2009

Neuropsychological functioning of U.S. Gulf War veterans 10 years after the war

Rosemary Toomey
Boston University

Renee Alpern
Hines Veterans Affairs Hospital

Jennifer J. Vasterling
Tulane University of Louisiana

Dewleen G. Baker
University of Cincinnati - Main Campus

Domenic J. Reda
Hines Veterans Affairs Hospital

See next page for additional authors

Journal of the International Neuropsychological Society (2009), **15**, 717–729.

doi:10.1017/S1355617709990294

Neuropsychological functioning of U.S. Gulf War veterans 10 years after the war

ROSEMARY TOOMEY,[1,2,3,4] RENEE ALPERN,[5] JENNIFER J. VASTERLING,[6,7] DEWLEEN G. BAKER,[8,9] DOMENIC J. REDA,[5] MICHAEL J. LYONS,[1,2,3,4] WILLIAM G. HENDERSON,[5] HAN K. KANG,[10] SETH A. EISEN,[11,12,13] AND FRANCES M. MURPHY[14,15]

[1]Research Service, Boston VA Healthcare System, Boston, Massachusetts
[2]The Massachusetts Mental Health Center Public Academic Psychiatry Division of the Beth Israel Deaconess Medical Center, Department of Psychiatry, Harvard Medical School, Department of Psychiatry, Boston, Massachusetts
[3]Harvard Institute of Psychiatric Epidemiology and Genetics, Boston, Massachusetts
[4]Psychology Department, Boston University, Boston, Massachusetts
[5]Cooperative Studies Program Coordinating Center, Hines Veterans Affairs Hospital, Hines, Illinois. (W.G.H. is currently at the University of Colorado Health Outcomes Program and the Department of Preventive Medicine and Biometrics, Aurora, Colorado.)
[6]VA South Central (VISN 16) Mental Illness Research, Education, and Clinical Center, New Orleans Veterans Affairs Medical Center, New Orleans, Louisiana
[7]Department of Psychiatry and Neurology, Tulane University School of Medicine, New Orleans, Louisiana. (Author is currently at the Psychology Service and VA National Center for PTSD, VA Boston Healthcare System and Department of Psychiatry, Boston University School of Medicine, Boston, Massachusetts.)
[8]Psychiatry Department, University of Cincinnati, Cincinnati, Ohio
[9]Psychiatry Service, Cincinnati VA Medical Center (Author is currently at Psychiatry Service, San Diego VA Healthcare System, San Diego, California, and the University of California, San Diego, California.)
[10]Environmental Epidemiology Service, Department of Veterans Affairs, Washington D.C.
[11]Medical and Research Services, St. Louis Department of Veterans Affairs Medical Center, St. Louis, Missouri.
[12]Department of Internal Medicine, Division of General Medical Sciences, Washington University School of Medicine, St. Louis, Missouri.
[13]Health Services Research and Development, Veterans Administration, Washington, D.C.
[14]Veterans Health Administration, U.S. Department of Veterans Affairs, Washington, D.C.
[15]Uniformed Services University of the Health Sciences, Washington, D.C.

(RECEIVED April 27, 2006; FINAL REVISION May 20, 2009; ACCEPTED May 26, 2009)

Abstract

Many U.S. Gulf War-era veterans complained of poor cognition following the war. This study assessed neuropsychological functioning in veterans 10 years after the war through objective tests. 2189 Gulf War-era veterans (1061 deployed, 1128 non-deployed) were examined at 1 of 16 U.S. Veterans Affairs medical centers. Outcomes included neuropsychological domains derived from factor analysis and individual test scores. Deployed veterans performed significantly worse than non-deployed veterans on 2 of 8 factors (motor speed & sustained attention, analysis not corrected for multiple comparisons) and on 4 of 27 individual test variables (Trails A & B, California Verbal Learning Test – List B, and Continuous Performance Test sensitivity, with only Trails B surviving Bonferroni correction). Within deployed veterans, Khamisiyah exposure was negatively correlated with motor speed after controlling for emotional distress. Depressive symptoms and self-reported exposure to toxicants were independently and significantly associated with worse sustained attention. Other factors were also associated with self-reported exposures. The findings were not a result of differential effort across groups. Gulf War deployment is associated with subtle declines of motor speed and sustained attention, despite overall intact neuropsychological functioning. Evidence suggests that toxicant exposures influence both these functions, and depressive symptoms also influence attention. (*JINS*, 2009, *15*, 717–729.)

Keywords: Cognition, Psychological tests, Neurotoxicity syndromes, Environmental medicine, Mental disorders, Psychological stress

INTRODUCTION

Approximately 700,000 U.S. military personnel deployed to Southwest Asia during Operation Desert Shield/Desert

Correspondence and reprint requests: Rosemary Toomey, Ph.D., Psychology Department, Boston University, 648 Beacon Street, 6th Floor, Boston, MA. 02215. E-mail: rosemary_toomey@hms.harvard.edu

Storm. Unique aspects of the Gulf War (GW) included deployment of a relatively large proportion of Reservists and National Guard units, exposures to potentially harmful natural and man-made environmental toxins (Presidential Advisory Committee on Gulf War Veterans' Illnesses, Final Report, 1996) and relatively few casualties. Large-scale epidemiological studies indicated that deployed veterans (DV) were more likely to report concentration and memory impairments than non-deployed veterans (NDV) (Fukuda et al., 1998; Iowa Persian Gulf Study Group, 1997; Ishoy et al., 1999; Kang et al., 2000; Unwin et al., 1999). Findings of smaller studies using objective neuropsychological tests have been less conclusive. Several studies reported a negligible relationship between subjective complaints of poor cognition and objectively measured cognitive performance (Binder et al., 1999; Lindem et al., 2003; Storzbach et al., 2000). Research suggesting that GW deployment is associated with neuropsychological deficits must be interpreted with caution, because most studies are based exclusively on self-reports of poor cognition (Fukuda et al., 1998; Iowa Persian Gulf Study Group, 1997; Ishoy et al., 1999; Kang et al., 2000; Unwin et al., 1999), or are not population based (Axelrod & Milner, 1997; Goldstein et al., 1996; Hom et al., 1997; Lindem et al., 2003; McDiarmid et al., 2000; Sillanpaa et al., 1997; White et al., 2001), or have small sample sizes (Anger et al., 1999; Binder et al., 1999; Storzbach et al., 2000).

