Psychological Medicine, 2002, **32**, 1357–1370.
DOI: 10.1017/S0033291702006359 Printed in the United Kingdom

Cognitive functioning and disturbances of mood in UK veterans of the Persian Gulf War: a comparative study

A. S. DAVID,[1] L. FARRIN, L. HULL, C. UNWIN, S. WESSELY AND T. WYKES

From the Gulf War Illnesses Research Unit, Guy's, King's and St Thomas' School of Medicine, King's College and Institute of Psychiatry, London

ABSTRACT

Background. Complaints of poor memory and concentration are common in veterans of the 1991 Persian Gulf War as are other symptoms. Despite a large research effort, such symptoms remain largely unexplained.

Method. A comprehensive battery of neuropsychological tests and rating scales was administered to 341 UK servicemen who were returnees from the Gulf War and peace keeping duties in Bosnia, plus non-deployed military controls. All were drawn from a large randomized survey. Most were selected on the basis of impaired physical functioning defined operationally.

Results. Group comparisons revealed an association between physical functioning and symptoms of depression, post-traumatic stress reactions, increased anger and subjective cognitive failures. Poorer performance on some general cognitive measures, sequencing and attention was also seen in association with being 'ill' but virtually all differences disappeared after adjusting for depressed mood or multiple comparisons. Deployment was also associated with symptoms of post-traumatic stress and subjective cognitive failures, independently of health status, as well as minor general cognitive and constructional impairment. The latter remained significantly poorer in the Gulf group even after adjusting for depressed mood.

Conclusions. Disturbances of mood are more prominent than quantifiable cognitive deficits in Gulf War veterans and probably lead to subjective underestimation of ability. Task performance deficits can themselves be explained by depressed mood although the direction of causality cannot be inferred confidently. Reduced constructional ability cannot be explained in this way and could be an effect of Gulf-specific exposures.

INTRODUCTION

Memory and concentration difficulties and disturbances of mood are among the most commonly reported symptoms associated with service in the 1990–1991 Persian Gulf War. In a comprehensive survey of UK military personnel who served in the Gulf War (GW), irritability and anger was reported by 55%, while forgetfulness and loss of concentration was reported by 50% and 45%, respectively (Unwin *et al.* 1999). Clinical evaluation confirms these reports (Coker *et al.* 1999) and surveys of US (Roy *et al.* 1998) and Danish servicemen (Suadicani *et al.* 1999) have produced similar figures.

Some studies have been published specifically investigating the neuropsychological or cognitive performance of veterans (Goldstein *et al.* 1996; Axelrod & Milner, 1997; Haley *et al.* 1997; Hom *et al.* 1997; Silanpaa *et al.* 1997; Anger *et al.* 1999; Storzbach *et al.* 2000), but few consistent findings have emerged. Hence while there is a well replicated high prevalence of cognitive and emotional disturbances in surveys of GW veterans several years after the conflict (The Iowa Persian Gulf Study Group, 1997; Roy *et al.* 1998; Unwin *et al.* 1999), evidence for

[1] Address for correspondence: Professor A. S. David, Section of Cognitive Neuropsychiatry, Institute of Psychiatry, De Crespigny Park, London SE5 8AF.

cognitive deficits on detailed testing is much less clear (Storzbach *et al.* 2000).

The possible reasons for objective cognitive impairment in this group (of whatever degree) include neurotoxic exposures – as advanced by Haley and others (Haley *et al.* 1997; Hom *et al.* 1997). Exposures such as organophosphates, radiation and infectious agents, could affect cognition directly by action on the central nervous system. Another cause is information processing deficits secondary to depression or stress reactions. Such deficits are well recognized (see e.g. Weingartner *et al.* 1981; Hartlage *et al.* 1993; Brown *et al.* 1994) in clinical groups with depression, and have been reported in GW veterans with post-traumatic stress disorder (PTSD) and depression (Vasterling *et al.* 1998; Wolfe *et al.* 1999). Of course toxic exposures could also cause neuropsychiatric disturbances including depression, which could in turn lead to neurocognitive deficits.

Subjective cognitive difficulties may still arise in the absence of definite or notable 'objective' cognitive deficits. One explanation for this subjective–objective discrepancy is a cognitive bias leading to negative performance evaluation by those veterans with low mood. This could include increased attention to minor slips and errors as well as a tendency to attribute such problems to a physical disease (see Robbins & Kirmayer, 1991). Again similar phenomena have been reported in civilian clinical groups (Cope *et al.* 1995; Gass & Apple 1997; Wagle *et al.* 1999) and symptomatic GW veterans (Binder *et al.* 1999).

In this study, potential participants were recruited by screening a large randomly selected cohort of servicemen and women deployed in the GW, in Bosnia and military personnel in service at the time of the Gulf conflict but deployed to neither theatre (Unwin *et al.* 1999, 2002). Given there is no accepted definition of GW illness we used a case definition based on a standard and well validated self-report index of physical functioning. This was intended to be as free as possible from any assumptions about possible aetiology.

A sample of 'ill' and 'well' GW veterans was then selected along with a comparison group of 'ill' controls, and invited to attend a clinical research centre for assessment. The aims were to determine: first, whether the pattern of neuropsychological deficits, if any, differed across symptomatic veterans from different deployments; secondly, whether such deficits could be attributed to active deployment or to specific deployment to the Gulf; and finally, which deficits were associated with general ill health. We were also interested to determine whether depressed mood would be an important variable in modulating objective and subjective cognitive performance. Because of our interest in patterns of deficits across groups with different deployment histories and hence their specificity, it was decided that recruitment of well people from the Bosnia and Era groups was not essential. The Gulf well group served as unaffected controls that were assumed to have shared at least some of the possible risk factors as their unwell counterparts.

Considerable information was recorded in our earlier survey on possible exposures. However, the reliability of such information is questionable because of recall and reporting biases (Southwick *et al.* 1997). Hence, in this study we use deployment as a crude but reliable proxy for a range of specific and non-specific exposures.

METHOD

Participants

This study forms part of the second phase of an initial epidemiological survey. A random sample of those who had served in either the Gulf or Bosnia in peace keeping operations, or who had been in the military at the time but served in neither theatre (the 'Era' group), were sent a postal questionnaire asking about their military experiences including deployment, exposures, illnesses and symptoms (Unwin *et al.* 1999). We used a generic measure of ill health. In phase 1, the Medical Outcome Study Short-Form 36 Physical Functioning (SF36-PF: Ware *et al.* 1993) subscale was completed by all respondents. This requires participants to rate any limitation they may have in physical activities in a typical day (not at all; a little; a lot) in relation to specific activities which are listed individually. The value at the first decile of the distribution of the SF36-PF subscale in the Era cohort (72·2) was used as the cut-off to define ill health in all three cohorts. The Era cohort was considered the most representative of the military as it comprised 80% of the UK Armed Forces. The rationale was first to prevent a selection bias in defining Gulf cases and non-Gulf controls. Using an alternative case definition based on specific symptoms, for example mental health problems, in Gulf veterans may have led to an erroneous assumption that the nature of ill health in Gulf veterans was the same in non-Gulf veterans. By applying the first decile (the lowest 10% scores) as a cut-off on the SF36-PF, those who were most disabled would be identified – hence we ended up with a functionally meaningful case definition. The number of Gulf, Bosnia and Era cases, and Gulf well controls that were eligible to participate, and from which random samples were drawn, was 406, 138, 278 and 3047, veterans respectively. If, after random selection, the subject reported a current serious physical illness defined as life limiting and requiring active treatment, he/she was excluded ($N = 3$: multiple sclerosis, carcinoma, head injury).

