

Stability of recall of military hazards over time

Evidence from the Persian Gulf War of 1991

S. WESSELY, C. UNWIN, M. HOTOFF, L. HULL, K. ISMAIL, V. NICOLAOU
and A. DAVID

Background Wartime traumatic events are related to subsequent psychological and physical health, but quantifying the association is problematic. Memory changes over time and is influenced by psychological status.

Aims To use a large, two-stage cohort study of members of the UK armed forces to study changes in recall of both traumatic and 'toxic' hazards.

Method A questionnaire-based follow-up study assessed 2370 UK military personnel, repeating earlier questions about exposure to military hazards.

Results The κ statistics for reporting of hazards were good for some exposures, but very low for others. Gulf veterans reported more exposures over time (no significant rise in the Bosnia cohort). In the Gulf cohort only, reporting new exposures was associated with worsening health perception, and forgetting previously reported exposures with improved perception. We found no association between physical health, psychological morbidity or post-traumatic stress disorder symptoms and endorsement or non-endorsement of exposures.

Conclusions Reporting of military hazards after a conflict is not static, and is associated with current self-rated perception of health. Self-report of exposures associated with media publicity needs to be treated with caution.

Declaration of interest None.
Funding detailed in Acknowledgements.

It is known that there is an association between traumatic event reporting and negative health outcomes, particularly post-traumatic stress disorder (PTSD) (Kaylor *et al*, 1987; Brewin *et al*, 2000). However, establishing the nature and magnitude of this association has been difficult, resulting in very different estimates. Most studies of the link between adversity and health are cross-sectional and rely on retrospective accounts of events and circumstances. There is consensus that retrospectively recalled accounts of trauma are problematic, and potentially subject to a variety of recall biases, but there is no consensus as to either the size of the problem or its implications (McFarlane, 1988).

Several studies have looked at the relationship between retrospective recall of exposures at several time points and their relationship to health outcomes, mainly PTSD. We have conducted a large-scale longitudinal study of the health of UK military personnel, based on three cohorts: those who saw service in the 1991 Persian Gulf conflict, those who were deployed on peacekeeping operations in Bosnia between 1991 and 1997, and those who were in the forces at the time of the 1991 Gulf conflict but were not deployed (Unwin *et al*, 1999). We now report on the results of a follow-up study, in which the same questions about specific military exposures related to the deployments on which the respondents had served were asked again. In this paper we examine the consistency of reporting of military traumas and hazards over the period. We also look at the predictors of any observed change in recall of traumatic events. In particular, we test the hypothesis that psychological distress prospectively increases the recall of traumatic events and hazards over time. This study is unique in that it permits the comparison of exposure consistency for both the Gulf War and the Bosnia deployments. We are thus able to compare the recall of military exposures relevant to both fighting and peacekeeping.

METHOD

A follow-up descriptive study examined the health status of a stratified sample of participants who had completed the first phase of the King's College London epidemiological health survey of military personnel (Unwin *et al*, 1999). The original study took place in 1997, which was 6 years after the end of the Gulf War, and 5 years after the start of the Bosnia deployment. This survey was succeeded by a series of detailed clinical case-control studies (stage 2: David *et al*, 2002; Higgins *et al*, 2002; Sharief *et al*, 2002). The follow-up study, stage 3, took place in 2000 and 2001, approximately 3 years later. During the follow-up study, participants were asked again about specific military exposures related to the deployments on which they had served. In this paper we compare responses between the two large epidemiological surveys of the same personnel at stage 1 and stage 3.

Participants

The target group was a stratified sample of the cohort who completed the stage 1 Health Survey of Military Personnel ($n=8195$). This cohort consisted of three groups: personnel who served in the Persian Gulf region between 1 September 1990 and 30 June 1991 (the Gulf Cohort); personnel who had served in Bosnia between 1 April 1992 and 6 February 1997 (the Bosnia cohort); and personnel who were serving in the armed forces on 1 January 1991 but who were not deployed to the Gulf conflict (the Era cohort). Special forces were excluded for security reasons. Two stratification variables were used: fatigue and gender.

We were primarily interested in examining the health of Gulf veterans. At stage 1, fatigue, along with being strongly associated with other health outcomes measured, was one of the most consistently reported symptoms in Gulf veterans, and was our *a priori* principal outcome measure, and the basis for the stratified sample strategy. In order to ensure that the most severely ill were well represented, all male veterans with a fatigue score of 9 or more (511 Gulf, 115 Bosnia and 120 Era were included). A 1:2 sample of male Gulf veterans with mid-range fatigue scores of 4–8 (484 veterans) along with all Bosnia veterans ($n=333$) and Era veterans ($n=364$) scoring in this range were selected. Finally, a 1:8 sample of veterans with fatigue scores less

than 4 was selected in order to represent asymptomatic individuals (250 in each group). All female veterans who completed the stage 1 questionnaire ($n=648$) were contacted, as women were oversampled in the original cohort. This also allowed us to look for any gender differences in follow-up variables. The total sample size was 3322.

Ethical approvals were obtained for all stages of the study. All respondents at stage 1 gave signed consent to later follow-up.

Questionnaire

The questionnaire mainly replicated the measures used at stage 1, including demographic details (age, gender, marital and educational status, alcohol and smoking habits), chronic fatigue scale, medical symptoms (50 items), self-reported medical disorders (39 items), the 12-item General Health Questionnaire (Goldberg, 1972), and the 36-item Medical Outcomes Study Short Form (SF-36) sub-scales for physical health, health perception and functional capacity (Stewart *et al*, 1988).

