commercial Road, Melbourne, VIC 3004, Australia; andrew.forbes@med.monash.edu.au

Accepted 13 June 2004

Background: A recent report showed that Australian veterans of the 1991 Gulf War displayed a greater prevalence of a multitude of self-reported symptoms than a randomly sampled comparison group of military personnel who were eligible for deployment but were not deployed to the Gulf.

Aims: To investigate whether the pattern, rather than frequency, of symptom reporting in these Australian Gulf War veterans differed from that of the comparison group personnel.

Methods: Factor analysis was used to determine whether the co-occurrence of 62 symptoms in 1322 male Gulf War veterans can be explained by a number of underlying dimensions, called factors. The methodology was also applied to 1459 male comparison group subjects and the factor solutions of the two groups were compared.

Results: For the Gulf War veterans, a three factor solution displayed replicability and construct validity. The three factors were labelled as psycho-physiological distress, somatic distress, and arthro-neuromuscular distress, and were broadly similar to those described in previous studies of Gulf War veterans. A concordant three factor solution was also found for the comparison group subjects, with strong convergence of the factor loadings and factor scores across the two groups being displayed.

Conclusion: Results did not display evidence of a unique pattern of self-reported symptoms among Gulf War veterans. Results also indicated that the differences between the groups lie in the degrees of expression of the three underlying factors, consistent with the well documented evidence of increased self-reported symptom prevalence in Gulf War veterans.

In the 1991 Gulf War, Australia deployed 1871 troops to the Gulf area as part of a larger multinational response to the invasion of Kuwait by Iraq. As reported in our companion paper,[1] we found that self-reported symptoms were more commonly reported in Australian veterans of the 1991 Gulf War than in a randomly sampled comparison group of military personnel who did not deploy to that conflict. Our findings were consistent with the results of several other studies of Gulf War veterans from the USA, UK, Denmark, and Canada (full references are provided in the companion paper[1]). In addition to this finding of increased prevalence, it is possible that the pattern of the self-reported symptoms, as reflected in their co-occurrence, is different between the two groups. In order to investigate whether such patterns differed between the Gulf War veterans and the comparison group, we used a technique known as factor analysis—the methodology and results of that analysis are the focus of this paper.

Developed by psychologists,[2] factor analysis was originally applied to the study of intelligence and personality,[3 4] although an early medical application is described by Andrews.[5] Factor analysis may be defined as a ''set of statistical methods for analysing the correlations among several variables in order to estimate the number of fundamental dimensions (factors) that underlie the observed data and to describe and measure those dimensions''.[6] In the context of the present study, factor analysis attempts to determine whether the co-occurrence of self-reported symptoms can be explained by a number of underlying and unobserved dimensions, known as factors, and whether the composition of such factors differ between the Gulf War veteran and comparison groups.

Several studies of 1991 Gulf War veterans have employed factor analysis in order to ascertain whether there are collections of symptoms specific to Gulf War veterans. Most of these studies have included a military comparison group whose members did not deploy to the Gulf,[7–12] although four were confined to Gulf War veterans only.[13–16] Of the studies that included comparison groups, all except one found no important differences in the pattern of self-reported symptoms between the study groups. These studies did not provide evidence to support the existence of a unique syndrome among Gulf War veterans beyond the possibility that Gulf War veterans may lie at different locations on the same underlying dimensions, when compared to members of the comparison group. In other words, the veterans may differ in the frequency of symptoms reported, but not in the patterns of symptoms. In regard to the only comparison group study that did find a difference,[9] this result was restricted to the composition of a single factor in a six factor solution. No other differences were found. It is interesting to note that the authors' interpretation and conclusions were questioned by the journal editor, in a footnote to the published article.[9] The recent report by Shapiro and colleagues[12] provides a summary of the methodology, results, and factors identified in a majority of the previous studies.

Conventional factor analysis may give misleading results if it is applied to the type of scales commonly found with questionnaires. Of the ten 1991 Gulf War factor analytic studies cited above, nine employed symptom questionnaires that involved items being scored using binary (''absent'', ''present'') or ordinal response scales. The latter type of scale may be defined as one in which the points are ordered (for example, according to severity: ''none'', ''mild'', ''moderate'',

Abbreviations: SF-12, 12-item version of the Short-Form Health Survey; PCS-12, Physical Component Summary scale of the SF-12; MCS-12, Mental Component Summary scale of the SF-12; TLI, Tucker-Lewis Index; CFI, comparative fit index; RMSEA, root mean squared error of approximation; SRMR, standardised root mean squared residual

Main messages

- Using factor analysis, three identifiable factors were extracted to represent the pattern of symptom reporting in Australian 1991 Gulf War veterans.
- A similar pattern of symptom reporting was found in a comparison group of non-deployed military personnel of that era, thereby not supporting the existence of a pattern of self-reported symptoms unique to Gulf War veterans.
- Despite the above similarity in the *pattern* of symptom reporting, the Gulf War veterans displayed a markedly more extreme *degree of expression* of the three underlying factors.