Evidence suggests that subgroups of DV who may be at heightened risk for neuropsychological compromise show impairment on objective neuropsychological tests. For example, ill DV performed more poorly on tasks of memory, attention, and response speed (Anger et al., 1999). Haley et al., (1997a) identified six syndromes based on self-reported symptoms via factor analysis. Cases characterized by the syndrome of "impaired cognition" were impaired on brain stem auditory evoked potentials (Haley et al., 1997b), and risk of this syndrome was increased in veterans wearing flea collars (Haley & Kurt, 1997c). Hom et al. (1997) found that veterans with the highest factor scores from six identified syndromes, when compared to healthy veterans, demonstrated deficits in intelligence, abstract thinking, mental flexibility, grip strength, and tactile and visual perception. Storzbach et al. (2000) found that veterans with persistent unexplained symptoms exhibited deficits on some tests of memory, attention, and response speed when compared to healthy veterans. McDiarmid et al. (2000) found that seven years after exposure to depleted uranium, veterans' urine uranium levels were positively associated with an impairment index derived from tests measuring processing efficiency, but not with an impairment index derived from a traditional neuropsychological test battery. White et al. (2001) found that DV reporting greater exposure to pesticides performed more poorly on attention and memory tasks compared to unexposed veterans. In contrast, several studies found that neuropsychological deficits among DV were more closely related to emotional functioning than to other war-related variables (Axelrod & Milner, 1997; Binder et al., 1999; David et al., 2002; Goldstein et al., 1996; Sillanpaa et al., 1997).

The "National Health Survey of Gulf War Era Veterans and Their Families" was designed to collect longitudinal epidemiological data to assess the prevalence of medical and psychological conditions in GW veterans. Beginning in 1995, mail and telephone surveys were conducted on 15,000 DV and 15,000 NDV. DV reported more psychological symptoms, more functional impairment, and poorer health related quality of life compared to NDV (Kang et al., 2000). The current study evaluated a subset of these veterans using in-person psychological and neuropsychological exams conducted between 1998 and 2001. In this sample, we previously reported that GW-era onset mental disorders were more prevalent in DV, compared to NDV, with depression and anxiety disorders remaining higher at the time of the exam, approximately 10 years after the war (Toomey et al., 2007). At the time of the exam, DV were more likely to have 4 out of 12 medical conditions (fibromyalgia, chronic fatigue syndrome, dermatologic conditions, and dyspepsia) (Eisen et al., 2005), but had no increased prevalence of distal symmetric polyneuropathy, as determined by electrophysiology and neurologic examinations (Davis et al., 2004).

METHODS

Study Population and Recruitment

Recruitment for the survey phase of the study, performed in 1995, is described elsewhere in detail (Eisen et al., 2005). Briefly, the Department of Defense's Defense Manpower Data Center (Monterey, CA) identified the entire cohort of 693,826 DV and approximately half (800,680) of NDV who were in military service between September, 1990 and May, 1991. A total of 15,000 DV and 15,000 NDV were solicited to participate in the study. To ensure that females, Reserve, and National Guard were adequately represented, a stratified random sampling methodology was applied so that one-fifth of each sample were women (3000), one-third Reservists (5000), and approximately one-quarter members of the National Guard (4000).

For our study's examination phase, a list of potential participants was created by randomly selecting from the 11,441 DV and 9,476 NDV that participated in the 1995 study, stratified by deployment status and region of last-known residence at the time of the original survey, based on home telephone area code. Potential participants were assigned to the participating VA medical center closest to their home. Participating medical centers were located in Albuquerque, Baltimore, Birmingham, Boston, Cincinnati, Hines (Chicago), Houston, Miami, Minneapolis, New Orleans, New York, Portland (Oregon), Richmond, Salt Lake City, San Diego, and St. Louis. Recruitment packages were mailed to the veterans and included an introductory letter, a detailed explanation of the purpose and nature of the study, a letter of intent form, and a preaddressed stamped return envelope. Because of lower participation rates among NDV, we solicited an additional 799 NDV to obtain groups of equal size and achieve the desired sample size of 1,000 per group.

Signed letters of intent were returned to the Hines VA Cooperative Studies Program Coordinating Center, which forwarded them to the participating VA Medical Center to which the veteran was assigned. Site personnel contacted these veterans and scheduled examinations. Travel, hotel, per diem costs, and a $200 honorarium were provided by the research project. The protocol and consent form were approved by the Hines Cooperative Studies Program Human Rights Committee and the individual site and Brockton VAMC Institutional Review Boards. The research was completed in accordance with the Helsinki Declaration. Participants gave signed informed consent before the start of the examination.

Measures

Neuropsychological functioning

The neuropsychological test battery was based on prior GW studies (White et al., 2001). A similar battery, used in a study examining two cohorts of GW veterans (Fort Devens, MA and New Orleans, LA) and a Germany deployed cohort, was designed to assess abilities across general intelligence, attention/executive function, motor ability, visual-spatial processing, and verbal and visual memory. The only domain without suggestion of differences between DV and NDV cohorts was visual-spatial functioning (White et al., 2001); therefore, it was not included in the present test battery. In addition, White et al. used a measure of general intellectual abilities that did not generate group differences. We estimated IQ in our battery with a reading test commonly used as an estimate of premorbid intellectual abilities (Griffin et al., 2002; Johnstone et al., 1996).

Factor analysis with orthogonal (varimax) rotation was employed to reduce data and derive variables reflecting neuropsychological domains of functioning. This analysis was performed by investigators blind to veteran deployment status. Selection of the 27 variables entered in the factor analysis was based on three criteria: (1) clinical significance, (2) elimination of test variables with non-normal frequency distributions (e.g., an error score on which most subjects received a 0 or 1), and (3) maximization of internal consistency of factors. Table 1 lists the variables selected by balancing these criteria, as well as tests used to assess insufficient effort and premorbid intellectual functioning.