Ninety-four per cent of the original cohorts had previously indicated willingness to participate in further research. A random sample of 'ill' potential participants from each of the three cohorts plus a random sample from the 'well' Gulf group, was then contacted (total $N = 738$). Two people had died, and 131 were untraceable – probably because they had moved without giving a forwarding address. This left 605 who were then asked if they wished to volunteer for phase 2 of the study. Out of this number, altogether 341 (56·4%) so volunteered and were able to attend, 98 (62·4%) who had been initially classified as 'Gulf well'; 111 (66·9%) classified as 'Gulf ill'; 78 (42·5%) as 'Era ill'; and 54 (56·8%) as 'Bosnia ill' – the Era ill group being significantly lower than the other groups ($\chi^2 = 24{\cdot}9$, $P < 0{\cdot}01$).

Participants and non-participants were compared on a number of social and demographic variables, health perception and psychiatric morbidity (Table 1). They did not differ in terms of gender, education or health perception. Participants' and non-participants' mean General Health Questionnaire-12 (Goldberg, 1972) scores recorded in phase 1 were 4·4 (3·9) and 3·9 (3·9) respectively ($t = 0{\cdot}074$, NS). A greater proportion of the Era group chose not to participate and Era group non-attenders contained proportionately more non-commissioned officers. Proportionately fewer Gulf well non-army personnel did not participate: 41·3% were still serving, with a mean duration of service across the sample being 182 months. The total phase 1 cohort contained 4·3% who had been medically discharged. Of those invited to phase 2 which was of course weighted to those who were 'ill', 9·4% of the participants and 17·3% of the non-participants were medically discharged ($\chi^2 = 4{\cdot}18$, $P = 0{\cdot}04$).

Participants ages ranged between 22 and 62 years, with the Gulf groups being the eldest and the Bosnia group the youngest, who correspondingly comprised more people still serving: 25 participants were female; 285 were Army; 56 were Royal Air Force, Navy or Marines; 252 were educated to O-level (high school) standard and 88 had A-levels (equivalent to a college diploma) or above.

Participants were required to attend the unit for a day to carry out both a full medical examination and neuropsychological assessment. They were reimbursed for their expenses. No subject with an established medical cause for cognitive impairment was included.

Neuropsychological testing

Two trained research assistants carried out all of the cognitive testing sessions. Both were blind to participants' group status. The order of the assessments (medical and psychological) was randomly assigned within each subject group. Scores on those items that have a subjective element in their scoring, that is the vocabulary and information subtests of the Wechsler Adult Intelligence Scale – Revised (Wechsler, 1981), were compared on a pilot sample until identical ratings were achieved consistently. Participants gave written, informed consent. Data were double entered and inputting errors eliminated. Testing lasted approximately two and half-hours, including two short breaks. The following tests were administered.

General functioning

Wechsler Adult Intelligence Scale – Revised (*WAIS-R*) (Wechsler, 1981)

Verbal subtests: Vocabulary, Digit Span, Arithmetic, Similarities; Performance subtests: Picture Arrangement, Block Design, Object Assembly; Digit Symbol. Scaled scores (not age corrected) were used in the analysis.

Table 1. *Demographic, military and health perception data on participants and non-participants across the military groups as originally defined*

Variable	Original classification	Participant ($N = 341$)	Non-participant ($N = 264$)
Sex	Gulf well		
	Male	92	55
	Female	6	4
	Total (%)	98 (62)	59 (38)
	Gulf ill		
	Male	105	53
	Female	6	2
	Total (%)	111 (67)	55 (33)
	Era ill		
	Male	72	93
	Female	6	14
	Total (%)	78 (42)	107 (58)
	Bosnia		
	Male	50	36
	Female	4	7
	Total (%)	54 (56)	43 (54)
Age (years)	Gulf well	33·5 (5·6)	33·3 (5·5)
	Gulf ill	35·9 (7·2)	36·1 (8·4)
	Era ill	38·7 (8·5)	35·9 (7·5)
	Bosnia ill	31·3 (7·0)	32·3 (7·8)
Service	Gulf well*		
	Army	89	41
	Marine, Navy, RAF	9	18
	Gulf ill		
	Army	89	47
	Marine, Navy, RAF	22	8
	Era ill		
	Army	53	83
	Marine, Navy, RAF	25	24
	Bosnia		
	Army	54	43
	Marine, Navy, RAF	0	0
Rank current/exit	Gulf well		
	Other	16	16
	NCO	63	34
	Officer	13	5
	Gulf ill		
	Other	25	7
	NCO	80	42
	Officer	2	3
	Era ill†		
	Other	8	23
	NCO	54	73
	Officer	9	3
	Bosnia ill		
	Other	20	14
	NCO	30	26
	Officer	2	3
Still serving	Gulf well		
	Yes	63	31
	No	32	28
	Gulf ill		
	Yes	45	18
	No	63	36
	Era ill		
	Yes	34	36
	No	44	71
	Bosnia ill		
	Yes	43	36
	No	10	7

Table 1. (*contd.*)

Variable	Original classification	Participant ($N = 341$)	Non-participant ($N = 264$)
Medically discharged	Gulf well		
	Yes	0	2
	No	38	27
	Gulf ill		
	Yes	8	5
	No	58	33
	Era ill		
	Yes	5	18
	No	39	57
	Bosnia ill		
	Yes	2	1
	No	9	7
Education	Gulf well		
	No qualifications	13	11
	School/College	80	46
	Gulf ill		
	No qualifications	30	17
	School/College	72	36
	Era ill		
	No qualifications	19	29
	School/College	55	72
	Bosnia ill		
	No qualifications	10	11
	School/College	41	31
Health perception‡	Gulf well	69·0 (23·5)	75·2 (20·1)
	Gulf ill	30·4 (19·6)	31·3 (20·8)
	Era ill	47·5 (25·2)	44·1 (23·9)
	Bosnia ill	51·1 (25·1)	58·8 (20·1)

* χ^2, $P < 0{\cdot}01$.
† χ^2, $P < 0{\cdot}05$.
‡ Analysis of variance. Main effect of group ($F = 85{\cdot}5$, $P < 0{\cdot}001$); but not participation ($F = 2{\cdot}07$, $P = 0{\cdot}15$). No significant interaction between participation and group ($F = 1{\cdot}79$, $P = 0{\cdot}146$).