As detailed in the original study report (Unwin *et al*, 1999) we created a brief measure labelled 'post-traumatic stress reaction' (PTSR). This was embedded in the wider questionnaire because we did not wish to have an overt PTSD scale, given the social context of 'Gulf War syndrome' at the time among the UK service community and also because of the need to keep measures to a minimum. Full details of this are contained elsewhere, but in essence it consisted of four simple stem questions covering the basic psychopathological features of PTSD (Unwin *et al*, 1999).

Military exposure history was investigated using the same checklist as at stage 1, again tailored for the appropriate deployment. In practice this meant that the Bosnia and Era groups were not asked about the following exposures specific to the Gulf War: smoke from oil-well fires; mustard gas or other blistering agents; having a Scud missile explode in the air or on the ground within 1 mile; hearing chemical alarms sounding; and chemical/nerve gas attack. The questionnaires were tailored according to whether the participant was still in service or not, as ascertained at stage 1.

Analyses

The reliability of the responses for each exposure at the two time points was quantified by the kappa statistic, which is a

measure of the degree of non-random agreement between measurements of the same categorical variable. If the measurements agree more often than expected by chance, κ is positive; if agreement is complete, κ is 1; if they disagree more often than expected by chance, κ is negative (Last, 1995). A paired *t*-test was used to examine the number of endorsed exposures at the two time points.

We followed the same analytical approach to exposure measurements over two time points as Southwick *et al* (1997). For each of the exposures asked about, variables were created indicating whether the exposure was:

- (a) always endorsed at both time points (YY);
- (b) never endorsed at either time point (NN);
- (c) endorsed at time 1 but no longer endorsed at time 2, i.e. no longer endorsed (YN);
- (d) not endorsed at time 1, later to be endorsed, i.e. newly endorsed (NY).

Risk factors for number of newly endorsed and no longer endorsed exposures were explored by examining their median and interquartile range (IQR). Change in health status was examined by creating a change variable (stage 3 minus stage 1) for each health outcome. This was then used as the dependent variable when exploring the effect of newly endorsed and no longer endorsed exposures on health reporting over time. The data were analysed using the Statistical Package for the Social Sciences, version 10 for Windows.

RESULTS

Address information was not available for 13 of the original participants. Valid responses were obtained from 2370 (72%) participants: 907 Gulf (response rate 73.0%), 638 Bosnia (70.2%), 643 Era (69.5%) and 182 Bosnia and Gulf (78.4%). There were 246 (7.4%) refusers. Owing to the absence of an accurate address, 259 (7.8%) never received the questionnaire, despite three mailing attempts, giving a true rate of 78% (Gulf 79%; Bosnia 77%; Era 6.0%; Bosnia and Gulf 82%). For the purpose of analysis participants who had been deployed to both the Gulf and Bosnia were combined with the Gulf-only group, as in previous analyses of this cohort (Unwin *et al*, 1999, 2002; Ismail *et al*, 2000;

Reid *et al*, 2001). Table 1 gives the distribution of demographic variables in the two study cohorts.

The mean number of reported exposures significantly increased over time for the Gulf cohort (Table 2), but the increase in the Bosnia cohort was modest and non-significant. The Pearson correlation between the number of reported events at both time points was low for both the Gulf cohort

Table 1 Demographic variables for the Gulf and Bosnia cohorts

	Gulf		Bosnia	
	<i>n</i>	(%)	<i>n</i>	(%)
Gender				
Male	900	(82.6)	477	(74.8)
Female	189	(17.4)	161	(25.2)
Age (years)				
Under 29	0	0	194	(30.4)
30–34	325	(29.8)	221	(34.6)
35–39	321	(29.5)	109	(17.1)
40–44	224	(20.6)	72	(11.3)
Over 45	219	(20.1)	42	(6.6)
Marital status				
Married	726	(67.1)	358	(56.8)
Living with partner	108	(10.0)	64	(10.2)
Never married	122	(11.3)	133	(21.1)
Separated	53	(4.9)	29	(4.6)
Divorced	69	(6.4)	46	(7.3)
Widowed	4	(0.4)	0	0
Alcohol use (units/week)				
None	151	(14.0)	48	(7.6)
1–3	271	(25.1)	149	(23.6)
4–10	331	(30.6)	199	(31.5)
11–20	197	(18.2)	142	(22.5)
21–30	99	(9.2)	63	(10.0)
30+	31	(2.9)	30	(4.8)

Table 2 Changes in mean number of exposures over time reported by the Gulf and Bosnia cohorts

Stage	Exposures (mean)	<i>n</i>	<i>t</i>	<i>r</i>
Gulf				
1	10.7	1089		
3	11.5	1089	–5.7**	0.66**
Bosnia				
1	7.1	820		
3	7.3	820	–1.07 [†]	0.57**

** $P < 0.01$; [†] $P = 0.29$.

Table 3 Frequency of recall categories for each exposure in the Gulf and Bosnia cohorts