Table 1. Neuropsychological tests and variables

Neuropsychological Tests	Variables
Digit Span subtest (Wechsler Adult Intelligence Scale–III) (Wechsler, 1997)	Raw scores for Digits forwards and Digits backwards
Trail Making Test (Army Individual Test Battery, 1944; Partington & Leiter, 1949)	Trails A time, Trails B time – Trails A time
Paced Auditory Serial Arithmetic Test (Gronwall, 1977)	Trial 1 total correct
Continuous Performance Test (Letz, 1991)	Mean reaction time on trials 13–60
	Sensitivity *
Wisconsin Card Sorting Test (Heaton, 1981)	Number of categories
	Number of perseveration errors
California Verbal Learning Test (Delis, 1987)	List A trials 1–5: Total correct and semantic organization scores
	Total correct for list B
	Total correct for short-delay free and cued recall
	Total correct, list A trial 5 compared to short-delay free recall
	Total correct for long-delay free and cued recall
	Number of recognition hits
Rey-Osterrieth Complex Figure Test (Osterrieth, 1944; Rey, 1941; Waber & Holmes, 1986)	Organization scores for copy, immediate and delayed recall conditions
	Accuracy scores for recall conditions
Finger Tapping Test (Halstead, 1947)	Mean number of taps for each hand
Purdue Pegboard Test (Tiffin, 1968)	Number of pegs for each hand
Test to Assess Malingering	**Exclusion criteria**
Test of Memory and Malingering (Tombaugh, 1996)	If Trial 1 < 25 or Trial 2 <45
Test to Estimate Premorbid IQ	**Variable**
Reading Subtest (Wide Range Achievement Test–III) (Jastak & Wilkinson, 1997)	Standard score

Note. *Because hit rates (HR) and false alarm rates (FA) are biased by the perceived payoffs for correct identification and perceived penalties for incorrect identification, a nonbiased sensitivity score was determined arithmetically from HR and FA using the following equation (Davies & Parasuraman, 1982): $A' = \frac{1}{2} + \frac{(HR - FA)(1 + HR - FA)}{4 \times HR(1 - FA)}$

The sign of some test variables was reversed so that higher scores indicated better performance on all variables. Examiners infrequently rated variables invalid when factors unrelated to the test clearly interfered with the subject's performance on the test, such that the scores were not considered an accurate reflection of the subject's ability in that domain (e.g., poor motor performance due to a hand injury). Regression imputation was utilized to replace missing values on valid tests. Factors were maintained with eigenvalue magnitudes greater than one.

We compared factor scores to the combined group mean using impairment thresholds commonly used in clinical neuropsychological assessment. One threshold was 1 standard deviation (*SD*) below the mean, reflecting possible impairment. The second threshold was 2 *SD* below the mean, reflecting clear evidence of definite impairment (Lezak et al., 2004). We examined both thresholds to balance the risk of false positives with the cutoff of 1 *SD* and false negatives with the cutoff of 2 *SD*. We also compared groups on factor score means and individual test score means.

Mental health assessment

The PTSD Checklist (PCL; Blanchard et al., 1996) was used to assess symptoms of post-traumatic stress disorder (PTSD) in the past month using 17 items, each rated on a scale of 1–5. The Beck Depression Inventory-II (BDI-II; Beck et al., 1996) was used to assess depressive symptoms in the past two weeks using 21 items, each rated on a scale of 0–3. The Beck Anxiety Inventory (BAI; Beck & Steer, 1993) was used to assess anxiety symptoms in the past week using 21 items, each rated on a scale of 0–3. Higher scores on each indicate greater symptoms and/or symptom severity.

Post-hoc analyses

We followed up the main group analyses with analysis of subgroups defined as follows:

Chronic Multisymptom Illness complex (CMI). CMI was defined, according to the Centers for Disease Control definition, as the presence, for six months or longer, of one or more symptoms from at least two of the following symptom clusters: general fatigue (Cluster-A), mood and cognitive abnormalities (Cluster-B), and musculoskeletal pain (Cluster-C) (Fukuda et al., 1998). Further details about the identification of CMI in our sample are reported by Blanchard et al. (2006). CMI prevalence was 28.9% among DV and 15.8% among NDV. Diagnoses of pre-war non-PTSD anxiety disorders and depression were related to the presence of CMI in both groups.

Khamisiyah exposure. In March 1991, U.S. demolition experts destroyed a munitions storage site at Khamisiyah, Iraq that was subsequently demonstrated to contain nerve agents, sarin and cyclosarin. The risk of individual veteran exposure was estimated by the Department of Defense by overlaying troop location data with meteorological and dispersion modeling (Winkenwerder, 2002).

Self-report of toxic exposure. In the 1991 survey, participants were asked "While in the Persian Gulf, do you believe you were exposed to or did you experience any of the following?" followed by several types of exposures: [Chemical Agent Resistant Compound (CARC) paint or other paint and/or solvent and/or petrochemical substances (2 questions); personal pesticides, including creams, sprays, or flea collars; nerve gas; immune globulin (IG); vaccines (5 questions); SCUD missile explosions; ate food contaminated with smoke, oil, or other chemicals, or bathed in or drank water contaminated with smoke, oil, or other chemicals (2 questions); pyridostigmine bromide pills (PB); ciprofloxacin; microwaves; burning trash/feces; smoke from oil well fires; depleted uranium; diesel, kerosene, or skin exposure to diesel (2 questions)].

Training and quality control

Dr. Toomey trained staff in neuropsychological testing and managed quality through weekly calls and periodic reviews.

Statistical analyses

Sample size requirements were estimated *a priori*. The target sample size of 1000 in each group provided 80% power to detect differences of 4.4% for neuropsychological impairment (assumed GW DV prevalence = 11.4%, NDV = 7.0%). Population prevalence estimates for both the unadjusted and adjusted analyses were obtained using SUDAAN (software developed for the analysis of complex survey data), release 8.0 (Research Triangle Institute, Research Triangle Park, NC). The sampling design was a stratified random sample with unequal probabilities of selection within combinations of the strata: deployment status, gender, and duty type. For continuous outcomes, *t* tests and linear regression models compared mean responses between groups. Logistic regression models were developed for dichotomous and ordinal polytomous outcomes. The demographic and military covariates considered in the multiple regression models were age, gender, race (white *vs.* other), years of education (less than 12 years *vs.* 12 or more), duty type (active *vs.* Reserve/National Guard), service branch (Army/Marine *vs.* Navy/Air Force), and rank (enlisted *vs.* officer). Candidate covariates were deleted for particular models when they caused computational problems preventing model calculation. The primary reason for this was low prevalence rates for certain outcome measures. Odds ratios, 95% confidence intervals, and *p* values are reported for dichotomous outcomes. Comparisons of categorical data and continuous data with adjustment for covariates are based on the Wald *F* statistic. For continuous data without adjustments, *p* values are based on the 2-sample *t* test. All statistical tests were two-sided and $p \leq .05$ was the criterion for statistical significance. This criterion was chosen for ease of comparison with other studies. Analysis of group differences on individual test variables is reported

with and without Bonferroni correction, again for ease of comparison to other studies.