National Adult Reading Test, 2nd Edition (*NART*) (Nelson, 1991)

A reading test of 'irregular' words, which gives a stable estimate of pre-morbid IQ was administered.

Letter Number Sequencing task (*WAIS III*) (Wechsler, 1997)

This task was administered as a test of working memory.

Attention

Paced Auditory Serial Addition Task (*PASAT*) (Gronwall & Wrightson, 1975)

Sixty single digits were presented auditorily (1 digit per 2 s). Participants were asked to add each digit to the one that comes before it and to give their answer aloud before the next digit had been auditorily presented. The score is the number of correct responses given (max. 60).

Sustained Attention to Response Task (*SART*) (Robertson *et al.* 1997)

A computer administered vigilance task. A score representing the number of errors of commission (max. 25) is calculated. Also, mean reaction time (RT) for correct responses is calculated.

Stroop Neuropsychological Screening Test (Trennery *et al.* 1989)

In this colour word naming test, a Stroop effect was calculated based on the difference in time to name the colour of ink of incongruent colour-word stimuli minus naming of colour blocks.

Trail Making Test, Parts A and B (Reitan, 1992)

This is a test of motor sequencing and set shifting in which the time (s) to complete the simple A trail and the alternating B trail were

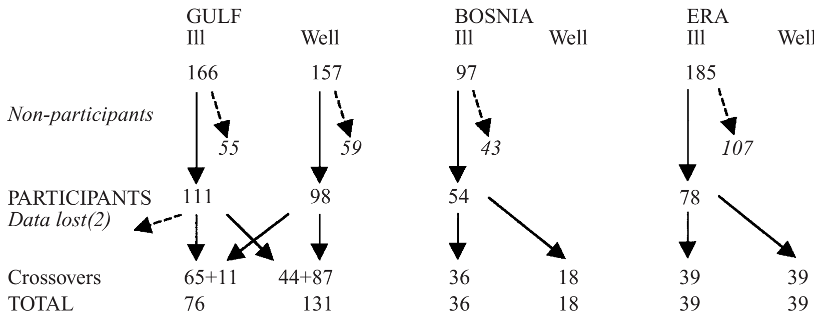

FIG. 1. Case crossovers: changes in health status between original recruitment and testing.

recorded, the difference (B − A) was taken as the dependent measure.

Memory

The Wechsler Memory Scales (WMS-R) (Wechsler, 1987) for Logical Memory, Immediate and Delayed Recall, Verbal Paired Associates Immediate and Delayed Recall; and Camden Recognition Memory Tests (Warrington, 1996). For faces and words (each scored out of 25) were employed.

Motor skills

The Purdue pegboard (Purdue Research Foundation, 1948) was used to assess motor skills. Assembly involves the placement of several small components (pins, collars and washers) into the pegboard, bimanually in the allotted time (right hand, left hand and both together).

Self-completion questionnaires

Four self completion questionnaires were administered: Cognitive Failures Questionnaire (CFQ) (Broadbent *et al.* 1982) range 0–100; Beck Depression Inventory (BDI) (Beck & Steer, 1993) range 0–63; State-Trait Anger Expression Inventory (STAXI – AX-EX) (Forgays *et al.* 1997) range 0–72; and Mississippi Combat Related PTSD Scale (Keane *et al.* 1988) adapted for the GW, 39 items – range 39–195.

Case crossovers

Due to the time lapse between filling out the postal questionnaire, and recruiting for phase 2 (12–18 months), the SF36-PF was administered again when participants attended the unit for testing (Fig. 1). The figure shows that the majority (89 %) of participants originally classified as Gulf well are, according to the same criterion, still 'well', with only 11 becoming 'ill'. However, 40 % of the original Gulf ill group are now classified as 'well', with fairly substantial crossovers in the Era and Bosnia groups too: 50 % and 33 % becoming well. These crossovers will be discussed in more detail in another paper.

Given this finding and our primary interest in the influence of ill health and psychopathology on cognitive performance, we decided to concentrate our analyses on the current classification of the groups. Statistical analysis was carried out using SPSS.PC v8.

RESULTS

Group comparison (based on SF36 responses at time of neuropsychological testing)

Using this classification, there are six groups ($N = 339$), reflecting those who are currently 'well' or 'ill' from the Gulf, Bosnia and Era cohorts. Demographic details are given in Table 2. Again the groups differ significantly in terms of age and pre-morbid IQ.

A 3×2 ANCOVA was used to compare the six groups on the test variables with deployment (Gulf, Era, Bosnia) and current health status (well, ill) as the factors. Again the analysis covaried for: (*i*) age, education and NART-estimated IQ; and (*ii*) for these potential confounders plus BDI score. Least significant difference *post hoc* procedure (alpha set at 0·05) was used to identify significant differences

Table 2. *Demographic and IQ data for military personnel by current SF36 case groupings*

	N	Age	Service (months)	Females %	NART IQ
Gulf ill	76	37·0 (7·5)	180·5 (84·9)	7·9	99·2 (11·1)
Gulf well	131	35·0 (5·9)	176·4 (78·7)	3·8	101·6 (10·3)
Era ill	39	39·4 (7·7)	203·1 (98·2)	7·7	102·2 (9·9)
Era well	39	39·8 (9·2)	224·8 (95·3)	7·7	104·3 (11·1)
Bosnia ill	36	31·6 (7·7)	140·4 (76·2)	5·6	104·3 (11·1)
Bosnia well	18	32·7 (5·3)	170·7 (84·3)	11·1	93·8 (11·1)
F	Total = 339	8·40**	4·408**		3·24**

Values are means with standard deviations in parentheses, unless stated otherwise.
$^{**}P < 0{\cdot}01$.

between groups (results are shown in Table 3, further details including standard deviations available on request).

Table 3 shows that the Gulf ill group had the highest mean scores on the psychopathology measures, and reported the most cognitive failures on the CFQ. They also tended to show poorer performance on most of the cognitive measures. Exceptions to this were on the WAIS-R verbal subtests and Verbal IQ, where the Bosnia well group was marginally worse than Gulf ill, and the SART errors variable, where Era ill were worse. The Bosnia well group also had the lowest recall score on WMS logical memory, although this was very similar to the Gulf ill group.