''severe''), but not necessarily equally spaced and not assigned numerical scores. None of the nine studies' published results used factor analytic techniques specifically intended for binary or ordinal scales, such as those discussed by Muthen,[17] although one study[11] did perform these techniques as a check on their published results and arrived at essentially the same findings (Ismail *et al*, personal communication May 2002).

As factor analysis may be seen as the regression of responses to items on one or more unobserved (latent) variables,[18] the problems arising from employing binary or ordinal dependent variables are readily apparent and mirror those associated with regression with observed predictors. Although the practice is not uncommon, the use of Pearson product-moment correlations calculated on binary or ordinal data can diminish the ability to detect the latent structure of such variables. Two responses may reflect the same latent construct, but have differing response distributions due to their location on this construct. One response may reflect a low or mild level of the construct while the other response may reflect a higher or more severe form. In this situation, a Pearson correlation will underestimate the true association of the two variables and its use in factor analysis may result in what have been referred to in the literature as ''difficulty factors''.[19]

The present paper investigates whether there are any differences in the co-occurrence or pattern of self-reported symptoms between Australian 1991 Gulf War veterans and a randomly sampled military comparison group who were not deployed to the Gulf. Contemporary methods of factor analysis, specifically intended for ordinally scaled data, are employed.

METHODS

The study was approved by the Standing Committee on Ethics in Research Involving Humans at Monash University, the Australian Government Department of Veterans' Affairs Human Research Ethics Committee, and the Australian Defence Human Research Ethics Committee.

Study population

Details of the study population, sampling scheme, and symptom questionnaire have been described in detail elsewhere,[1] therefore we present only the main aspects here. The study population was the entire cohort of 1871 Australian Gulf War veterans. A comparison group of 2924 subjects was selected as a stratified random sample from 26 411 Australian Defence Force personnel who were in operational units at the time of the Gulf War, but were not deployed to that conflict. They were frequency matched to the Gulf War group by service type, gender, and three-year age band. Participants, recruited from August 2000 to April 2002, completed a self-administered postal questionnaire that included a 63 item self-report symptom questionnaire which asked about the occurrence of symptoms in the past month and whether the severity of those symptoms was ''mild'', ''moderate'', or ''severe'', and a health assessment. The symptom questionnaire was based on that used by the King's College Gulf War Illness Research Unit,[20] which in turn was based on the Hopkins Symptom Checklist.[21] Due to the very small number of female veterans, all analyses were limited to male subjects. A total of 2970 male subjects completed the symptom questionnaire (1421 Gulf War veterans and 1549 comparison group subjects, representing response rates of 78% and 54%, respectively). The 63 self-reported symptoms used in this study were screened in order to determine whether there were any low prevalences that could make the factor analytic procedures computationally unstable and unreliable. As part of this process one symptom (seizures in past month) was omitted from further consideration because its prevalence was very low, defined as less than 1% (nine persons in total, 0.3%). Any individuals who were missing any of the remaining 62 symptom responses were also excluded, leaving 2781 subjects (1322 Gulf War, 1459 comparison group) for analysis.

Statistical methods

As the response scale for the 62 self-reported symptom questions to be used in the factor analysis was of an ordinal (''none'', ''mild'', ''moderate'', ''severe'') nature, factor analytic methods appropriate for this type of data,[17] involving polychoric correlations,[18] were applied using Mplus statistical software (version 2.1).[22]

Factor analysis of polychoric correlations

In brief, although the ordinal response scale for each symptom has only a small number of discrete points, the underling factor model[23] assumes that each symptom varies on an underlying continuous scale and each person can be located on that continuum. Persons located above a certain threshold on that continuum will endorse the symptom as being present. When considering multiple response categories—for example, ''none'', ''mild'', ''moderate'', ''severe''—there is an additional threshold for each transition between categories. Each underlying continuum is assumed to follow a Gaussian (that is, ''normal'') distribution. Polychoric correlation coefficients are the bivariate correlations between two such underlying continua,[18] and are computed by an iterative procedure[24] using the observed ordinal responses for each pair of individual symptoms. Recent research with binary responses[25] has indicated the general robustness of polychoric correlations to skewed non-normal distributions of the underlying continua.

The computation of polychoric correlation coefficients for pairs of symptoms may encounter difficulties if no individual has that particular combination of reported symptoms. To reduce this possibility, cross-tabulations of all pairs of symptoms were examined and individual symptom categories were combined where theoretically meaningful to do so, giving either three point or two point scales. In such a fashion expected cell frequencies of at least one in all cells of the resulting tables could be ensured. This resulted in 28 self-reported symptom items being recoded to three categories and 25 items being recoded to two categories, with nine items retaining their original four point scale.

Exploratory factor analysis was applied to the polychoric correlation matrix and factor solutions extracted using diagonally weighted least squares, with a robust estimate of the covariance matrix of the parameter estimates and a mean and variance corrected χ^2 statistic. This was implemented using the WLSMV estimation method in the MPlus software.[22]

Methods of factor rotation

A crucial step in any factor analysis is the determination of the number of factors to retain for rotation, in order that an interpretable solution may be found.[23] In our study, this step involved examination of the eigenvalues (which can be thought of as the amount of variation accounted for by each factor) of the polychoric correlation matrix, the percentage of the total variance explained by each possible number of factors and the associated scree plot (defined below), the reproducibility of the factors, and the clinical meaningfulness of the factors extracted. Two forms of factor rotation were applied to each factor solution:[23 26] Varimax, which produces orthogonal or uncorrelated factors, and Promax (with power constant = 2), which allows factors to be oblique, or correlated with each other. An arbitrary but conventional threshold of 0.40 for the factor loadings was applied when interpreting and providing brief descriptive labels for the factors.