RESULTS

Participants

Table 2 lists the sociodemographic and military characteristics of 2189 Veterans who participated in the study. At the time of their research examination, DV significantly differed from NDV participants in that they were slightly younger, less likely to be Caucasian, less likely to have education past the high school level, less likely to be officers, less likely to be married, and reported a lower annual family income than NDV. There was no difference in the percentage of male participants and the percentage on active military duty. Although there were some group differences, these differences existed between the actual DV and NDV cohorts, and thus, are not unique to this sample (Kang & Bullman, 2001).

Participation Rates

Of the 1996 DV who were solicited to participate, 53% (1061) were examined; of 2883 NDV who were solicited to participate, 39% (1128) were examined. Despite intensive efforts, 12.8% of DV and 15.2% of NDV were not located. Of those who were located, 34.1% of DV and 45.6% of NDV either never returned their participation letter, or an examination could not be scheduled.

Participation Bias

Historical military service data, obtained in 1991 on all solicited veterans from the U.S. Department of Defense's Manpower Data Center, were used to evaluate participation bias on sociodemographic variables. We compared participants and nonparticipants in each group (DV and NDV) and used the Breslow-Day homogeneity of odds ratios test to determine whether the odds ratios for DV and NDV were equal.

Participation bias for demographic characteristics is reported in detail elsewhere (Eisen et al., 2005). Briefly, we found that participants were nearly two years older than nonparticipants, and Caucasians, women, Reservists, and National Guard members were significantly more likely to participate. Additionally, officers and Army personnel were more likely to participate than non-officer and non-Army personnel, although these differences were not uniformly significant. We calculated participation bias for self-reports of neuropsychological functioning from the assessment of

Table 2. Sociodemographic and military service characteristics of deployed and non-deployed participants at the research examination

Characteristic		Deployed (n = 1061)	Non-deployed (n = 1128)	p value Deployed *vs.* Non-deployed
Mean Age (*SD*)		38.9 (8.8)	40.7 (9.6)	0.001
Sex, %	Male	78.0	78.0	0.99
	Female	22.0	22.0	
Race, %	Caucasian	76.4	80.0	0.03
	African American	19.9	15.7	
	Other	3.7	4.3	
Highest Education, %	< High School graduate	1.8	2.0	0.001
	High School graduate	65.7	56.0	
	College graduate	19.8	22.1	
	Postgraduate	12.7	19.9	
Active Military Duty		7.8	8.5	> 0.2
Rank, %	Enlisted	85.7	80.4	0.001
	Officer	14.3	19.6	
Branch of Service	Army	64.6	62.9	0.22
	Navy	12.0	13.6	
	Air Force	11.9	13.7	
	Marines	11.6	9.8	
Unit, %	Reserves	36.3	36.9	0.78
	Active	35.2	35.9	
	National Guard	28.6	27.2	
Mean Income In $1,000's (*SD*)		46.8 (32.6)	52.0 (44.3)	0.003
Marital Status, %	Married	67.5	72.3	0.02
	Never Married	17.0	12.2	
	Divorced	12.5	12.5	
	Other	3.0	2.9	

Table 3. Rotated factor loadings for neuropsychological test variables

Neuropsychological Test Variables	Verbal Memory	Attention/ Working Memory	Visual Memory	Executive Functioning	Perceptual Motor Speed	Visual Organization	Motor Speed	Sustained Attention
CVLT Long Delay Cued Recall	**.91**	.09	.11	.08	.08	.04	–.001	.06
CVLT Long Delay Free Recall	**.90**	.09	.14	.07	.09	.05	–.02	.07
CVLT Short Delay Cued Recall	**.90**	.08	.10	.08	.10	.04	–.01	.06
CVLT Short Delay Free Recall	**.90**	.09	.13	.10	.09	.05	–.03	.08
CVLT Total Hits Trials 1–5	**.86**	.18	.01	.07	.11	.08	.01	–.04
CVLT Semantic Clustering Trials 1–5	**.77**	.11	–.04	.07	.05	.08	–.01	–.07
CVLT Recognition Hits	**.60**	.06	.08	.003	–.02	–.01	.03	.08
CVLT Short Free Recall vs. Trial 5	**.54**	–.05	.19	.09	.04	–.004	–.05	.17
CVLT Hits List B	**.46**	.28	–.03	.01	.11	.03	.05	–.12
Digit Span Forward	.01	**.80**	.03	–.01	.02	.05	.05	.004
Digit Span Backwards	.17	**.80**	.05	.04	.05	.06	.01	–.01
PASAT Trial 1 Total Correct	.23	**.61**	.06	.22	.12	.02	.08	.21
Trails B – Trails A Time	.23	**.45**	.06	.28	.08	.07	–.004	.13
Rey Immediate Recall Accuracy	.21	.08	**.90**	.06	.04	.18	.04	–.01
Rey Delayed Recall Accuracy	.23	.09	**.89**	.07	.05	.18	.05	.02
WCST Number of Categories	.13	.12	.05	**.91**	.08	.05	.04	.02
WCST Perseverative Responses	.14	.14	.06	**.90**	.07	.06	.06	–.01
Purdue Pegboard Dominant Hand	.13	.04	.01	.05	**.85**	.01	.02	.06
Purdue Pegboard Non-dominant Hand	.13	.09	.0002	.05	**.84**	.07	.06	.06
Trails A Time Score	.16	.19	.17	.13	**.43**	–.11	.19	.20
Rey Delayed Recall Organization	.07	.02	.23	.05	.01	**.77**	.03	–.03
Rey Immediate Recall Org.	.07	.04	.34	.05	.02	**.73**	.04	.01
Rey Copy Organization	.06	.10	–.12	.03	–.01	**.72**	.03	.03
FTT Mean Non-dominant Hand	–0.02	.04	.01	.03	.09	.03	**.91**	.07
FTT Mean Dominant Hand	–0.04	.07	.06	.06	.07	–.002	**.90**	.05
CPT Sensitivity	.02	.11	.08	.02	.02	–.01	.03	**.80**
CPT Mean Reaction Time	.14	.04	–.10	–.01	.18	.04	.10	**.72**
Cronbach's Alpha	.92	.70	.95	.87	.65	.65	.82	.46
Eigenvalue	5.81	2.15	1.96	1.85	1.77	1.77	1.72	1.35
Variance Accounted for (Total = 68%)	21.5%	8.0%	7.3%	6.9%	6.6%	6.6%	6.4%	5.0%