Health status

Not surprisingly, there were significant main effects of health status, on measures of emotional dysfunction in that people who were ill had significantly higher levels of depression on the BDI and higher scores on the Mississippi PTSD scale. They showed higher trait anger and reported more cognitive failures. In addition, there were main effects of health status with ill-health being associated with poorer performance on WAIS-R Performance IQ, although the digit symbol subtest (sensitive to attention) was the only one on which scores were significantly lower. Performance was also significantly poorer on the Trails test and SART accuracy. Given that the ill-health status group were significantly more depressed, it was likely that this factor could contribute to the cognitive impairments noted. After adjustment for BDI, the only differences that remained significant were on Mississippi scores and the digit symbol test. Carrying out a Bonferroni correction for the number of contrasts described in Table 3, raises the statistical threshold to ($P < 0{\cdot}0017$). BDI score remains significantly associated with poorer health. Prior to adjustment for depression, the only scales significantly associated with poorer health after correction for multiple comparisons were the Mississippi scores and digit symbol substitution subtest. Only the Mississippi survived after BDI was taken into account as well.

Deployment

In terms of main effects of deployment, only a few measures were significantly different across groups after adjusting for age, pre-morbid IQ and education, namely Mississippi scale scores, STAXI anger expression scores with depression scores not significantly different. The differences were accounted for by higher (more pathological) scores in the Gulf deployed group.

Regarding cognitive measures, there were main effects of deployment on the WAIS-R Verbal and Performance IQ (in particular, the block design subtest), the delayed Wechsler logical memory test and the Purdue Pegboard assembly task. The self-report CFQ also showed a significant main effect of deployment. *Post hoc* tests revealed that the Gulf group had significantly lower Verbal IQ and Performance IQ scores (particularly block design) than the Era group, and the lowest pegboard performance. The main effect of deployment on CFQ was, however, due to the Era group reporting less cognitive failures than both the actively deployed groups. The Era group also had superior logical memory scores than both deployed groups.

Despite depression scores not differing signifi-

Table 3. *Analysis of covariance based on new SF36 classification. Means for psychometric measures for each group*

Test Mean score or %	Gulf ill $N = 76$	Era ill $N = 39$	Bosnia ill $N = 36$	Gulf well $N = 131$	Era well $N = 39$	Bosnia well $N = 18$, M	Main effects (F) Deployment	 Health
BDI	14·5	10·3	10·2	6·96	6·21	8·00	2·60	19·28**
Cases > 9	25%	7%	8%	10%	7%	3%	—	—
Mississippi	95·2	81·6	82·4	73·7	69·7	73·1	4·59*	29·33**††
Cases > 106	20%	6%	7%	8%	2%	0	—	—
STAXI Trait	20·6	18·0	19·7	17·6	16·7	19·3	1·60	5·31*
STAXI Express	29·2	23·0	25·7	23·3	21·4	26·6	3·25*	3·15
CFQ	54·5	46·8	49·8	42·7	36·7	45·6	4·86**†	16·9**
WAIS: VIQ	94·6	101·9	98·7	98·3	102·2	94·4	4·61*†	1·03
PIQ	102·1	109·1	106·9	108·1	113·3	104·9	3·54*	5·68*
Digit Span	9·6	9·8	10·3	10·0	10·2	10·4	1·89	2·13
Picture Arr.	9·3	10·4	9·5	9·8	9·9	8·7	1·20	0·01
Vocabulary	9·1	10·0	9·8	9·7	10·6	8·4	1·70	1·36
Block Design	10·5	11·5	12·4	11·8	12·2	11·5	3·21*	2·88
Arithmetic	9·2	11·1	10·1	10·2	10·6	8·9	2·40	0·05
Object Assem.	10·0	10·7	11·0	10·8	11·5	11·0	2·53	3·78
Digit Symbol	9·0	9·4	9·9	10·2	10·6	10·1	1·29	11·5**††
Similarities	8·5	9·3	9·4	9·0	9·4	8·5	1·89	0·001
Trails A	30·6	28·9	27·8	27·9	26·6	23·9	2·86	2·45*
Trails B	66·2	63·0	59·9	58·2	56·3	54·1	1·66	2·78*
Trails B-A	35·4	34·0	32·2	30·4	29·4	30·2	0·226	4·51*
Stroop	69·3	65·8	66·3	62·4	61·4	65·2	0·78	3·03
PASAT	30·4	33·2	29·7	31·7	32·6	31·0	0·04	1·01
SART errors	12·1	13·6	11·6	11·0	10·7	10·9	1·01	3·89*
SART RT	347·3	337·0	342·9	323·0	354·6	332·9	0·23	0·335
Log. Memory 1	23·8	25·9	24·8	25·8	26·6	23·5	2·12	1·22
Log. Memory 2	18·6	22·1	20·8	20·9	22·3	18·8	4·41*	0·25
Verbal p-a 1	19·3	20·4	20·7	19·3	20·1	19·2	2·41	1·44
Verbal p-a 2	7·5	7·6	7·7	7·6	7·8	7·7	1·56	1·00
Purdue R+L+B	38·2	38·9	40·1	38·5	39·4	39·5	2·95	0·63
Assembly	35·2	37·4	37·9	36·2	37·5	38·0	6·40**††	0·86

* $P < 0{\cdot}05$; ** $P < 0{\cdot}01$ after controlling for age, education and NART IQ – corresponds to F value in Table; † significant at $P < 0{\cdot}05$; †† $P < 0{\cdot}01$ after controlling for BDI. Only tests where there was a suggestion of group differences are reported in the Table; LNS, Camden, not shown.

The proportion of subjects doing each test ranged from 95–100%.

cantly across the three deployment groups, we chose to adjust for BDI since trends were evident suggesting differences may be present that could affect cognition. After adjustment, very few of these results remained significant. The Era cohort still had significantly lower (i.e. less pathological) scores on the CFQ than the Gulf and Bosnia cohorts, and the Gulf group still showed a difference in Verbal IQ from the Era group. The Purdue Pegboard assembly score remained significantly lower for the Gulf deployed group *versus* both the other groups. Again, none of these contrasts remained significant after Bonferroni correction, even before adjustment for BDI score which attenuated all the contrasts.

Importantly, there were no significant or even near significant interactions between deployment and health status for any of the variables.

We attempted to assess the cross-sectional predictors of subjective cognitive impairment using multiple regression. The CFQ score was the dependent variable and the following independent variables were entered into the model: age, education status (college/no college), deployment (Gulf/non-Gulf), Full Scale IQ, Mississippi score and BDI score. The overall R^2 was 0·547. The only variables which significantly contributed to the model were BDI score (β (S.E.)) 0·74 (0·125), $t = 5{\cdot}92$, $P < 0{\cdot}0001$; and Mississippi score (β (S.E.)) 0·29 (0·048), $t = 6{\cdot}095$, $P < 0{\cdot}0001$. The standardized coefficients for the two scales were similar: 0·381 and 0·401 respectively (all the others were $< 0{\cdot}05$).