Method of assessing factor reproducibility

The reproducibility of the factor solution is highly important, as extracting more than the required number of factors may lead to extraction of ''noise'' that is specific to a particular sample. The resulting factor structure may therefore not be replicable in other samples, or even in samples from the same population.[27] To assess the reproducibility of our obtained factor solutions, the Gulf War veteran group was randomly split in half. Subsequently two, three, and four factor solutions were obtained for each half, and the congruence of the solutions for the two halves were assessed via Pearson product-moment and one-way random-effects intra-class correlation coefficients[28] of the factor loadings, with the latter providing a more direct measure of reproducibility.[29] As the two types of correlation coefficient produced almost identical results, only the Pearson correlations are reported.

Method of factor validation

Factor scores were computed for each individual for each of the obtained factors using the iterative factor scoring procedure as implemented in the Mplus software.[22] These scores estimate quantitatively where each individual lies on each of the dimensions tapped by the factors. The construct validity[30] of the obtained solution was assessed by correlating the factor scores with the Physical and Mental Component Summary scales (PCS-12 and MCS-12) of the 12 item version of the Short-Form Health Survey (SF-12).[31] Internal consistency of each of the factors was assessed using Cronbach's coefficient alpha,[30 32] based on unit weighting (item weights set to one) for the items with loadings greater than 0.40 in each factor.

Two different methods were employed to assess whether a similar factor structure existed in the comparison group. The principal method employed an exploratory factor analysis of the comparison group using the same methods and rotation as described above for the Gulf War veterans. Congruence of the factor loadings for each factor between the Gulf War and comparison group was assessed using intra-class correlation coefficients. Factor scores were computed for each of the factors for the comparison group subjects using both the factor solution from the comparison group, and also the factor solution obtained from the Gulf War group. These scores were then correlated with each other, with a high correlation providing evidence of similarity of the factor structure underlying the symptom patterns observed in the two groups. A similar method has been employed in a previous study of Gulf War veterans.[8]

The second method of investigating factor structure similarity involved a more formal assessment of the invariance of the obtained solution across the Gulf War and comparison groups.[33] This was performed by implementing a two-group structural equation model[23] and assessing its goodness of fit. This model for each group allowed the factors to be correlated with each other, loaded all symptoms on all factors, and constrained the loadings, covariance structure, scale, and threshold parameters so as to be equal across groups. This constrained model also enabled the difference in the mean *amount* of each of the obtained factors between the two groups to be computed. Differences in means were rescaled into units of standard deviations (that is, effect sizes[34]) of the factors. For comparison purposes, an additional model was fit which relaxed the constraints of equality of factor loadings across groups. The fit of these models to the symptom data was assessed with conventional goodness of fit indices, together with recent guidelines for their interpretation.[35]

RESULTS

Gulf War veteran group

Exploratory factor analysis of the 1322 Gulf War veterans with complete data on the 62 symptoms yielded 41.1% of the total variance accounted for by the first factor. The second through fifth factors contributed incremental amounts of 3.3%, 2.7%, 1.9%, and 1.7% respectively, yielding 47.1% of the variance explained by three factors, and 50.7% of the variance explained by five factors. Accordingly, examination of the ''scree plot'',[36] a plot with eigenvalues on the vertical axis and factor numbers (for example, the first factor extracted, second, third, fourth, and so on) on the horizontal axis, showed a very large or dominant first factor with possible minor contributions from a further two factors. The remaining factors appeared to be ''scree'', a geological term denoting debris or litter. The above analyses indicated that additional factors beyond three contributed very little to explaining the variance of the patterns of symptom reporting. There were 13 factors with eigenvalues greater than one and hence met Kaiser's criterion,[23] which is commonly used but typically overestimates the required number of factors in situations with large numbers of items.[36]

Varimax and Promax rotations were applied to the two, three, and four factor solutions. The Promax solutions provided more distinct and interpretable factor solutions than did the orthogonal Varimax rotations, and indicated that the underlying factors were moderately correlated. Interfactor correlations for the three factor solution were 0.52, 0.47, and 0.44 between factors 1 and 2, 1 and 3, and 2 and 3, respectively, meaning that persons with a greater than average amount of one factor tended to have greater than average amounts of the other factors.