Note. CVLT = California Verbal Learning Test, PASAT = Paced Auditory Serial Arithmetic Test, CPT = Continuous Performance Test, WCST = Wisconsin Card Sorting Test, FTT = Finger Tapping Test, Cronbach's alpha was calculated for variables with loading ≥ .40 on each factor (factor loading scores in bold).

Table 4. Prevalence of cognitive impairment at two levels compared to combined group mean

Neuropsychological Impairment	Possible Impairment (–1 SD)				Definite Impairment (–2 SD)			
	Deployed (n = 1043)	Non-deployed (n = 1114)	p value[1]	Odds Ratio (CI)	Deployed (n = 1043)	Non-deployed (n = 1114)	p value[1]	Odds Ratio (CI)
Verbal memory	15.1%	16.3%	0.96	0.99 (0.69, 1.41)	4.1%	3.0%	0.17	1.72 (0.80, 3.73)
Attention/working memory	18.3%	13.2%	0.35	1.19 (0.83, 1.71)	1.8%	0.7%	0.41	1.57 (0.54, 4.63)
Visual memory	13.0%	15.6%	0.71	0.94 (0.66, 1.33)	2.2%	4.3%	0.20	0.63 (0.31, 1.27)
Executive Functioning	12.7%	14.1%	0.58	0.90 (0.63, 1.29)	4.3%	5.9%	0.18	0.70 (0.41, 1.18)
Perceptual Motor Speed	12.3%	13.9%	0.99	1.00 (0.70, 1.43)	1.8%	2.0%	0.42	1.42 (0.60, 3.37)
Visual organization	15.3%	16.4%	0.80	0.96 (0.68, 1.34)	3.0%	5.0%	0.046	0.52 (0.28, 0.99)
Motor speed	12.8%	13.5%	0.55	1.11 (0.79, 1.57)	2.6%	1.4%	0.02	2.35 (1.16, 4.75)
Sustained Attention	10.9%	10.2%	0.62	1.11 (0.74, 1.68)	3.7%	1.7%	0.02	2.64 (1.17, 5.96)

Note. [1]Adjusted p -values are listed for all factors. CI = confidence interval. Analyses were adjusted for the Wide Range Achievement Test-3 (Reading subtest, Standard Score), age, gender, race (white *vs.* other), years of education (less than 12 *vs.* 12 or more), duty type (active *vs.* reserve/guard), service branch (army/marine *vs.* navy/air force), and rank (enlisted *vs.* officer).

depressive symptoms ["Have you experienced (in the past year): difficulty concentrating or reasoning, memory loss"] in our 1995 survey. Participants and nonparticipants did not differ significantly within DV and NDV in their reports of these symptoms. The only characteristic that yielded a statistically significant difference in the odds ratios comparing participants to nonparticipants within NDV and DV was 1995 active duty status (Eisen et al., 2005). Participants were less likely to be on active duty. Overall, considering all variables examined, the degree of participation bias was independent of deployment status.

Neuropsychological Functioning

The groups differed significantly on the WRAT-III reading subtest. NDV displayed greater oral reading proficiency than DV (DV mean = 98.02, *SE*= 0.55; NDV mean = 100.39, *SE* = 0.43; weighted and adjusted p = .004). Because this test can be used to estimate level of premorbid intellectual functioning and it differed between DV and NDV, WRAT-III Reading was included as a covariate in neuropsychological analyses. Sixteen veterans (10 DV, 6 NDV) were excluded because they met criteria for insufficient effort (n = 8) on the Test of Memory and Malingering (TOMM) or had one or both trials missing (n = 8). Thirteen subjects were excluded for one or more invalid test scores (5 DV and 13 NDV). The tests that were invalid on these 13 subjects were as follows: grooved pegboard (3 subjects), CVLT (3), PASAT (3), Trails (1), Rey (1), and multiple test exclusions for 2 subjects. Some examples of the reasons for the invalid ratings include hand injuries or severe carpal tunnel syndrome interfering with pegboard performance, inability to understand directions, complete the practice items sufficiently, or clearly giving up effort part way through on the PASAT, and in some cases, examiner error (e.g., not stopping and redirecting a veteran making errors on the Trails test). One subject was excluded for an invalid reading score (DV). Table 3 presents the factor analysis results. Eight factors were generated, accounting for 68% of the variance: verbal memory, attention/working memory, visual memory, executive functioning, perceptual motor speed, visual organization, motor speed, and sustained attention.

Table 4 shows the group comparisons on factor impairment. At the cutoff of 1 *SD* below the mean (possible impairment), the percentage of DV reaching this threshold on any one factor did not differ significantly from the percentage of NDV. At the cutoff of 2 *SD* below the mean (definite impairment) (Lezak et al., 2004), DV performed significantly worse than NDV on factor 7 (motor speed) (DV = 2.6%, NDV = 1.4%, p = 0.02) and factor 8 (sustained attention) (DV = 3.7%, NDV = 1.7%, p = 0.02). In contrast, NDV performed worse compared to DV on factor 6 (visual organization) (DV = 3.0%, NDV = 5.0%, p = 0.046).

Given that there were 7.2% more NDVs with postgraduate education compared to DVs, the above analyses were repeated to see if the results changed when excluding these subjects. This sample size included 1801 veterans (354 postgraduate veterans and 2 veterans with missing education were deleted). None of the previously nonsignificant p values changed to significant. DV still performed significantly worse than NDV on factor 7 (motor speed) (DV = 2.7%, NDV = 1.6%, p = 0.03) and factor 8 (sustained attention) (DV = 3.7%, NDV = 1.7%, p = 0.04). The significant p value for factor 6 (visual organization) shifted to a trend (DV = 2.9%, NDV = 4.4%, p = 0.055).