Table 4. *Analysis of covariance between Gulf ill 'Fukuda' cases and Gulf well 'non-Fukuda' cases, means and standard deviations for psychometric measures for each group*

Test	Fukuda GWS ($N = 65$) Mean (S.D.)	Non-GWS ($N = 33$) Mean (S.D.)	ANOVA P
BDI	15·7 (10·5)	4·1 (5·6)	**
Mississippi	98·3 (25·3)	61·6 (9·7)	**‡‡
STAXI Trait	21·3 (6·0)	16·4 (3·7)	**
STAXI Expression	30·0 (11·6)	20·3 (8·1)	**
CFQ	56·4 (16·8)	33·30 (14·29)	**‡‡
WAIS VIQ	94·5 (10·0)	100·8 (10·3)	**
WAIS PIQ	101·8 (13·8)	110·6 (11·7)	**
Digit Span	9·6 (2·8)	10·6 (2·5)	—
Picture Arrangement	9·2 (3·1)	10·7 (2·5)	—
Vocabulary	9·0 (1·9)	10·3 (2·5)	‡
Block Design	10·5 (2·8)	12·1 (2·7)	—
Arithmetic	9·1 (2·8)	10·8 (2·8)	*
Object Assembly	9·9 (2·6)	10·7 (1·7)	—
Digit Symbol	8·9 (2·1)	10·6 (2·1)	**‡
Similarities	8·6 (1·7)	9·2 (1·4)	—
Trails B-A	36·0 (18·9)	24·2 (11·3)	**
Stroop	71·1 (19·8)	57·3 (14·4)	**
PASAT	29·7 (9·0)	36·3 (11·9)	**‡
SART errors	11·8 (6·39)	8·4 (6·0)	*
SART RT	345·6 (81·7)	330·7 (78·8)	—
Logical Memory 1	23·6 (6·3)	25·8 (4·9)	—
Logical Memory 2	18·3 (6·6)	21·2 (5·5)	—
Verbal Associates 1	19·3 (3·2)	18·5 (3·6)	**†
Verbal Associates 2	7·5 (0·89)	7·7 (0·74)	—
Purdue R+L+B	38·2 (3·5)	38·9 (4·5)	—
Purdue Assembly	34·8 (5·75)	36·1 (5·7)	—

Only tests where there was a suggestion of group differences are reported in the Table; LNS, Camden, not shown.

* $P < 0·05$; ** $P < 0·01$ after controlling for age, education and NART IQ; † near significant at $P = 0·07$; ‡ $P < 0·05$; ‡‡ $P < 0·01$ after controlling for BDI.

Fukuda cases: multi-symptom case definition (Fukuda *et al.* 1998).

'Gulf War syndrome'

The Center for Disease Control has derived a working definition for 'Gulf war syndrome' (GWS) (Fukuda *et al.* 1998) based on the presence of symptoms from at least two of the three symptom-clusters: fatigue, mood/cognition and musculo-skeletal, lasting 6 months or more in returnees from the GW. Using the symptom classification recorded at phase 1 combined with the duration criterion of persisting illness according to the SF36 at the time of testing (and excluding crossovers), we operationalized our own working definition of GWS. We were thus able to compare cases ($N = 65$) so defined, with Gulf veterans who were entirely 'well' – i.e. they had no symptoms at all at time 1 and were currently well ($N = 33$) (Table 4). The cases were slightly older (36·7 *v.* 35·1 years, $P < 0·005$); with 75·0% educated to below college standard *v.* 69·7%, and this is reflected in NART IQ scores of 99·8 *v.* 103·8 ($P < 0·001$).

ANCOVA (Table 4), covarying for age, NART, and education, showed that there were several significant differences between the two groups. The GWS cases showed considerably more emotional distress in terms of depression, PTSD symptoms, anger and subjective cognitive impairment than the control group. There were also impairments on the WAIS-R Verbal and Performance IQ scales, the Trails test, Stroop, PASAT and SART errors. After adjustment for depression, a few significant differences remained, namely Mississippi, CFQ, the vocabulary and digit symbol subtests of the WAIS-R and PASAT with the Trails and Stroop just failing at the 5% level. The increase in Mississippi scale score was the only one to remain significantly associated with GWS after Bonferroni correction.

DISCUSSION

The data we report are based on a random sample of GW veterans plus both deployed and non-deployed military controls. Of course deployment is not a random process which may have introduced biases, but all subjects at least have military service in common. Furthermore, the Gulf well and ill groups both underwent similar deployments. The response of the Gulf and Bosnia groups to an invitation to attend for a comprehensive medical and psychological assessment was satisfactory although the rate for the non-deployed participants was, understandably, lower. Those that did attend had not been obviously more impaired than the remainder when they originally participated in the questionnaire survey and fewer had been medically discharged, but we cannot exclude selection bias affecting the responders.

This study, which represents phase two of an ongoing programme of research, was not designed to give the prevalence of cognitive deficits and emotional disorders, since an estimate of this has already been provided by the much larger survey (Unwin *et al.* 1999). Most symptoms (including cognitive difficulties) and medical conditions were reported in approximately twice as many GW veterans as in either of the control groups. Here we were more concerned with the pattern and precise nature of any deficits using standardized tests.

It became clear to us that the original groupings were no longer valid and that a large number of 'ill' participants were now 'well' by our definition. This prompted analyses with six groups of unequal numbers. It should be noted that the 'well' Bosnia and Era participants are best described as 'recovered' while the Gulf well group include recovered and 'always well'. Similarly, the Gulf ill now includes a small number of 'newly ill'. Results from the small Bosnia well group were somewhat variable and were inferior to the ill group on some tests. The long-term health effects of peace-keeping are only now beginning to be studied (Bramsen *et al.* 2000). Nevertheless, while adding complexity to the analysis this approach had the advantage, apart from its face validity, that the effect of place of deployment could be looked at separately from the presence of illness, and the interaction between these two factors examined. The results showed that deployment occasionally emerges as a significant main effect in the ANOVA, after controlling for potential confounders. However, *post hoc* examination of this points to deployment to the Gulf being the more adverse for some measures including PTSD symptoms and Purdue pegboard, while active deployment (i.e. Gulf and Bosnia) was more adverse for cognitive failures and WAIS-R measures. After adjustment for depression, Purdue assembly was the only measure significantly affected by deployment to the Gulf – regardless of whether a person was 'ill' or 'well' (see below). None of these statistical differences was robust and all disappeared after controlling for multiple comparisons.

Further examination of the main effect for health status revealed that four out of the five self-report measures were related to ill-health plus four of the cognitive measures. So in general, being 'well' means having fewer psychological and cognitive symptoms or deficits. Of the latter, only one subtest of the WAIS-R remained significant after adjustment for depression. There was no interaction between these two factors – there is nothing to suggest that those whose physical functioning fell below the cut-off, following deployment to the Gulf, were any worse than those similarly affected following deployment elsewhere.