	Recall category ¹								κ
	YY		YN		NN		NY		
	<i>n</i>	(%)	<i>n</i>	(%)	<i>n</i>	(%)	<i>n</i>	(%)	
Gulf									
Smoke from oil-well fires	698	(70.8)	23	(2.3)	208	(21.1)	57	(5.8)	0.79
Handled POW	461	(46.8)	59	(6.0)	382	(38.8)	83	(8.4)	0.71
Small arms fire	165	(16.8)	40	(4.1)	708	(72.0)	70	(7.1)	0.68
Scud missile exploding within 1 mile	241	(25.2)	66	(6.9)	574	(60.0)	74	(7.7)	0.67
Dismembered bodies	543	(54.9)	90	(9.1)	280	(28.3)	76	(7.7)	0.64
Burning rubbish/faeces	501	(51.7)	81	(8.4)	286	(29.5)	101	(10.4)	0.61
Maimed soldiers	405	(41.1)	85	(8.6)	382	(38.7)	114	(11.6)	0.60
Diesel/petrochemical fuel on skin	451	(46.8)	80	(8.3)	318	(33.0)	113	(11.7)	0.59
Witnessed anyone dying	131	(13.2)	58	(5.8)	712	(71.5)	94	(9.4)	0.54
Dead animals	260	(26.4)	104	(10.6)	496	(50.4)	124	(12.6)	0.51
Combat-related injury	38	(3.8)	28	(2.8)	899	(89.8)	36	(3.6)	0.51
NBC suits	817	(83.2)	39	(4.0)	62	(6.3)	64	(6.5)	0.49
Heard chemical alarms sounding	677	(69.8)	67	(6.9)	124	(12.8)	102	(10.5)	0.49
Depleted uranium	87	(10.2)	18	(2.1)	632	(74.0)	117	(13.7)	0.48
Chemical/nerve gas attack	54	(6.2)	28	(3.2)	718	(82.7)	66	(7.6)	0.48
Other paints/solvents	396	(41.9)	82	(8.7)	290	(30.7)	176	(18.6)	0.45
Pesticides on clothing/bedding	251	(26.6)	84	(8.9)	440	(46.7)	166	(17.6)	0.45
Food contaminated with smoke	139	(15.0)	72	(7.8)	588	(63.4)	127	(13.7)	0.44
Artillery close by	179	(18.3)	97	(9.9)	566	(57.8)	136	(13.9)	0.44
Personal pesticides	518	(54.5)	83	(8.7)	189	(19.9)	160	(16.8)	0.42
Local food	448	(47.3)	124	(13.1)	236	(24.9)	139	(14.7)	0.42
Exhaust from heaters/generators	623	(64.3)	63	(6.5)	128	(13.2)	155	(16.0)	0.40
Diesel/petrochemical fumes	690	(71.7)	55	(5.7)	93	(9.7)	124	(12.9)	0.40
Heat illness	110	(11.1)	115	(11.6)	678	(68.5)	87	(8.8)	0.39
CARC paint	163	(17.7)	61	(6.6)	500	(54.2)	198	(21.5)	0.37
Bathed in local pond/river	27	(2.7)	38	(3.8)	885	(89.3)	40	(4.0)	0.37
Bathed in/drank local water	57	(6.0)	57	(6.0)	753	(78.9)	87	(9.1)	0.36
Mustard gas	2	(0.2)	12	(1.3)	876	(95.5)	25	(2.7)	0.08
Bosnia									
Small arms fire	234	(34.7)	57	(8.4)	311	(46.1)	73	(10.8)	0.61
Witnessed anyone dying	85	(12.5)	28	(4.1)	508	(74.6)	60	(8.8)	0.58
Artillery close by	127	(18.6)	50	(7.3)	440	(64.5)	65	(9.5)	0.57
Dismembered bodies	193	(28.5)	66	(9.7)	337	(49.8)	81	(12.0)	0.55
Maimed soldiers	147	(21.7)	57	(8.4)	390	(57.5)	84	(12.4)	0.52
Diesel/petrochemical fuel on skin	308	(45.9)	59	(8.8)	202	(30.1)	102	(15.2)	0.51
Bathed in local pond/river	89	(13.2)	60	(8.9)	462	(68.5)	63	(9.3)	0.47
Heat illness	49	(7.2)	51	(7.5)	540	(79.3)	41	(6.0)	0.44
Dead animals	235	(34.8)	67	(9.9)	246	(36.4)	128	(18.9)	0.43
Bathed in/drank local water	41	(6.1)	40	(5.9)	545	(81.0)	47	(7.0)	0.41
Burning rubbish/faeces	274	(41.0)	100	(14.9)	198	(29.6)	97	(14.5)	0.40
Food contaminated with smoke	41	(6.1)	26	(3.9)	537	(79.8)	69	(10.3)	0.39
Local food	362	(54.1)	72	(10.8)	127	(19.0)	108	(16.1)	0.39
Other paints/solvents	255	(38.1)	66	(9.9)	191	(28.5)	158	(23.6)	0.34
Handled POW	116	(17.3)	78	(11.7)	363	(54.3)	112	(16.7)	0.34
Exhaust from heaters/generators	458	(68.0)	55	(8.2)	64	(9.5)	97	(14.4)	0.32
Pesticides on clothing/bedding	82	(12.3)	50	(7.5)	405	(60.8)	127	(19.1)	0.31
Diesel/petrochemical fumes	434	(64.5)	48	(7.1)	69	(10.3)	120	(17.8)	0.30

(continued)

Table 3 (continued)

	Recall category ¹								κ
	YY		YN		NN		NY		
	n	(%)	n	(%)	n	(%)	n	(%)	
Bosnia (continued)									
Personal pesticides	206	(31.0)	77	(11.6)	217	(32.7)	163	(24.5)	0.29
Combat-related injury	14	(2.0)	27	(4.0)	596	(87.3)	46	(6.7)	0.22
CARC paint	18	(2.7)	31	(4.7)	538	(81.5)	73	(11.1)	0.18
NBC suits	3	(0.4)	13	(1.9)	606	(90.6)	47	(7.0)	0.06
Depleted uranium	1	(0.0)	10	(0.5)	618	(98.4)	17	(0.8)	0.05

CARC, chemical agent resistant coating; NBC, nuclear, biological and chemical; POW, prisoners of war.

1. YY, endorsed at both time points; NN, not endorsed at either time point; YN, no longer endorsed; NY, not endorsed at time 1, endorsed at time 2.