Group comparison of mean factor scores revealed no significant differences. Group comparisons of the mean scores of the 27 individual variables included in the factor analysis yielded four variables with significant differences between groups (Table 5): Trails A time (DV mean = 30.0, *SE* = 0.49; NDV mean = 29.1, *SE* = 0.43; p = .01); Trails B-A time (DV mean = 36.6, *SE* = 1.26; NDV mean = 31.9, *SE* = 0.84; p = .002); CVLT List B number correct (DV mean = 6.5, *SE* = 0.09; NDV mean = 6.8, *SE* = 0.07; p = .03); and CPT sensitivity (DV mean = 0.98, *SE* = .002; NDV mean = 0.99, *SE* = .001; p = .01). For all four variables, DV performed less

Table 5. Group means and standard errors (SE) on individual neuropsychological test variables

Neuropsychological Test Variables	Deployed (n = 1061) Mean (SE)	Non-deployed (n = 1128) Mean (SE)	p value Deployed *vs.* Non-deployed	Adjusted p value*
CVLT Long Delay Cued Recall	12.3 (0.11)	12.4 (0.11)	0.59	0.53
CVLT Long Delay Free Recall	11.6 (0.13)	11.7 (0.12)	0.41	0.18
CVLT Short Delay Cued Recall	12.1 (0.11)	12.3 (0.10)	0.22	0.15
CVLT Short Delay Free Recall	11.3 (0.13)	11.4 (0.11)	0.64	0.28
CVLT Total Hits Trials 1–5	52.8 (0.42)	53.6 (0.36)	0.16	0.13
CVLT Semantic Clustering Trials 1–5	18.4 (0.49)	19.0 (0.43)	0.33	0.60
CVLT Recognition Hits	14.8 (0.06)	14.8 (0.06)	0.77	0.65
CVLT Short Free Recall *vs.* Trial 5	–10.9 (0.71)	–10.6 (0.58)	0.78	0.17
CVLT Hits List B	6.5 (0.09)	6.8 (0.07)	0.01	0.03
Digit Span Forward	10.5 (0.11)	10.8 (0.09)	0.06	0.47
Digit Span Backwards	6.9 (0.11)	7.2 (0.09)	0.046	0.23
PASAT Trial 1 Total Correct	39.7 (0.56)	40.8 (0.46)	0.13	0.36
Trails B – Trails A Time	–36.6 (1.3)	–31.9 (0.8)	0.002	0.002
Rey Immediate Recall Accuracy	51.4 (0.40)	50.6 (0.36)	0.13	0.60
Rey Delayed Recall Accuracy	51.7 (0.36)	51.5 (0.32)	0.65	0.58
WCST Number of Categories	5.3 (0.07)	5.3 (0.06)	0.98	0.80
WCST Perseverative Responses	–15.2 (0.62)	–14.9 (0.52)	0.72	0.69
Purdue Pegboard Dominant Hand	14.2 (0.09)	14.2 (0.08)	0.95	0.42
Purdue Pegboard Non-dominant Hand	13.6 (0.08)	13.5 (0.07)	0.41	0.26
Trails A Time Score	–30 (0.49)	–29.1 (0.43)	0.16	.01
Rey Delayed Recall Organization	9.4 (0.16)	9.1 (0.14)	0.28	0.70
Rey Immediate Recall Organization	9.4 (0.16)	9.2 (0.14)	0.33	0.71
Rey Copy Organization	10.0 (0.13)	10.0 (0.12)	0.98	0.88
FTT Mean Non-dominant Hand	46.0 (0.32)	46.3 (0.28)	0.46	0.16
FTT Mean Dominant Hand	50.8 (0.31)	50.9 (0.29)	0.72	0.22
CPT Sensitivity	0.98 (0.002)	0.99 (0.001)	0.007	0.01
CPT Mean Reaction Time	–399.4 (1.79)	–398.8 (1.54)	0.81	0.42

Note. CVLT = California Verbal Learning Test, PASAT = Paced Auditory Serial Arithmetic Test, CPT = Continuous Performance Test, WCST = Wisconsin Card Sorting Test, FTT = Finger Tapping Test,
*Analyses were adjusted for the Wide Range Achievement Test-3 (Reading subtest, Standard Score), age, gender, race (white *vs.* other), years of education (less than 12 *vs.* 12 or more), duty type (active *vs.* reserve/guard), service branch (army/marine *vs.* navy/air force), and rank (enlisted *vs.* officer).

proficiently than NDV. After using a Bonferroni corrected p value to adjust for multiple tests ($p < .002$), only Trails B-A remained significant.

To address whether significant neuropsychological declines remained in DV compared to NDV when emotional factors were controlled, we conducted three regression analyses to model the mean neuropsychological scores for factors 7 and 8 and for Trails B-A, using the explanatory variables of deployment and the continuous scores of the BDI, BAI, and PCL. Initially, a model was run with all explanatory variables entered simultaneously to evaluate the independent contributions of the explanatory variables on the outcomes. The WRAT-3 was also included as a covariate in all models. Because there were high correlations among BDI, BAI, and PCL ($p < .0001$), none were significant when all three were entered in the three models. Thus, we tried all combinations and tested for interactions. For factor 7, none of the variables were significant in any combination. For

factor 8, deployment was not significant in any combination, but when each of the PCL and BDI were paired alone with deployment, these symptom scores were significant. For Trails B-A, deployment and WRAT-3 were significant in all models, but symptoms were not significant in any model.

Post-hoc Analyses

To investigate relationships between specific deployment-related factors and neuropsychological functioning, we conducted a series of logistic regression analyses within the deployed group only. Mean scores on all eight neuropsychological factors were the dependent variables, and the independent variables included PCL, BDI, BAI, CMI status, Khamisiyah exposure status, and self-reported exposure to toxic substances. In addition, age, sex, race, education, and WRAT-3 reading variables were used as covariates.