PTSD scores also remained significantly greater in all 'ill' veterans across deployment categories despite adjustment for depression suggesting that such symptoms are, to some extent, distinct from low mood and may contribute to perceived loss of physical and mental functioning. The Mississippi scale has a cut-off of 107 for probable cases of PTSD. Just over a quarter of the Gulf ill group scored above this compared to 12·8% across the whole sample. However, ongoing studies on the same individuals looking at clinical diagnosis following a standardized examination show very low levels of DSM-IV PTSD (around 3%: see Ismail *et al.* 2002). Recent studies have highlighted a strong association between PTSD symptoms and unexplained physical complaints (Engel *et al.* 2000). Whatever the precise mechanism underlying this association (e.g. alcohol use, changes in social supports (Wagner *et al.* (2000)), it appears to apply to unexplained cognitive complaints as well.

We selected out a group who conformed to a working definition of 'Gulf War syndrome'. In doing so we may simply have selected a more symptomatic subgroup differing only quantitatively from the remainder. Nevertheless, when depression was covaried in the analysis (which was on average in the mild clinical range), very few neuropsychological performance differences between the GWS group and well Gulf controls remained. These differences seemed to spare memory and motor tasks and affect those more demanding of attention.

One striking pattern was the tendency for psychopathology measures to be strongly associated with both deployment to the Gulf and ill-health (including GWS) and that deficits on neuropsychological measures were less prominent. More media attention seems to be given to the latter perhaps because of a supposed direct link to neurological damage. Furthermore, it may be more socially acceptable for an individual to complain of difficulties in memory and concentration than low mood. However, psychopathology is an obvious and important confound in measures of cognition, given the well recognized detrimental effect depression has on effortful information processing (Weingartner *et al.* 1981; Hartlage *et al.* 1993; Brown *et al.* 1994; Purcell *et al.* 1997). By adjusting for depression, few performance measures remained significantly different across the groups, however defined. In this we are assuming that depression leads to impaired cognition although it is possible that the direction of causality is in the opposite direction in some cases. It is also conceivable that adverse exposures cause disturbances of mood and anxiety, which in turn cause cognitive problems.

A high level of anxiety has long been known to affect performance detrimentally (Yerkes & Dodson, 1908). We did not have a pure measure of the anxiety, analogous to the BDI, so could not examine this further. However, the BDI and equivalent anxiety inventory correlate highly (Enns *et al.* 1998) and anxiety and depression overlap considerably in non-clinical populations (Judd *et al.* 1998). Hence, it is reasonable to assume that adjusting the performance data for anxiety would have yielded a similar pattern of results as adjusting for depression.

The CFQ was devised to provide a proxy measure of everyday cognitive lapses or failures (Broadbent *et al.* 1982) and has been shown to correlate with accidents in US Naval recruits (Larsen *et al.* 1997). However, the questionnaire also taps into negative self-evaluation – a cardinal feature of depression. Clinical studies have shown the CFQ to be more related to depression than cognitive performance (Wagle *et al.* 1999) and this has been confirmed in GW veterans by Binder and colleagues (1999) as well as our own group (Farrin *et al.* 2002). In fact the Spearman correlation between BDI and CFQ was 0·66 in the current sample (while correlation with WAIS-R PIQ was $-0{\cdot}147$ and between CFQ and WAIS-R PIQ was $-0{\cdot}084$). In our regression analyses, depression and PTSD symptom scores strongly predicted CFQ but cognitive measures and deployment did not approach significance. Having said that CFQ scores were high in affected veterans compared with healthy norms ($N = 2379$; mean score (S.D.) = 33·6 (12·8) Larson *et al.* 1997), and remained significantly related to active deployment even after depression was taken into account – unlike most of the observer-rated cognitive test scores. This is most likely a reflection of increased symptom monitoring as a consequence of stressful military duties perhaps accentuated by well publicized health fears peculiar to the Gulf conflict.

The Purdue assembly finding, though isolated, merits further comment. Performance was only mildly inferior to published norms: males, age range 15–40 years, mean items in the time allowed (S.D.) = 39·0 (4·9) *v.* 35·8 (4·7) for total Gulf group. One explanation for this, is that personnel deployed to the Gulf have suffered some subtle neurotoxic damage (see Haley *et al.* 1997) specifically affecting bi-manual constructional dexterity. A precise mechanism and pathophysiology, which could account for this pure deficit, is however, elusive. An attractive possibility is organophosphate pesticide toxicity, although exposure to pesticides was similarly high in both Bosnia and Gulf cohorts in our study (Unwin *et al.* 1999). Purdue performance more generally was found to be compromised in a group of 44 GW veterans who came forward for health screening as reported by Axelrod & Milner (1997), but in that study it was merely one aspect of mild generalized deficits on timed motor tests and was confounded by emotional distress (see also Silanpaa *et al.* 1997).

The performance of 'ill' veterans, while inferior to 'well' on many tests, was usually within normal range derived from published norms. For example, performance on the Trails A & B was within 40–50th and 30–40th centiles respectively in the Gulf 'ill' veterans (Lezak, 1995; cf Haley *et al.* 1997). PASAT scores (correct responses) in all the participants tested, both well and ill, were somewhat poorer than published 'well educated' healthy norms ($N = 30$, age 30–49), mean (S.D.): 41·9 (10·2) (Lezak, 1995). This suggests that we should not assume that past and present military personnel at this age perform better than average on measures of this kind, i.e. there is no substantial 'healthy warrior' effect (Haley, 1998; Kang & Bullman, 1998).

Conclusions

There is no evidence of major neuropsychological impairment in GW veterans. Those weak effects that were detected were patchy in terms of the cognitive systems implicated. Furthermore, they were just as likely to be attributable to any active deployment and hence not likely to be related to specific Gulf-related exposures – with the exception of the Purdue Assembly measure. Test performance in unwell veterans was impaired relative to well controls but generally within the normal range. More noteworthy was the increase in self-reported depression, anger and PTSD symptoms and subjective cognitive failures in association with ill-health. Although we attempted to control for depression in our analysis of the neuropsychological data, we propose that this level of psychopathology is the most parsimonious account for the bulk of reported subjective cognitive deficits, as well as the more sparse objective deficits. It is possible that subjective (and minor objective) cognitive deficits lead to low mood and increased 'stress'. This hypothesis has been examined in relation to unexplained symptoms in the civilian population, such as chronic fatigue sufferers. The findings show that even after controlling for functional disability by employing comparison groups with multiple sclerosis, neuromuscular conditions, rheumatoid arthritis, etc., levels of psychiatric disturbance remain significantly higher in the 'unexplained' illness group (see David, 1991 for a review). Hence depression as a reaction to functional impairment or subjective ill-health is an inadequate explanation on its own.