Factor reproducibility

The assessment of the reproducibility of two, three, and four factor Promax solutions using the two randomly split halves of the Gulf War veteran data indicated that the two and three factor solutions each displayed the same symptoms dominating the loadings on each factor in each split-half. The Pearson product moment correlations between the symptom loadings of the factors across the two split-halves for the two factor solution were 0.78 and 0.86. For the three factor solution these correlations were 0.78, 0.85, and 0.66. The four factor solution displayed reproducibility for only two of its factors (correlations 0.71 and 0.73), with the remaining two factors not being defined consistently by dominant symptoms across split-halves (correlations ranging from −0.19 to 0.57 across the four possible pairings of the remaining two factors in each split-half). As the three factor solution provided interpretable factors in addition to its amount of explained variance and reproducibility, it was adopted as the best representation of the symptom pattern in the Gulf War veterans.

The factor loadings in the Gulf War veteran sample ($n = 1322$) are displayed in the central panel of table 1. The ordering of symptoms in the table is determined by the size of the coefficients within each of the three factors, and coefficients greater than 0.40 are presented in bold type.

To assist with factor interpretation, brief descriptive labels are suggested, based on the patterns of loadings within each factor. These labels are not intended to reflect clinical diagnoses. The first factor involves symptoms of psychophysiological distress, is labelled as such, and explains 23.7% of the total variance. The second factor involves adverse cognitive symptoms and is labelled "cognitive distress"; it explains 25.2% of the variance. The third factor involves arthritic, muscular, and neurological symptoms and is labelled "arthro-neuromuscular" distress; it explains 10.8% of the variance. Due to the correlation between the factors, there is a total overlap of 12.6% in explained variance, divided into an overlap of 7.1% between factors 1 and 2, 2.7% between 1 and 3, and 2.8% between factors 2 and 3.

Factor validity and reliability

The construct validity of the three factors was supported by computing the Pearson product-moment correlation between the factor scores computed from the three factor model and the physical and mental health summary subscale scores of the SF-12. Table 2 presents the resulting correlations.

All correlations were negative, moderate to large in magnitude, and highly statistically significant ($p < 0.0001$). The negative correlation indicates that the higher the location of each factor a person possesses, the lower the SF-12 score, indicating poorer health. Factor 1 has both physiological and psychological components and accordingly correlates moderately with both the physical and mental SF-12 subscales. Factor 2, cognitive distress, correlates more highly with the SF-12 mental subscale than the physical subscale, while for factor 3, arthro-neuromuscular distress, the reverse is the case.

Assessment of internal consistency, with unit weights applied to items with loadings greater than 0.40 within each factor, produced Cronbach's coefficient alphas of 0.85, 0.93, and 0.82 for factors 1, 2, and 3 respectively, indicating satisfactory internal consistency. The coefficient alphas could not be improved by deleting any item from any factor.

Comparison group

An exploratory factor analysis was carried out in the comparison group subjects using the same methodology as for the Gulf War veterans in order to assess the similarity or otherwise of the underlying factor structure. The three factor Promax solution displayed very substantial concordance with that of the Gulf War veterans, and is presented in the right hand column of table 1. The inter-factor correlations for the comparison group were 0.48, 0.43, and 0.49 between factors 1 and 2, 1 and 3, and 2 and 3, respectively. The eigenvalues and scree plot for the comparison group subjects essentially reproduced those found for the Gulf War veteran data. The intra-class correlations between the factor loadings of each factor between the two groups were 0.94, 0.97, and 0.96 for factors 1, 2, and 3 respectively, indicating substantial similarity of factor structure. The factor scores computed for the comparison group subjects were likewise extremely similar to the scores for these subjects computed using the Gulf War factor analysis solution, with intra-class correlations of 0.98, 0.98, and 0.99 for factors 1, 2, and 3 respectively.

The assessment of the adequacy of the fit of the two group structural equation models used the following indices (with abbreviations and recommended thresholds for adequate fit[35]): Tucker-Lewis Index (TLI, >0.95), comparative fit index (CFI, >0.95), root mean squared error of approximation (RMSEA, <0.06), and standardised root mean squared residual (SRMR, <0.08). The first three indices are overall measures of model fit which penalise the model in various ways for the number of parameters being estimated. The fourth index (SRMR) compares observed and estimated polychoric correlations without any penalty for the number of parameters estimated, and as such will always decrease in value as model complexity increases. The results for the fully constrained model displayed an adequate fit (TLI = 0.98, CFI = 0.98, RMSEA = 0.033, SRMR = 0.071). The model with factor loadings not constrained to be equal across groups displayed an inferior fit (TLI = 0.98, CFI = 0.92, RMSEA = 0.036, SRMR = 0.062) due to the unnecessary estimation of separate loadings for each group when a common loading would have sufficed. For the constrained model, the estimated differences in the mean amount of each of the three factors between the Gulf War and comparison groups, expressed in units of effect size, were 0.32 (psychophysiological), 0.40 (cognitive), and 0.16 (arthro-neuromuscular). The first two effect sizes are between a medium (0.50) and a small (0.20) effect size as defined by Cohen.[34] For all three factors, the Gulf War veterans displayed higher scores (more extreme levels of each factor) than the comparison group, and these differences were statistically significant ($p < 0.003$).

DISCUSSION

In our study involving the factor analysis of symptoms reported by Australian Gulf War veterans, a reproducible factor solution with three moderately correlated factors was identified for the latent structure underlying the pattern of symptom reporting. A similar three factor solution was found for the comparison group subjects. This suggests that the underlying factors affecting patterns of symptom reporting in the Gulf War veterans are unlikely to be different from those of the comparison group subjects.