In the initial analyses, nonsymptom independent variables were examined individually to determine their significance. Significant explanatory variables were as follows: Verbal Memory (Khamisiyah, vaccines), Visual Memory (CARC/other paint, Immune Globulin), Perceptual Motor Speed (Scud missile explosions, vaccines), Motor Speed (Khamisiyah), and Sustained Attention (CMI, Pesticides, Nerve Gas, Contaminated water and food, Scud missile explosions, and PB pills). The following self-reported exposures were not significant for any factors: Ciprofloxacin, microwaves, burning trash/feces, smoke from oil well fires, depleted uranium, diesel, kerosene + skin exposure to diesel. For three factors (Attention/Working memory, Executive Functioning, Visual Organization), none of these explanatory variables were significant.

Follow-up analyses were conducted to determine if these predictors remained significant when including psychological symptoms in the model. For four of the five models, there was no change, while there was a change in the model for Sustained Attention. CMI, pesticides, scud missile explosions, and PB pills were no longer significant explanatory variables in the Sustained Attention model with the inclusion of psychological symptoms. Specifically, depression was a significant explanatory variable in all models for Sustained Attention. Only self-reported exposure to contaminated food or water remained significant with depression in the model.

In cases in which predictors remained significant after controlling for psychological symptoms and there was more than one significant predictor, we created models including the demographic variables and the significant toxicant exposure variables. For Verbal Memory, Khamisiyah exposure remained significant (along with age, gender, and WRAT reading), whereas vaccines were no longer significant (along with race and education). For Visual Memory, both CARC/paint and IG remained significant (along with age and gender), while race, education, and WRAT reading were not significant. For Psychomotor Speed, both Scud missiles and vaccines remained significant (along with age, gender, race, and education), whereas the only nonsignificant variable was WRAT reading. For Sustained Attention, significant variables in the model were contaminated water/food, education, and BDI. Nonsignificant variables were age, gender, race, and WRAT reading. In this model, contaminated water/food was somewhat correlated with nerve gas. When both variables were entered in a model together with depressive symptoms and education, contaminated water/food was no longer significant ($p = .29$), nerve gas was a trend ($p = .09$), education remained significant ($p = .02$), and BDI was the most significant ($p = .001$).

DISCUSSION

Gulf War deployed veterans have reported poor concentration and memory loss after their service (Fukuda et al., 1998; Iowa Persian Gulf Study Group, 1997; Ishoy et al., 1999; Kang et al., 2000; Vasterling et al., 2003; Unwin et al., 1999), but prior research indicated that self-report does not necessarily relate to performance deficits (Axelrod & Milner, 1997; Binder et al., 1999; Silanpaa et al., 1997; Vasterling et al., 2003). We used objective neuropsychological tests to compare DV and NDV and to examine subgroups within DV. Using a possible impairment threshold of –1 *SD*, the percentage meeting this threshold in both groups was similar to that expected in a normative sample, and in some cases lower than would be expected. Using a definite impairment threshold of –2 *SD*, the percentage meeting this threshold in both groups tended to be higher than expected in a normative group. Using this latter threshold, DV performed worse on factors of motor speed and sustained attention compared to NDV. These results are based on analysis with no statistical correction for multiple comparisons. Such a correction was made on the comparison between groups on individual test variables, and DV performed worse on an individual test variable measuring flexibility of attention compared to NDV.

DV performed less well than NDV on three other test variables that did not remain significant after Bonferroni correction. These variables measured visual attention, verbal memory, and sustained attention. Of these four variables, three have norms available that vary in the size and nature of the normative sample. The mean of Trails A (DV = 30.0, NDV = 29.1) and Trails B-A (DV = 36.6, NDV = 31.9) are within normal limits (25th–75th percentile) compared to group norms (Strauss et al., 2006). CVLT variables with available group norms also are within normal limits for both groups compared to male and female norms (Delis et al., 1987). Lezak et al. (2004) emphasize that there are times when another comparison is more appropriate than a comparison with a normative sample, particularly when the normative sample may differ from an individual or clinical sample. In contrast to available normative groups, our control group of NDV is much larger and more closely matches the DV in important demographic and military characteristics. Though mean scores for both groups are within the normal range based on group norms, there are differences between groups in the percentage of individuals meeting a predefined impairment

threshold for motor speed and sustained attention. Our definition of impairment is defined in terms of distance from the mean. Because the normal range spans half of people (25th–75th percentile), there is room for significant strengths and weaknesses within the normal range. On an individual level, a drop in functioning may be relevant even if performance remains in the average range.

Our findings of DV performing worse on attention tasks (Trails B-A and the sustained attention factor comprised of CPT variables) compared to NDV are consistent with those of White et al. (2001), who used a Continuous Performance Test, the Trails error score, and another measure of attention, although their differences did not survive Bonferroni correction. Our finding of DV performing worse on verbal memory compared to NDV is consistent with David et al. (2002). For David et al., this difference no longer remained when controlling for current depression; in our study, the difference did not survive the Bonferroni correction. Vasterling et al. (2006) also found decrements in verbal learning and visual-spatial memory in DV. Also consistent with other studies, we found differences in motor functioning. We used the Finger Tapping Test while David et al. (2002) and White et al. (2001) both used the Purdue Pegboard test. In contrast to other findings, we did not find differences on executive function (White et al., 2001, WCST). Axelrod and Miller (1997) compared DV performance to population means and found DV dysfunction on motor (Purdue Pegboard) and executive skills (the Stroop Test). Thus, we replicated small differences between groups on attention tasks and motor functioning, but not on executive function tasks.

Previous research demonstrating neuropsychological deficits among GW subgroups suggests these deficits are related to psychological disorders or distress (Axelrod and Milner, 1997; Binder et al., 1999; David et al., 2002; Goldstein et al., 1996; Sillanpaa et al., 1997; Vasterling et al., 2003). This is not surprising given that difficulty concentrating is part of the diagnostic criterion for both depression and PTSD (American Psychiatric Association, 2000). Deficits in sustained attention have been observed in major depression (Egeland et al., 2003; van der Meere et al., 2007) and deficits in attention and working memory are associated with PTSD (e.g., Jenkins et al., 2000; Vasterling et al., 1998). The effects of PTSD on cognition are reviewed by Vasterling and Brailey (2005). Vietnam veterans with PTSD were shown to be impaired specifically on Trails B compared to Vietnam veterans without PTSD (Beckham et al., 1998). Previous GW investigators also controlled for emotional disorders when examining cognitive deficits. For example, Lange et al. (2001) compared Gulf War veterans with and without fatiguing illness and found that ill veterans displayed problems with response speed and mental flexibility that remained after controlling for emotional disorders. Complicating interpretation of the relationship between emotional factors and neuropsychological functioning is the possibility that psychological symptoms themselves could be caused by toxic exposures. Exposure to solvents and fuel can be associated with reports of depression, and, in some cases, this distress is independent of any cognitive dysfunction. Similarly, chronic exposure to pesticides is associated with reports of symptoms of depression and anxiety (Lezak et al., 2004). Alternatively, psychological symptoms may be causally related to other effects of toxic exposures.