This broadly concurs with earlier studies (Goldstein *et al.* 1996; Axelrod & Milner, 1997; Silanpaa *et al.* 1997). Haley *et al.* (1997) and Hom *et al.* (1997) however reported what they considered to be evidence of chronic neuropsychological impairment in their somewhat older sample, which they attributed to neurotoxic wartime exposures. All these studies have been poorly controlled and have used small numbers of symptomatic participants, without taking into account adequately the non-specific effects of illness on test performance. An exception is the preliminary report by Anger *et al.* (1999). The authors isolated a subgroup of the cases who performed more slowly on an automated working memory test. The nature of the implied deficit finds no counterpart in the current study where digit span, SART, PASAT and letter number span failed to reveal group differences independently of depressed mood. The full report from this study (Storzbach *et al.* 2000) attempted to classify Gulf War veterans with and without a cluster of unexplained symptoms. They confirmed that 'cases' and 'controls' differed substantially on measures of psychopathology while differences on cognitive tests had modest effect sizes. The measures that best distinguished cases and controls comprised four of the psychopathology and 'distress' measures but none of the cognitive measures.

In summary, the results from our studies have demonstrated yet again that emotional and psychological disorder is common in GW veterans and, in a minority, likely to be clinically significant. This could in itself explain the subjective cognitive problems reported by veterans, although the direction of causality cannot be assumed from such data. Finally, the increased prevalence of mood disorders in Gulf veterans, is an obvious cause for concern, a target for therapeutic intervention and for research into possible preventative action. Furthermore, subjective cognitive failures, regardless of their origin, do appear to be related to accidents (Larsen *et al.* 1997) and this may contribute to a causal chain resulting in the reported increased mortality (MacFarlane *et al.* 2000) in this group.

The study was funded by the U.S. Department of Defense. We are grateful to the Royal British Legion for their support; to veterans and their families for their co-operation; and to Tom Manley and Ian Robertson for advice on the SART.

REFERENCES

Anger, W. K., Storzbach, D., Binder, L. M., Campbell, K. A., Rohlman, D. S., McCauley, L., Kovera, C. A. & Davis, K. L. (1999). Neurobehavioral deficits in Persian Gulf veterans: evidence from a population-based study. *Journal of the International Neuropsychological Society* **5**, 203–212.

Axelrod, B. N. & Milner, I. B. (1997). Neuropsychological findings in a sample of Operation Desert Storm veterans. *Journal of Neuropsychiatry and Clinical Neurosciences* **9**, 23–28.

Beck, A. T. & Steer, R. A. (1993). *Beck Depression Inventory*. The Psychological Corporation: San Antonio, TX.

Binder, L. M., Storzbach, D., Anger, W. K., Campbell, K. A., Rohlman, D. S. & the Portland Environmental Hazards Research Center (1999). Subjective cognitive complaints, affective distress, and objective cognitive performance in Persian Gulf War veterans. *Archives of Clinical Neuropsychology* **14**, 531–536.

Bramsen, I., Dirkzwager, A. J. E. & van der Ploeg, H. M. (2000). Predeployment personality traits and exposure to trauma as predictors of posttraumatic stress symptoms: a prospective study of former peacekeepers. *American Journal of Psychiatry* **157**, 1115–1119.

Broadbent, D. E., Cooper, P. F., Fitzgerald, P. & Parkes, K. R. (1982). The Cognitive Failures Questionnaire (CFQ) and its correlates. *British Journal of Clinical Psychology* **21**, 1–16.

Brown, R. G., Scott, L. C., Bench, C. J. & Dolan, R. J. (1994). Cognitive function in depression: its relationship to the presence and severity of intellectual decline. *Psychological Medicine* **24**, 829–847.

Coker, W. J., Bhatt, B. M., Blatchley, N. F. & Graham, J. T. (1999). Clinical findings for the first 1000 Gulf war veterans in the Ministry of Defence's medical assessment programme. *British Medical Journal* **318**, 290–294.

Cope, H., Pernet, A., Kendell, B. & David, A. (1995). Cognitive functioning and magnetic resonance imaging in chronic fatigue. *British Journal of Psychiatry* **167**, 86–94.

David, A. S. (1991). The post viral fatigue syndrome and psychiatry. *British Medical Bulletin* **47**, 966–988.

Engel, C. C., Liu, X., McCarthy, B. D., Miller, R. F. & Ursano, R. (2000). Relationship of physical symptoms to post traumatic stress disorder among veterans seeking care for Gulf War-related health concerns. *Psychosomatic Medicine* **62**, 739–745.

Enns, M. W., Cox, B. J., Parker, J. D. & Guertin, J. E. (1998). Confirmatory factor analysis of the Beck Anxiety and Depression Inventories in patients with major depression. *Journal of Affective Disorders* **47**, 195–200.

Farrin, L., Hull, L., Unwin, C. & David, A. (2002). Objective and subjective measures of attention: effects of depressed mood. *Journal of Neuropsychiatry and Clinical Neuroscience* (in the press).

Forgays, D. G., Forgays, D. K. & Spielberger, C. D. (1997). Factor structure of the State–Trait Anger Expression Inventory. *Journal of Personality Assessment* **69**, 497–507.

Fukuda, K., Nisenbaum, R., Stewart, G., Thompson, W. W., Robin, L., Washko, R. M., Noah, D. L., Barrett, D. H., Randall, B., Herwaldt, B. L., Mawle, A. C. & Reeves, W. C. (1998). Chronic multi-symptom illness affecting air force veterans of the Gulf War. *Journal of the American Medical Association* **280**, 981–988.

Gass, C. S. & Apple, C. (1997). Cognitive complaints in closed-head injury: relationship to memory test performance and emotional disturbance. *Journal of Clinical and Experimental Neuropsychology* **19**, 290–299.

Goldberg, D. (1972). *The Detection of Psychiatric Illness by Questionnaire*. Oxford University Press: London.

Goldstein, G., Beers, S. R., Morrow, L. A., Shemansky, W. J. & Steinhauer, S. R. (1996). A preliminary neuropsychological study of Persian Gulf veterans. *Journal of the International Neuropsychological Society* **2**, 368–371.

Gronwall, D. & Wrightson, P. (1975). Cumulative effect of concussion. *Lancet* ii, 995–997.

Haley, R. W., Hom, J., Roland, P. S., Bryan, W. W., van Ness, P. C., Bonte, F. J., Devous, M. D., Mathews, D., Fleckenstein, J. L., Wians, F. H., Wolfe, G. I. & Kurt, T. L. (1997). Evaluation of neurologic function in Gulf war veterans. A blinded case-control study. *Journal of the American Medical Association* **277**, 223–230.

Haley, R. W. (1998). Point: bias from the 'Healthy-Warrior Effect' and unequal follow-up in three Government Studies of health effects of the Gulf War. *American Journal of Epidemiology* **148**, 315–323.

Haley, R. W., Kurt, T. L. & Hom, J. (1997). Is there a Gulf war syndrome? Searching for syndromes by factor analysis of symptoms. *Journal of the American Medical Association* **277**, 215–222.

Hartlage, S., Alloy, L. B., Vazquez, C. & Dykman, B. (1993). Automatic and effortful processing in depression. *Psychological Bulletin* **113**, 247–278.

Hom, J., Haley, R. W. & Kurt, T. L. (1997). Neuropsychological correlates of Gulf war syndrome. *Archives of Clinical Neuropsychology* **12**, 531–544.