2. Denominator used for percentages is total n for each exposure.

and Bosnia cohort ($r=0.66$ and 0.57 respectively). Table 3 shows the percentage responses of YY, NN, YN and NY, along with the κ values, for each of the variables in the questionnaires given to the Gulf and Bosnia cohorts. In the Gulf cohort the most reliably recalled exposures were: smoke from oil-well fires ($\kappa=0.79$); handled prisoners of war ($\kappa=0.71$); small arms fire ($\kappa=0.68$); Scud missile exploding within 1 mile ($\kappa=0.67$); and seeing dismembered bodies ($\kappa=0.64$). For the Bosnia

cohort the most reliably recalled exposures were: small arms fire ($\kappa=0.61$); witnessing anyone dying ($\kappa=0.58$); artillery close by ($\kappa=0.57$); seeing dismembered bodies ($\kappa=0.55$); and diesel or petrochemical fuel on skin. For the 24 exposures common to the two cohorts, the Gulf cohort had higher κ values for all except 4 (bathed in/drank local water; bathed in local pond/river; heat illness; and witnessed anyone dying).

On average the Gulf cohort had more newly endorsed (NY) than no longer

endorsed (YN) exposures (mean NY=2.90, s.d.=2.39; mean YN=1.80, s.d.=1.75), a pattern repeated in the Bosnia cohort (mean NY=2.89, s.d.=2.36; mean YN=1.74, s.d.=1.78) and indicating an overall rise in the number of exposures recalled over time (Table 4). Table 5 gives the risk factors for numbers of newly endorsed and no longer endorsed exposures. For the number of no longer endorsed items, the most significant risk factor was serving status, with those in service having a higher median value for no longer endorsed exposures than those not in service. This pattern held for both the Gulf and Bosnia cohorts. For the newly endorsed exposures over time, being male and younger were associated with higher median values for both the Gulf and Bosnia cohorts, whereas living with a partner was associated with a higher median value in the Bosnia cohort only.

Table 6 shows the mean changes in health outcomes for the Bosnia and Gulf cohorts and their association with newly endorsed or no longer endorsed exposures. For the purpose of these analyses, the no longer endorsed (YN) and newly endorsed (NY) exposure recall variables have been recoded to combine the tailed distribution into one group. There was a pattern of increased (i.e. improved) health perception and increased no longer endorsed (i.e. forgotten) exposures over time in the Gulf cohort, which was not replicated in the Bosnia cohort. Conversely, there was a pattern of worsening health perception and increasing new endorsement of exposure variables over time in the Gulf cohort but not the Bosnia cohort. There was no discernible pattern of association between

Table 4 Frequency of newly endorsed ('no to yes') and no longer endorsed ('yes to no') exposure recall in the Gulf and Bosnia cohorts

Number of changes	Gulf				Bosnia			
	Yes to no		No to yes		Yes to no		No to yes	
	n	(%)	n	(%)	n	(%)	n	(%)
0	260	(25.8)	131	(13.0)	195	(28.4)	93	(13.6)
1	261	(25.9)	204	(20.3)	177	(25.8)	123	(17.9)
2	209	(20.8)	183	(18.2)	142	(20.7)	145	(21.1)
3	118	(11.7)	151	(15.0)	70	(10.2)	100	(14.6)
4	89	(8.8)	123	(12.2)	46	(6.7)	79	(11.5)
5	33	(3.3)	86	(8.5)	27	(3.9)	58	(8.5)
6	20	(2.0)	46	(4.6)	14	(2.0)	28	(4.1)
7	6	(0.6)	32	(3.2)	7	(1.0)	21	(3.1)
8	6	(0.6)	24	(2.4)	4	(0.6)	19	(2.8)
9	1	(0.1)	10	(0.4)	2	(0.3)	12	(1.7)
10	2	(0.2)	8	(0.8)	2	(0.3)	1	(0.1)
11	–	–	3	(0.3)	–	–	5	(0.7)
12	–	–	1	(0.1)	–	–	1	(0.1)
13	–	–	1	(0.1)	–	–	1	(0.1)
14	–	–	1	(0.1)	–	–	–	–
15	–	–	2	(0.2)	–	–	–	–
16	1	(0.1)	–	–	–	–	–	–

Table 5 Association between demographic factors and exposure change variables (YN, no longer endorsed; NY, newly endorsed)

	Gulf							Bosnia						
	n	YN			NY			n	YN			NY		
		Median	IQR	P	Median	IQR	P		Median	IQR	P	Median	IQR	P
Gender				0.35 ¹			0.02 ¹				0.17 ¹			<0.01 ¹
Male	839	1.0	3.00		3.0	3.00		552	1.0	3.00		3.0	3.00	
Female	167	1.0	3.00		2.0	3.00		134	1.0	2.00		2.0	2.00	
Age (years)				0.12 ²			0.01 ²				0.90 ²			<0.01 ²
Under 29	0	–	–		–	–		168	1.0	2.00		3.0	3.00	
30–34	306	1.0	3.00		3.0	3.00		272	1.0	3.00		2.0	3.00	
35–39	299	1.0	2.00		3.0	3.00		117	2.0	3.00		2.0	3.00	
40–44	206	1.0	2.00		2.0	3.00		87	1.0	1.00		2.0	3.00	
Over 45	195	2.0	2.00		2.0	3.00		42	1.5	3.00		2.0	3.25	
Marital status				0.67 ²			0.09 ²				0.55 ²			0.04 ²
Married	667	1.0	3.00		2.0	3.00		407	1.0	3.00		2.00	3.00	
Living with partner	105	2.0	2.00		3.0	4.00		76	1.0	2.00		3.0	3.00	
Never married	112	1.0	2.00		2.0	3.00		118	1.0	2.00		2.0	4.00	
Separated	50	2.0	2.00		3.0	4.00		32	1.0	1.00		2.0	2.75	
Divorced	61	2.0	2.50		3.0	3.00		44	1.0	3.00		2.0	3.00	
Widowed	4	1.5	2.50		3.0	4.25		0	–	–		–	–	
Serving status				0.03 ¹			0.41 ¹				0.00 ¹			0.68 ¹
Still serving	347	2.0	2.00		3.0	3.00		416	2.0	3.00		2.0	3.00	
Not serving	642	1.0	2.00		2.0	3.00		251	1.0	2.00		2.0	3.00	

IQR, interquartile range.