Our finding does not support the existence of a unique pattern of symptoms affecting Gulf War veterans only, and is consistent with the results of five of the six previous studies of Gulf War veterans in which a comparison group was used in the research design.[7 8 10–12] The other study[9] reported similar three, four, and five factor solutions between Gulf War and comparison groups, but reported a difference in the composition of one factor in their six factor solution—an interpretation that has been disputed.[9]

We have labelled the three factors identified in the Gulf War and comparison group subjects in this study as psychophysiological distress, cognitive distress, and arthro-neuromuscular distress. These factors display some similarity with those from each of the previous ten studies (six with a comparison group,[7–12] and four without a comparison group[13–16]). Exact comparisons involving these studies are difficult because of the use of different symptom questionnaires. Of the three factors in our results, the cognitive distress factor is the most reproducible across studies, with nine of the ten studies producing a similar factor. Most studies have also produced either a musculoskeletal factor or a neurological factor, with three studies[13 15 16] also reporting a combination factor similar to that found in our study. Other studies[10 13] have reported various somatic factors similar to components of the psychophysiological distress factor described here. Interestingly, the similarity of our results

Table 1 Factor loadings (pattern coefficients) × 100 among the Gulf War veterans (n = 1322) and comparison group subjects (n = 1459)

Symptom	Gulf War veterans			Comparison group		
	Factor 1 (Psy-phys)*	Factor 2 (Cogn)†	Factor 3 (A-N-M) ‡	Factor 1 (Psy-phys)*	Factor 2 (Cogn)†	Factor 3 (A-N-M) ‡
Vomiting	**80**	−11	−10	**81**	−8	11
Nausea	**79**	3	−1	**67**	19	−4
Stomach cramps	**70**	1	1	**51**	11	11
Diarrhoea	**64**	3	−6	**58**	7	−2
Wheezing	**61**	4	1	**61**	−4	2
Indigestion	**60**	−4	13	**40**	12	14
Shortness of breath	**58**	14	12	**46**	18	15
Dry mouth	**54**	19	12	**48**	22	1
Feeling feverish	**54**	17	15	**67**	13	8
Swelling of lymph glands	**49**	3	16	**51**	2	15
Lump in throat	**49**	9	11	**56**	−11	10
Persistent cough	**49**	17	−3	**61**	−9	0
Pain on passing urine	**47**	24	1	31	29	9
Constipation	**45**	16	9	21	21	12
Difficulty speaking	**44**	**46**	2	27	**49**	−2
Dizziness, fainting, or blackouts	**44**	30	7	39	35	8
Loss of balance or coordination	**44**	35	19	39	**43**	14
Sore throat	**43**	−1	9	**51**	−4	−1
Flatulence/burping	**43**	12	14	31	14	16
Loss control bladder/bowels	**43**	26	−6	38	32	21
Burning in sex organs	**41**	26	8	31	**40**	15
Skin ulcers	**40**	14	9	**51**	17	6
Loss of, or decrease in, appetite	**40**	**42**	−2	46	37	−3
Loss of concentration	10	**80**	5	12	**79**	3
Feeling distant from others	11	**78**	−3	8	**81**	−5
Unrefreshed after sleep	−5	**74**	28	−17	**79**	25
Forgetfulness	9	**73**	4	15	**71**	−2
Loss of interest in sex	9	**69**	−2	12	**55**	13
Sleeping difficulties	3	**69**	16	−12	**71**	24
Avoiding things/situations	20	**69**	5	9	**72**	9
Feeling jumpy or easily startled	26	**65**	2	13	**65**	6
Problems sexual functioning	7	**65**	−1	15	**51**	20
Distressing dreams	15	**65**	7	19	**61**	3
Fatigue	1	**65**	30	−4	**68**	22
Irritability or outbursts of anger	14	**64**	5	8	**63**	9
Difficulty finding right word	17	**59**	9	19	**64**	−2
Feeling disorientated	33	**58**	5	38	**59**	−4
Increased sensitivity to noise	22	**52**	5	32	**42**	13
Shaking	34	**46**	4	30	39	11
Increased sensitivity to light	29	**43**	11	40	31	9
Increased sensitivity to smells/odours	17	**41**	14	24	**42**	9
Stiffness in several joints	6	5	**88**	6	−1	**91**
Pain in several joints (no swelling or redness)	4	10	**81**	2	2	**84**
General muscle aches/pains	19	4	**67**	9	14	**60**
Loss sensation hands/feet	22	18	**48**	27	11	**52**
Low back pain	15	11	**45**	3	13	**56**
Tingling/burning in hands/feet	24	19	**43**	26	15	**43**
Ringing in the ears	24	22	20	14	21	28
Rash or skin irritation	22	21	19	26	11	18
Itchy or painful eyes	26	24	18	28	19	20
Night sweats that soak sheets	23	36	17	34	33	19
Unintended weight gain >4 kg	23	22	17	15	31	16
Chest pain	35	19	17	35	18	23
Toothache	37	1	15	26	12	10
Skin infections e.g. boils	25	15	14	24	9	25
Mouth ulcers	23	7	13	21	2	13
Rapid or pounding heart beat	37	33	13	37	33	5
Double vision	27	38	12	34	34	1
Intolerance to alcohol	32	37	10	26	**41**	6
Headaches	30	26	9	22	27	16
Passing urine more often	32	32	9	35	25	14
Unintended weight loss >4 kg	34	32	0	30	40	−17