Iowa Persian Gulf Study Group (1997). Self-reported illness and health status among Persian Gulf war veterans: a population-based study. *Journal of the American Medical Association* **277**, 238–245.

Ismail, K., Davies, K., Brugha, T. S., Hotopf, M., Seed, P., Palmer, I., Reid, S., Unwin, C., David, A. S. & Wessely, S. (2002). The mental health of United Kingdom Gulf War veterans. *British Medical Journal* (in the press).

Judd, L. L., Kessler, R. C., Paulus, M. P., Zeller, P. V., Wittchen, H. U. & Kunovac, J. L. (1998). Comorbidity as a fundamental feature of generalized anxiety disorders: results from the National Comorbidity Study (NCS). *Acta Psychiatrica Scandinavica* (suppl. 393), 6–11.

Kang, H. K. & Bullman, T. (1998). Counterpoint: negligible 'Healthy-Warrior Effect' on Gulf War veterans' mortality. *American Journal of Epidemiology* **148**, 324–325.

Keane, T., Caddell, J. & Taylor, K. (1988). The Mississippi Scale for combat-related post-traumatic stress disorder: three studies in reliability and validity. *Journal of Consulting and Clinical Psychology* **65**, 85–90.

Larsen, G. E., Alderton, D. L., Neideffer, M. & Underhill, E. (1997). Further evidence on dimensionality and correlates of the Cognitive Failures Questionnaire. *British Journal of Psychology* **88**, 29–38.

Lezak, M. D. (ed.) 1995 *Neuropsychological Assessment, 3rd edn*. Oxford University Press: New York.

MacFarlane, G., Thomas, E. & Cherry, N. (2000). Mortality amongst United Kingdom Gulf war veterans. *Lancet* **356**, 17–21.

Nelson, H. E. (1991). *National Adult Reading Test* (*NART*). NFER-Nelson: Windsor.

Purcell, R., Maruff, P., Kyrios, M. & Pantelis, C. (1997). Neuropsychological function in young patients with unipolar major depression. *Psychological Medicine* **27**, 1277–1285.

Purdue Research Foundation (1948). *Purdue Pegboard Test*, Lafayette Instrument Co: Lafayette, IN.

Reitan, R. M. (1992). *Trail Making Test: Manual for Administration and Scoring*. Reitan Neuropsychology Laboratory: Tucson, AZ.

Robbins, J. M. & Kirmayer, L. J. (1991). Attributions of common somatic symptoms. *Psychological Medicine* **21**, 1029–1045.

Robertson, I. H., Manly, T., Andrade, J., Baddeley, B. T. & Yiend, J. (1997). 'Oops!': performance correlates of everyday attentional failures in traumatic brain injured and normal subjects. *Neuropsychologia* **35**, 757–768.

Roy, M., Koslowe, P. A., Kroenke, K. & Magruder, C. (1998). Signs, symptoms and ill-defined conditions in Persian Gulf war veterans: findings from the Comprehensive Clinical Evaluation Program. *Psychosomatic Medicine* **60**, 663–668.

Silanpaa, M. C., Agar, L. M., Milner, I. B., Podany, E. C., Axelrod, B. N. & Brown, G. G. (1997). Gulf war veterans: a neuropsychological examination. *Journal of Clinical and Experimental Neuropsychology* **19**, 211–219.

Southwick, S., Morgan, C. I., Nicolaou, A. & Charney, D. (1997). Consistency of memory for combat-related traumatic events in veterans of Operation Desert Storm. *American Journal of Psychiatry* **154**, 173–177.

Storzbach, D., Campbell, K. A., Binder, L. M., McCauley, L., Anger, W. K., Rohlman, D. S., Kovera, C. A. & Other members of the Portland Environmental Hazards Research Center (2000). Pyschological differences between veterans with and without Gulf War unexplained symptoms. *Psychosomatic Medicine* **62**, 726–735.

Suadicani, P., Ishøy, T., Guldager, B. & Appleyard, M. (1999). Determinants of long-term neuropsychological symptoms. *Danish Medical Bulletin* **46**, 423–427.

Trenerry, M. R., Crosson, B., Deboe, J. & Leber, W. R. (1989). *Stroop Neuropsychological Screening Test*. Psychological Assessment Resources Inc: Odessa, FL.

Unwin, C., Blatchley, N., Coker, W., Ferry, S., Hotopf, M., Hull, L., Ismail, K., Palmer, I., David, A. & Wessely, S. (1999). Health of UK servicemen who served in the Persian Gulf war. *Lancet* **353**, 169–182.

Unwin, C., Hotopf, M., Hull, L., Ismail, K., David, A. & Wessely, S. (2002). Women in the Gulf: lack of gender differences in long-term health effects of service in UK armed forces in the 1991 Gulf War. *Military Medicine* (in the press).

Vasterling, J. J., Brailey, K., Constans, J. I. & Sutker, P. B. (1998). Attention and memory dysfunction in post-traumatic stress disorder. *Neuropsychology* **12**, 125–133.

Wagle, A. C., Berrios, G. E. & Ho, L. (1999). The cognitive failures questionnaire in psychiatry. *Comprehensive Psychiatry* **40**, 478–484.

Wagner, A. W., Wolfe, J., Rotnitsky, A., Proctor, S. P. & Erickson, D. J. (2000). An investigation of the impact of posttraumatic stress disorder on physical health. *Journal of Traumatic Stress* **13**, 41–55.

Ware, J. E. J., Snow, K. K., Kosinski, M. & Gandek, B. (1993). *SF-36 Health Survey: Manual and Interpretation Guide*. B1–B5. The Health Institute, New England: Boston, MA.

Warrington, E. K. (1996). *The Camden Memory Tests Manual*. Psychology Press, Taylor and Francis Ltd: Chichester.

Wechsler, D. (1981). *The Wechsler Adult Intelligence Scale-Revised Manual*. The Psychological Corporation: San Antonio, TX.

Wechsler, D. (1987). *The Wechsler Memory Scale-Revised Manual*. The Psychological Corporation: San Antonio, TX.

Wechsler, D. (1997). *The WAIS-III*. The Psychological Corporation: San Antonio, TX.

Weingartner, H., Cohen, R. M. & Murphey, D. L. (1981). Cognitive processes in depression. *Archives of General Psychiatry* **38**, 42–47.

Wolfe, J., Proctor, S. P., Erickson, D. J., Heeren, T., Friedman, M. J., Huang, M. T., Sutker, P. B., Vasterling, J. J. & White, R. F. (1999). Relationship of psychiatric status to Gulf War veterans' health problems. *Psychosomatic Medicine* **61**, 532–540.

Yerkes, W. H. & Dodson, J. D. (1908). The relation between the strength of stimulus to rapidity of habit-formation. *Journal of Comparative Neurology and Psychiatry* **18**, 459–482.