1. Mann–Whitney *U* test for rank difference.

2. Kruskal–Wallis test for rank difference.

physical health, psychological morbidity or PTSD symptoms (PTSR) and endorsement or non-endorsement of exposures, for either the Gulf or the Bosnia cohort. These analyses were repeated for the Gulf cohort, omitting the five Gulf-specific exposures (smoke from oil-well fires; mustard gas or other blistering agents; having a Scud missile explode in the air or on the ground within 1 mile; chemical/nerve gas attack; and hearing chemical alarms sounding), with no difference in findings (Table 7).

Table 8 shows the regression analyses results for the effects of newly endorsed ('newly remembered') and no longer endorsed ('forgotten') exposures on health, controlling for age, gender and number of endorsed exposures at stage 1. For the Gulf cohort, the total of newly endorsed exposures was associated with a reduction in health perception and increased psychological morbidity, whereas the total of no longer endorsed exposures was significantly associated with improved health perception. This pattern was not

replicated in the Bosnia cohort (data not shown).

DISCUSSION

We already know that there is poor agreement between reporting of military events and contemporaneous records of the same events (Keane *et al*, 1989). In an ideal world we would have objective, independent, contemporary records of exposures and hazards, but this is rarely (if ever) possible, given the 'friction' of war, and the impossibility of monitoring all hazards, both known and unknown, at the time. For that reason it is likely that self-report of hazards and exposures will continue to be the basis of the assessment of the consequences of war and military deployments for the foreseeable future.

Roemer *et al* (1998) documented consistent increases in reports of exposure to seven specific war-related stressors over time in a sample of 460 service personnel deployed to Somalia in a peacekeeping

operation. These men and women were assessed in the first year after their return, and then 1–3 years later. At the second assessment PTSD symptoms uniquely contributed to reported exposure scores. Southwick *et al* (1997) administered on two occasions a 19-item war zone exposure questionnaire and the Mississippi Scale for Combat Related PTSD to 59 members of the National Guard who had been activated for Gulf War duty. They analysed the extent that recall and forgetting of exposures altered over time, and found that the number of 'no' to 'yes' changes was significantly and positively related to PTSD symptom severity at the later assessment. In contrast, Bramsen and colleagues, in a study of Dutch peacekeepers, did not find either an increase in reported items over time, or an association between number of changes between the first and second assessments and symptoms of PTSD (Bramsen *et al*, 2001). Meanwhile, other researchers have used the experience of the Gulf War to study consistency of

Table 6 Mean change in health outcomes categorised by no longer endorsed and newly endorsed exposures for the Bosnia and Gulf cohort

	Health perception ¹		Physical health ¹		GHQ ²		PTSR ³	
	Mean	(s.d.)	Mean	(s.d.)	Mean	(s.d.)	Mean	(s.d.)
Gulf								
Total	0.8	(20.90)	-1.7	(16.00)	-0.8	(5.65)	-0.2	(1.42)
No longer endorsed (YN)								
0	-0.9	(20.71)	-3.2	(17.42)	-0.7	(6.12)	0.008	(1.39)
1	0.6	(21.83)	-1.2	(14.54)	-0.9	(5.75)	-0.2	(1.49)
2	0.2	(20.40)	-1.2	(14.67)	-1.0	(5.17)	-0.3	(1.34)
3	1.2	(20.01)	-2.5	(14.12)	-0.4	(5.03)	-0.2	(1.33)
4 or 5	1.6	(20.32)	0.6	(19.72)	-1.5	(5.64)	-0.5	(1.46)
6+	13.8	(19.96)	-2.3	(13.31)	0.0	(6.06)	-0.4	(1.44)
P ⁴	0.01		0.33		0.67		0.01	
Newly endorsed (NY)								
0	3.8	(22.82)	-2.1	(15.92)	-1.9	(5.29)	-0.4	(1.31)
1	2.3	(20.94)	-0.03	(15.58)	-0.9	(5.23)	-0.4	(1.38)
2	2.47	(19.07)	-1.7	(14.92)	-1.5	(5.03)	-0.4	(1.51)
3	-1.16	(22.54)	-2.2	(18.97)	-0.5	(6.34)	-0.01	(1.43)
4 or 5	-1.27	(18.86)	-2.1	(15.20)	0.004	(5.90)	0.05	(1.33)
6+	-1.55	(22.04)	-2.6	(15.60)	-0.6	(6.01)	-0.02	(1.50)
P ⁴	<0.01		0.14		0.01		<0.01	
Bosnia								
Total	-2.07	(22.56)	-2.97	(15.76)	-0.7	(5.99)	0.04	(1.6)
No longer endorsed (YN)								
0	-1.8	(22.27)	-5.1	(18.09)	-0.4	(6.02)	0.3	(1.64)
1	-2.1	(22.20)	-3.1	(15.36)	-0.7	(6.13)	0.04	(1.63)
2	-2.2	(22.47)	0.7	(13.19)	-1.2	(6.34)	-0.1	(1.33)
3	-0.7	(24.19)	-3.4	(14.66)	-1.1	(5.15)	-0.3	(1.70)
4 or 5	-4.4	(21.50)	-3.1	(15.12)	-0.007	(5.25)	0.2	(1.61)
6+	-0.5	(26.74)	2.5	(15.40)	-2.1	(6.94)	-0.7	(1.87)
P ⁴	0.56		0.03		0.11		0.01	
Newly endorsed (NY)								
0	0.7	(20.56)	-0.5	(13.79)	-1.3	(4.64)	-0.3	(1.59)
1	1.5	(22.98)	-3.0	(16.20)	-0.6	(5.74)	-0.05	(1.53)
2	-4.1	(21.03)	-1.6	(12.32)	-0.8	(6.07)	-0.03	(1.48)
3	-5.7	(24.83)	-3.6	(17.57)	-0.6	(5.83)	0.2	(1.57)
4 or 5	-4.4	(22.08)	-4.9	(18.29)	-0.5	(6.28)	0.2	(1.62)
6+	1.4	(23.49)	-3.9	(15.74)	-0.7	(7.24)	0.2	(1.89)
P ⁴	0.17		0.31		0.21		0.05	