Coefficients greater than 0.40 are presented in bold type.
*The label given to factor 1 is ''psychophysiological distress''.
†The label given to factor 2 is ''cognitive distress''.
‡The label given to factor 3 is ''arthro-neuromuscular distress''.

with those of other studies has occurred despite methodological differences between studies in factor analytic estimation methods employed, factor loading threshold values, response scales, sample sizes, and participation rates.

The correlation between these three factors and the mental and physical subscales of the SF-12 gives some insight into their characteristics. The cognitive distress factor (factor 2) most clearly differentiated between these two dimensions. The items in the second factor are similar to many of the symptoms of post-traumatic stress disorder,[37] and a body of work has now shown the abnormalities in working memory or other cognitive processes, associated with this disorder.[38]

Table 2 Pearson correlations between SF-12 mental and physical subscale scores and Gulf War veteran symptom factor scores

Scale	Factor 1 (psychophysiological)	Factor 2 (cognitive)	Factor 3 (arthro-neuromuscular)
SF-12 mental	−0.44	−0.73	−0.31
SF-12 physical	−0.47	−0.37	−0.55

The Pearson correlation between SF-12 mental and physical subscale scores was 0.09.

The third factor, arthro-neuromuscular distress, had the strongest correlation with the SF-12 physical subscale, suggesting that this component was mostly accounted for by somatic processes. In contrast, the first factor, involving items including gastrointestinal, respiratory, and genitourinary symptoms, appeared to be equally attributable to physical and psychological pathophysiological elements. Hence the three factor solution suggests that the health outcomes of both the Gulf War and comparison group subjects are best accounted for by a complex amalgam of somatic and psychological processes. It would seem unlikely that there exists a simple agent specific to Gulf War veterans that could explain such intertwining of the somatic and psychological symptoms.

The analysis of self-reported symptoms in our companion paper[1] indicated increased prevalences for all 62 symptoms among Gulf War veterans relative to comparison group subjects. The constrained two group structural equation model in the present paper identified three factors underlying these symptoms, with the Gulf War veterans possessing statistically significant increased amounts of each of these three factors relative to the comparison group. These two findings are consistent, as according to the factor analysis model the elevated amounts of each factor are postulated to explain the increased symptom prevalences in Gulf War veterans.

As mentioned earlier, nine of the ten previous Gulf War factor analytic studies employed symptom questionnaires using an ordinal or binary response scale for each symptom item, but used conventional methods of factor analysis not specifically intended for such scales. Performing an exploratory factor analysis on the Gulf War veteran data in the present study, and treating the ordinal data as if they were interval, the use of Pearson correlations produced attenuated correlation coefficients, as might be expected.[39] Despite this, three broadly similar factors to those presented here emerged. However, the explained variance from the three factors was reduced (28%), as were the sizes of the factor loadings. It is therefore plausible that other studies using such methods may also have attenuated loadings, and a clearer factor structure may have emerged had methods for ordinal data been applied.

As argued above, we believe that a three factor solution extracts the maximum number of dimensions that can be identified reliably in the data. However, we also computed four, five, and six factor solutions in order to assess whether our data support the existence of the cluster of four symptoms (blurred vision, loss of balance/dizziness, speech difficulty, and tremors/shaking) reported by Kang and colleagues[9] as being present in their sample of Gulf War veterans but not in their comparison group. Using a loading threshold of 0.30 adopted by Kang and colleagues,[9] for each of the four, five, and six factor solutions, our analyses indicated that all four symptoms loaded on one of the factors in our sample of Gulf War veterans (with some cross-loading on other factors). However, unlike Kang and colleagues,[9] all symptoms also correspondingly loaded on a single factor in the comparison group for each of these solutions. Therefore our data do not provide support for this cluster of four symptoms as being unique to Gulf War veterans.

Any factor analysis involves a number of analytic decisions (for example, the number of factors to retain), as detailed above. Potential limitations of the analysis outlined in the present paper include non-participation bias among the comparison group subjects, information bias among Gulf War veterans, and non-randomly missing symptom data. It is possible, although purely speculative, that the pattern of symptom reporting among comparison group non-participants differs from that of participants, in which case differences in the underlying factor structures of Gulf War veterans and comparison group subjects may arise. It is also possible that the media coverage given to the health problems of Gulf War veterans may have influenced their patterns of symptom reporting.[40] However, our finding that the comparison group exhibited a very similar pattern of reporting does not support this possibility. It appears more likely that any information bias may only be affecting the frequency and severity rather than the pattern of symptom reporting. The percentage of subjects excluded from the factor analysis due to missing symptom item responses was low (7.0% Gulf War, 5.8% comparison group). Although individuals with missing symptom data tended to have a higher rate of symptom reporting than individuals with complete data, this occurred in both groups. This latter finding, together with the overall low percentage of missing data, indicates that missing data is unlikely to have influenced our results in any more than a minor way.

While symptom patterns may be found to be similar in two populations, different underlying illnesses may produce these similar symptom patterns (for example, influenza and the common cold both resulting in runny nose and fever). So, as is pointed out by Ferguson,[41] the presence or absence of separate illnesses with different causative agents in two different populations cannot be determined solely from the factor analysis of self-reported symptoms. Although we cannot rule out such a phenomenon in this study, we do not consider it very plausible given the large number of symptoms incorporated into the factor analysis model and the breadth of the symptomatology covered.

Evidence against the existence of a unique symptom complex among Gulf War veterans is also provided by the results of cluster analytic studies[7,42] involving Gulf War veterans. Whereas factor analysis is concerned with the similarity and co-occurrence of symptoms themselves, cluster analysis is generally concerned with the grouping of *individuals* into subgroups or clusters, based on similar profiles or patterns of symptoms.[42,43] Six clusters were found by Cherry and colleagues,[7] and five by Everitt and colleagues;[42] however there were no substantially sized clusters that were predominantly or exclusively made up of Gulf War veterans, suggesting that there was not a clinical profile of symptoms that was unique to Gulf War veterans.

In conclusion, using factor analytic methods devised specifically for ordinal scaled data, we found that the pattern of self-reported symptoms did not differ between 1991 male Australian Gulf War veterans and a military comparison group not deployed to the Gulf. This finding casts doubt on the existence of a symptom complex that is unique to Australian Gulf War veterans, and accords with the results of the majority of previous studies using other factor analytic methods. However, our results also indicated that Gulf War veterans displayed a more extreme degree of expression of three underlying dimensions relative to the comparison group, consistent with the vast body of accumulated evidence from many studies of the increased self-reported symptom prevalence among Gulf War veterans.

ACKNOWLEDGEMENTS

The Australian Gulf War Veterans' Health Study was funded by the Australian Government–Department of Veterans' Affairs. The study was overseen by a Scientific Advisory Committee, headed by Professor Terry Dwyer, and by a Veterans' Consultative Forum. We are grateful to members of both groups for their contribution and support, and in particular to Dr Leigh Blizzard for his valuable comments. We gratefully acknowledge the contribution of Dr Keith Horsley, Dr Warren Harrex, Mr Bob Connolly, and his contact and recruitment team at the Australian Government–Department of Veterans' Affairs. Lastly, but most importantly, we sincerely thank the Gulf War veterans and comparison group members for their participation.

Authors' affiliations

A B Forbes, D P McKenzie, H L Kelsall, J F Ikin, D C Glass, M R Sim, Department of Epidemiology and Preventive Medicine, Monash University, Melbourne, Victoria, Australia
A J Mackinnon, Department of Psychological Medicine, Monash University, and Biostatistics and Psychometrics Unit, Mental Health Research Institute, Melbourne, Victoria, Australia
A C McFarlane, Department of Psychiatry, University of Adelaide, South Australia, Australia

Funding: The study was financially supported by the Australian Government–Department of Veterans' Affairs.