1. Positive value indicates improvement in health.

2. General Health Questionnaire (GHQ): positive value indicates increase in psychological morbidity.

3. Post-traumatic stress reaction (PTSR): positive value indicates increase in symptoms.

4. Non-parametric test for trend.

reports of hazardous 'toxic' exposures over time (McCauley *et al*, 1999; Wolfe *et al*, 2002), but these studies did not consider the influence of psychological variables on changing patterns of recall.

We now consider two major findings. The first relates to the stability of recall of military hazards, and the second concerns

the direction of any observed changes and the influence of psychosocial factors on those changes.

Stability of recall

There was relatively low agreement for reporting of war exposures over time, as

shown by the majority of exposures having a κ under 0.6. In general our findings are very similar to those of the only study that used a similar design to look at consistency of recall in smaller numbers of US Gulf War veterans (McCauley *et al*, 1999). When the questions that we asked were almost identical to those asked in the US survey, the consistency of recall was likewise similar. Hence, in both studies, reporting exposure to smoke from oil fires was associated with good reliability, hearing Scuds detonate was also reasonably reliable, being aware of chemical alarms sounding was moderately reliable as was believing oneself exposed to chemical attack, whereas reporting drinking local water, exposure to chemical agent resistant coating (CARC) paint and exposure to depleted uranium was very unreliable in both studies. The recall of depleted uranium exposure is particularly problematic in the Bosnia cohort, indicated by the lowest κ value (0.05). Perhaps this is a reflection of the enormous publicity given to reports of cancers occurring in peacekeepers from several European nations that happened between the two phases of our study. Likewise, chemical exposures such as CARC paint, other paints/solvents and pesticides on clothing/bedding were associated with the greatest number of 'no' to 'yes' changes (increased recall) in both Gulf and Bosnia cohorts, and correspondingly low κ values. There has also been intense media concern over all these exposures in the British press in the past decade, including, but not restricted to, the military context.

Change in recall over time

Looking now at the general pattern of change, previous studies have shown that the mean number of events reported over time can either increase (Southwick *et al*, 1997; Roemer *et al*, 1998; King *et al*, 2000) or stay the same (Bramsen *et al*, 2001). Our study produced an increase in the number of events reported over time in the Gulf cohort, but no significant increase in the Bosnia cohort. What this illustrates is the importance of not assuming that all conflicts are the same in terms of their social and psychological impact. Results in our peacekeeping cohort are similar to those of Bramsen and colleagues looking at Dutch peacekeepers, and our Gulf results, although different

Table 7 Mean change in health outcomes categorised by newly endorsed and no longer endorsed exposures for generic military exposures in the Gulf cohort

	Health perception ¹		Physical health ¹		GHQ ²		PTSR ³	
	Mean	(s.d.)	Mean	(s.d.)	Mean	(s.d.)	Mean	(s.d.)
Total	0.7	(20.93)	-1.6	(15.71)	-0.8	(5.80)	-0.2	(1.44)
No longer endorsed (YN)								
0	-0.3	(21.23)	-2.7	(16.32)	-0.7	(6.52)	-0.02	(1.47)
1	0.4	(20.88)	-0.8	(15.24)	-0.7	(5.44)	-0.1	(1.45)
2	-0.3	(20.55)	-0.8	(14.18)	-0.9	(5.18)	-0.3	(1.35)
3	2.3	(19.71)	-2.4	(13.72)	-0.8	(5.53)	-0.3	(1.30)
4 or 5	2.1	(21.26)	-1.3	(19.88)	-1.0	(5.59)	-0.6	(1.58)
6+	18.3	(19.30)	2.3	(10.68)	-1.6	(5.66)	-0.4	(1.09)
P ⁴	0.01		0.38		0.47		<0.01	
Newly endorsed (NY)								
0	2.8	(22.25)	-1.1	(14.71)	-1.2	(6.07)	-0.3	(1.43)
1	2.0	(21.22)	-0.7	(14.99)	-1.4	(5.45)	-0.4	(1.48)
2	1.7	(18.64)	-1.1	(15.42)	-0.6	(5.37)	-0.2	(1.40)
3	-1.3	(21.31)	-3.3	(18.51)	-0.3	(5.94)	-0.1	(1.49)
4 or 5	-1.4	(20.07)	-3.0	(16.18)	-0.3	(6.09)	-0.05	(1.32)
6+	-1.7	(22.02)	-1.3	(14.91)	-0.9	(6.03)	-0.2	(1.52)
P ⁴	<0.01		0.06		0.12		0.01	

1. Positive value indicates improvement in health.

2. General Health Questionnaire (GHQ): positive value indicates increase in psychological morbidity.

3. Post-traumatic stress reaction (PTSR): positive value indicates increase in symptoms.

4. Non-parametric test for trend.

from those in the peacekeepers, are similar to the findings of Southwick and colleagues in US Gulf veterans. On the other hand, neither we nor Bramsen *et al* (2001) are able to confirm the substantial increase in reporting of events recorded by Roemer *et al* (1998) in US peacekeepers in Somalia. However, the US operation in Somalia was beset by difficulties, and involved rather more than peacekeeping, with periods of actual combat.

A second reason why the literature is not entirely consistent is that previous studies have been concerned with either post-traumatic type events and symptoms, or more 'toxic' hazards, but rarely with both. In our Gulf studies we have always taken a broader view of hazards and exposures, incorporating exposures such as vaccinations, smoke from oil fires, depleted uranium and so on, which are not traumatic in the customary use of the word, but certainly came to prominence after the 1991 Gulf War. This means that we included more measures of these kinds of hazards than the other studies, but

conversely our measure of post-traumatic stress symptoms is less sophisticated.

Recall of exposures and current health perception

We found an association between health perception and both increased reporting and also forgetting of exposures, but this was not true of psychological morbidity or physical health. This association held for the Gulf cohort, but not for the Bosnia cohort. In general we found that the main pattern of change was of increased reporting ('no' to 'yes') rather than forgetting ('yes' to 'no').

There may be several explanations for changes in reporting of an event over time. The recall of events might simply become inflated over time; conversely, individuals might have underestimated their reports initially and later given more accurate appraisals. Over the specific interval of this study, there was considerable media attention to the Gulf War and its possible health effects. No doubt this information was incorporated to some lesser or greater

degree into the participants' perspectives on their experiences of the war. The acquisition of new knowledge, from whatever source, could explain both types of changes in item endorsement: elucidating and clarifying events and circumstances that did occur but were not previously known (accounting for 'no' to 'yes' changes), and delimiting details of experiences previously held to be true (accounting for 'yes' to 'no' changes). We should also be careful not to assume that changes in reporting equate with changes in memory – it might be that events previously seen as irrelevant and not endorsed on a questionnaire have increased in importance and salience over time, perhaps because of media coverage, rather than being newly remembered. On the other hand, there is evidence from this study that the reporting of events is influenced by current health perception. We found an association between changes in endorsement – both positive and negative – of hazards, and current health perception. Remembering more exposures over time was associated with worsening perception of health; conversely, improved perception of health was associated with forgetting previously recalled exposures. This pattern held after we had removed exposures specific to the Gulf War from the analysis, indicating that the finding was not due to certain key Gulf War exposures that might be strongly associated with health. This finding was not replicated with measures of mental health in general, or PTSD symptoms in particular.

Another important finding is that the association between health perception and recall or non-recall of hazards and exposures was found in the Gulf cohort but not in the Bosnia cohort. This was not because of a differential effect of exposures only encountered in the Gulf and also the subject of intense media scrutiny, since removing these Gulf-specific exposures did not alter the association.

We draw attention to the finding that, contrary to our original predictions, change in recall of exposures was not associated with changes in post-traumatic stress symptoms, but with health perception. However, this is in keeping with our own nested case-control study in which we interviewed both ill and healthy veterans, this time using a standardised psychiatric interview to enable us to make firm diagnoses of PTSD, which was not the case in the epidemiological study. Although psychological morbidity was increased, modestly, in the Gulf

Table 8 Prediction of change in health status and psychological morbidity by newly endorsed and no longer endorsed exposure recall: hierarchical regression analysis controlling for age, gender and exposure at time I for the Gulf cohort

Predictors	R	R ²	Step 1			Step 2			Step 3		
			B	s.e.	β	B	s.e.	β	B	s.e.	β
Health perception and newly endorsed (NY)											
Step 1	0.07	0.005									
Age			-0.21	0.10	-0.07*	-0.20	0.01	-0.07*	-0.25	0.10	-0.08*
Gender			-0.22	1.77	-0.00	0.12	1.85	0.00	-0.64	1.86	-0.01
Step 2	0.07	0.00									
Time I exposure						0.009	0.14	0.02	0.001	0.14	0.00
Step 3	0.12**	0.010**									
NY									-0.90	0.29	-0.10**
Health perception and no longer endorsed exposures (YN)											
Step 1	0.07	0.005									
Age			-0.21	0.10	-0.07*	-0.20	0.10	-0.07*	-0.24	0.10	-0.08*
Gender			-0.22	1.78	-0.00	0.12	1.85	0.00	-0.30	1.84	-0.00
Step 2	0.07	0.000									
Time I exposure						0.009	0.14	0.02	-0.008	0.15	-0.02
Step 3	0.13**	0.011**									
YN									1.36	0.40	0.11*
Psychological morbidity and newly endorsed exposures (NY)											
Step 1	0.03	0.001									
Age			0.01	0.03	0.01	0.007	0.03	0.009	0.02	0.03	0.02
Gender			0.44	0.49	0.03	0.34	0.51	0.02	0.47	0.51	0.03
Step 2	0.04	0.002									
Time I exposure						0.03	0.04	-0.02	0.01	0.04	-0.01
Step 3	0.08	0.006									
NY									0.16	0.08	0.07*

* $P < 0.05$; ** $P < 0.01$.

cohort, this was rarely due to PTSD (Ismail *et al.*, 2002). This is unsurprising. The Gulf War was not particularly traumatic in the conventional sense for the coalition forces, particularly in the context of past military campaigns that were associated with high rates of classic war-related psychiatric injury. The perceived hazards of the Gulf tended to be those not usually associated with the military setting, and not encapsulated in the formulations of PTSD. Instead, most revolved around fears of environmental exposure and contamination. We have argued elsewhere that the indisputable increase in ill health seen after the Gulf War is better understood as part of the literature on unexplained symptoms and syndromes, rather than conventional PTSD. This may

help to explain why changes in recall of exposures were associated more with changes in health perception than with symptoms of post-traumatic stress.