REFERENCES

1 **Kelsall HL**, Sim MR, Forbes AB, *et al.* Symptoms and medical conditions in Australian veterans of the 1991 Gulf War: relation to immunisations and other Gulf War exposures. *Occup Environ Med* 2004;**61**:1006–13.
2 **Spearman C**. General intelligence objectively determined and measured. *Am J Psychol* 1904;**15**:202–93.
3 **Cattell RB**. The integration of factor analysis with psychology. *J Educ Psychol* 1948;**39**:227–36.
4 **Eysenck HJ**. *The dimensions of personality.* London: Routledge and Kegan Paul, 1947.
5 **Andrews T**. Statistical studies in allergy. II. A factorial analysis. *J Allergy* 1948;**19**:43–6.
6 **Last JM**. *A dictionary of epidemiology*, 3rd edn. New York: Oxford University Press, 1995.
7 **Cherry N**, Creed F, Silman A, *et al.* Health and exposures of United Kingdom Gulf War veterans. Part I: The pattern and extent of ill health. *Occup Environ Med* 2001;**58**:291–8.
8 **Doebbeling BN**, Clarke WR, Watson D, *et al.* Is there a Persian Gulf War syndrome? Evidence from a large population-based survey of veterans and nondeployed controls. *Am J Med* 2000;**108**:1–11.
9 **Kang HK**, Mahan CM, Murphy FM, *et al.* Evidence for a deployment-related Gulf War syndrome by factor analysis. *Arch Environ Health* 2002;**57**:61–8.
10 **Knoke JD**, Smith TC, Gray GC, *et al.* Factor analysis of self-reported symptoms: does it identify a Gulf War syndrome. *Am J Epidemiol* 2000;**152**:379–88.
11 **Ismail K**, Everitt B, Blatchley N, *et al.* Is there a Gulf War syndrome? *Lancet* 1999;**353**:179–82.
12 **Shapiro SE**, Lasarev MR, McCauley L. Factor analysis of Gulf War illness: what does it add to our understanding of possible health effects of deployment? *Am J Epidemiol* 2002;**156**:578–85.
13 **Bourdette DN**, McCauley LA, Barkhuizen A, *et al.* Symptom factor analysis, clinical findings and functional status in a population-based case control study of Gulf War unexplained illness. *J Occup Environ Med* 2001;**43**:1026–40.
14 **Fukuda K**, Nisenbaum R, Stewart G, *et al.* Chronic multisymptom illness affecting Air Force veterans of the Gulf War. *JAMA* 1998;**280**:981–8.
15 **Haley RW**, Kurt TL, Hom J. Is there a Gulf War syndrome? Searching for syndromes by factor analysis of symptoms. *JAMA* 1997;**277**:215–22.
16 **Hallman WK**, Kipen HM, Diefenbach M, *et al.* Symptom patterns among Gulf War registry veterans. *Am J Public Health* 2003;**93**:624–30.
17 **Muthen B**. A general structural equation model with dichotomous, ordered categorical, and continuous latent variable indicators. *Psychometrika* 1984;**49**:115–32.
18 **Bartholomew DJ**, Knott M. *Latent variable models and factor analysis.* London: Arnold, 1999.
19 **Christoffersen A**. Factor analysis of dichotomized variables. *Psychometrika* 1975;**40**:5–22.
20 **Unwin C**, Blatchley N, Coker W, *et al.* Health of UK servicemen who served in Persian Gulf War. *Lancet* 1999;**353**:169–78.
21 **Derogatis LR**, Lipman RS, Rickels K, *et al.* The Hopkins Symptom Checklist (HSL): a self-report symptom inventory. *Behav Sci* 1974;**19**:1–15.
22 **Muthen B**, Muthen B. *Mplus: the comprehensive modeling program for applied researchers: user's guide. Version 2.1.* Los Angeles: Muthen & Muthen, 2002.
23 **Kline RB**. *Principles and practice of structural equation modeling.* New York: Guilford Press, 1998.
24 **Olsson U**. Maximum likelihood estimation of the polychoric correlation coefficient. *Psychometrika* 1979;**44**:443–60.
25 **Greer T**, Dunlap WP, Beatty GO. A Monte Carlo evaluation of the tetrachoric correlation coefficient. *Educational and Psychological Measurement* 2003;**63**:931–50.
26 **Rencher AC**. *Multivariate statistical inference and applications.* New York: Wiley, 1998.
27 **Floyd FJ**, Widaman KF. Factor analysis in the development and refinement of clinical assessment instruments. *Psychological Assessment* 1995;**7**:286–99.
28 **Shrout PE**, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. *Psychol Bull* 1979;**86**:420–8.
29 **Streiner DL**. Learning how to differ: agreement and reliability statistics in psychiatry. *Can J Psychiatry* 1995;**40**:60–6.
30 **Anastasi A**, Urbina S. *Psychological testing*, 7th edn. Upper Saddle River, NJ: Prentice-Hall, 1997.
31 **Ware JE**, Kosinski M, Keller SD. A 12-item Short-Form Health Survey. Construction of scales and preliminary tests of reliability and validity. *Med Care* 1996;**34**:220–33.
32 **Cronbach L**. Coefficient alpha and the internal structure of tests. *Psychometrika* 1951;**16**:297–334.
33 **Vandenberg RJ**, Lance CE. A review and synthesis of the measurement invariance literature: suggestions, practices, and recommendations for organizational research. *Organizational Research Methods* 2000;**3**:4–69.
34 **Cohen J**. *Statistical power analysis for the behavioral sciences*, 2nd edn. Hillsdale, NJ: Erlbaum, 1988.
35 **Hu L**, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Structural Equation Modeling* 1999;**6**:1–55.
36 **Cattell RB**. *The scientific use of factor analysis in behavioral and life sciences.* New York: Plenum Press, 1978.
37 **American psychiatric association**. *diagnostic and Statistical Manual of Mental Disorders (DSM-IV)*, 4th edn. Washington, DC: American Psychiatric Association, 1994.
38 **Clark CR**, McFarlane AC, Morris P, *et al.* Cerebral function in posttraumatic stress disorder during verbal working memory updating: a positron emission tomography study. *Biol Psychiatry* 2003;**53**:474–81.
39 **Waller N**, Tellgen A, McDonald R, *et al.* Exploring nonlinear models in personality assessment: development and preliminary validation of a negative emotionality scale. *J Pers* 1996;**64**:545–76.
40 **Hyams KC**, Scott K. Adding to our understanding of Gulf War health issues. *Psychol Med* 2002;**32**:1335–7.
41 **Ferguson E**. Is there a Gulf War syndrome? *Lancet* 1999;**353**:1182–3.
42 **Everitt B**, Ismail K, David AS, *et al.* Searching for a Gulf War syndrome using cluster analysis. *Psychol Med* 2002;**32**:1371–8.
43 **Everitt BS**, Landau S, Leese M. *Cluster analysis*, 4th edn. London: Arnold, 2001.