ACKNOWLEDGEMENTS

This study was funded by the US Department of Defense and the UK Medical Research Council. Neither organisation has had any input into the design, conduct, analysis or reporting of the study. The views expressed are the authors' own.

We thank all the servicemen and service women who gave freely of their time. We also thank Dr Inge Bramsen and Professor Ariel Shalev for their helpful advice.

REFERENCES

- Bramsen, I., Dirkzwager, A., Van Esch, S., *et al* (2001) Consistency of self reports of traumatic events in a population of Dutch peacekeepers: reason for optimism. *Journal of Traumatic Stress*, **14**, 733–740.
- Brewin, C., Andrews, B. & Valentine, J. (2000) Meta-analysis of risk factors for posttraumatic stress disorder in trauma exposed adults. *Journal of Consulting and Clinical Psychology*, **68**, 748–766.
- David, A., Farrin, L., Hull, L., *et al* (2002) Cognitive functioning and disturbances of mood in UK veterans of the Persian Gulf War: a comparative study. *Psychological Medicine*, **32**, 1357–1360.
- Goldberg, D. (1972) *The Detection of Psychiatric Illness by Questionnaire*. London: Oxford University Press.
- Higgins, E., Ismail, K., Kant, K., *et al* (2002) Skin disease in Gulf War Veterans. *Quarterly Journal of Medicine*, **95**, 67–676.

Ismail, K., Blatchley, N., Hotopf, M., et al (2000)

Occupational risk factors for ill health in UK Gulf war veterans. *Journal of Epidemiology and Community Health*, **54**, 834–838.

—, **Kent, K., Brugha, T., et al (2002)** The mental health of UK Gulf War veterans: phase 2 of a two-phase cohort study. *BMJ*, **325**, 576–579.

Kaylor, J., King, D. & King, L. (1987) Psychological effects of military service in Vietnam: a meta analysis. *Psychological Bulletin*, **102**, 257–271.

Keane, T., Fairbank, J., Caddell, J., et al (1989) Clinical evaluation of a measure to assess combat exposure. *Journal of Consulting and Clinical Psychology*, **1**, 53–55.

King, D., King, L., Erickson, D., et al (2000) Post traumatic stress disorder and retrospectively reported stressor exposure: a longitudinal prediction model. *Journal of Abnormal Psychology*, **109**, 624–633.

Last, J. (1995) *Dictionary of Epidemiology*. Oxford: Blackwell Scientific.

McCauley, L., Joos, S., Spencer, P., et al (1999) Strategies to assess validity of self reported exposures during the Persian Gulf War. *Environmental Research Section A*, **81**, 195–205.

McFarlane, A. C. (1988) The longitudinal course of posttraumatic morbidity. The range of outcomes and their predictors. *Journal of Nervous and Mental Disease*, **176**, 30–3.

Reid, S., Hotopf, M., Hull, L., et al (2001) Multiple chemical sensitivity and chronic fatigue syndrome in UK Gulf War veterans. *American Journal of Epidemiology*, **153**, 604–609.

Roemer, L., Litz, B., Orsillo, S., et al (1998) Increases in retrospective accounts of war-zone exposure over time: the role of PTSD symptom severity. *Journal of Traumatic Stress*, **11**, 597–605.

Sharief, M., Delamont, R., Rose, M., et al (2002) Neurophysiologic evaluation of neuromuscular symptoms in UK Gulf War veterans. A controlled study. *Neurology*, **59**, 1518–1525.

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

CLINICAL IMPLICATIONS

- Stability of recall of hazardous exposures during military operations differs according to the nature of the exposure – those extensively publicised in the media are particularly problematic.
- The number of exposures recalled has increased with the passage of time.
- The total number of hazardous exposures reported may reflect not only actual exposures but also current distress.

LIMITATIONS

- All military conflicts differ in various ways – it cannot be assumed that these findings extrapolate beyond the 1991 Gulf War.
- The measurement of self-reported exposures was fairly crude.
- Measurement of post-traumatic symptoms was by self-report and not by structured interview.

S. WESSELY, FRCP, C. UNWIN, MSc, M. HOTOPF, PhD, L. HULL, BSc, K. ISMAIL, MRCPsych, V. NICOLAOU, BSc, A. DAVID, MD, King's Centre for Military Health Research, King's College London, UK

Correspondence: Professor S. Wessely, Department of Psychological Medicine, GKT School of Medicine, 103 Denmark Hill, London SE5 8AF, UK. E-mail: s.wessely@iop.kcl.ac.uk

(First received 15 January 2003, final revision 22 April 2003, accepted 6 May 2003)

Stewart, A., Hays, R. & Ware, J. (1988) The MOS Short-form General Health Survey: reliability and validity in a patient population. *Medical Care*, **26**, 724–732.

Unwin, C., Blatchley, N., Coker, W., et al (1999) The health of United Kingdom servicemen who served in the Persian Gulf War. *Lancet*, **353**, 169–178.

—, **Hull, L., Ismail, K., et al (2002)** Women in the Persian Gulf: lack of gender differences in long-term health effects of service in United Kingdom Armed Forces in the 1991 Persian Gulf War. *Military Medicine*, **167**, 406–413.

Wolfe, J., Proctor, S., Erickson, D., et al (2002) Risk factors for multisymptom illness in US Army veterans of the Gulf War. *Journal of Occupational and Environmental Medicine*, **44**, 271–281.