Subtle deficits in DV may be multiply determined by such factors as the psychological and physical environments, in combination with individual vulnerability factors and immunologic factors (Vasterling & Bremner, 2006). In addition to the effect of psychological distress, previous research has demonstrated that objective toxic exposure (McDiarmid et al., 2000), or self-report of GW chemical exposure (White et al., 2001) may explain impaired neuropsychological functioning in GW veterans. Indeed, we found that chemical exposure significantly predicted some areas of neuropsychological functioning in DV. Khamisiyah exposure alone predicted slower motor speed, even after controlling for emotional symptoms, which were not significant predictors of motor speed. Consistent with our data, Proctor et al., (2006) found significant dose-response (high, moderate, none) relationships between Khamisiyah exposure and psychomotor dexterity (Purdue Pegboard), but not finger tapping. They speculated that motor speed as assessed by the Finger Tapping Test may have represented a skill that soldiers in the high exposure group were more likely to have practiced in their line of work. In contrast to Proctor et al., our sample is larger and we examined Khamisiyah exposure as a dichotomous variable. Their research group also compared 26 Khamisiyah exposed veterans to 13 unexposed veterans using brain magnetic resonance imaging (MRIs) to evaluate central nervous system pathology related to sarin/cyclosarin exposure. Linear trend analyses showed a significant association between Khamisiyah exposure and reduced white matter and increased right and left lateral ventricle volumes (Heaton et al. 2007).

In our analyses, both current depressive symptoms and self-reported exposure to contaminated food and water explained unique variance in sustained attention. Garfield and Leu (2000) cite data from the Iraqi ministry of health reporting a rapid rise in childhood illnesses related to contaminated water from 1990 to 1994. Other factors associated with sustained attention (CMI, Pesticides, Nerve gas, Scud missile explosion, PB pills) no longer remained significant after controlling for emotional distress.

Given that certain relationships may be obscured in group comparisons of DV with NDV, we examined relationships between deployment attributes and the other cognitive factors within the deployed group. For three other factors, environmental toxicants explained functioning beyond emotional distress and demographic variables: Khamisiyah exposure was associated with poorer verbal memory, CARC/other paint predicted poorer visual memory, and vaccines and scud missile explosions both explained poorer perceptual motor speed. A partial explanation for the lack of group differences may be that vaccines and IG exposure were not

uncommon among NDV (vaccines in DV 37% and in NDV 25%, IG in DV 31% and in NDV 13%). Although verbal memory deficits may not characterize the DV group as a whole, they are relevant for the subset of DV with Khamisiyah exposure. Similarly, deficits in perceptual motor speed may be relevant for only the subset of DV exposed to scud missile explosions.

In November of 2008, the Research Advisory Committee on Gulf War Veterans' Illnesses published a comprehensive report on the health of Gulf War veterans. In a summarizing statement, the report states that evidence across numerous studies indicates that GW illness is not caused by stress alone, and that psychological factors are not a central cause of GW illness. This is consistent with the fact that our neuropsychological findings remained significant even after controlling for emotional distress. In most research designs comparing DV to NDV, however, cognitive effects were diminished or eliminated when controlling for emotional functioning. The Institute of Medicine suggested that several of these studies overcorrected for depression and multiple comparisons, resulting in the possible masking of more subtle deficits. Our findings in DV *versus* NDV are indeed subtle, with DV performance still within the average range compared to normative samples. The Advisory Committee's report discusses how studies evaluating symptomatic veterans consistently found differences in several neuropsychological domains that remained significant after emotional adjustment, although they were not large (examples include response speed, attention, executive function, memory, visuospatial, and psychomotor function). Studies revealed slowed response latencies across several cognitive domains. Of note, our areas of lower performance in our DV occurred on timed tasks. The Advisory Committee also reviewed animal studies and fewer human studies that examined the interaction of stress and chemical exposures. There is good evidence on the synergistic effects of stress and PB and hypotheses that conditions of stress may allow PB to cross the blood brain barrier directly or may indirectly allow PB to have greater central and autonomic nervous system effects. The chemical exposures with the strongest evidence to date of association with GW illness include PB and pesticides, although to date, there is little evidence on the synergistic effects of the multiple exposures GW veterans may have experienced in their service.

Limitations to the study include the low study participation rates, which may have biased results. Studies suggest that nonparticipants have higher rates of psychiatric disorders, and thus, epidemiologic studies may underestimate the prevalence of psychiatric conditions (Haapea, 2008). We were able to determine that participants and nonparticipants did not differ in their report of the depressive symptoms of having difficulty with concentrating, reasoning, or memory loss, but we were unable to fully assess the differences between participants and nonparticipants on psychiatric and neuropsychological functioning. Given that we inquired about combat and other stresses, examiners were not blinded to the deployment status of the veterans. Furthermore, the cross-sectional design of the study precludes concluding causality regarding the relationships between neuropsychological functioning, poor physical or mental health, and exposure to environmental toxicants. We do not have objective verification of the self-report of toxic exposure in order to determine their validity. The Khamisiyah exposure was estimated by matching meteorological data and estimates of atmospheric transport of detonated munitions to data on troop locations in these geographical areas. Thus, these are estimates rather than exact determinations of who was exposed and the levels of exposures.

In summary, we found that DV performed worse on sustained attention and motor functioning compared to NDV, despite mean functioning in both groups remaining within the normal range. Depressive symptoms negatively influenced sustained attention, but not motor functioning. Chronic multisymptom illness and self-reported exposure to toxicants also influenced attention, coupled with emotional distress. In contrast, estimated Khamisiyah exposure was the only significant predictor for motor slowing. Some self-reported exposures to toxicants were associated with aspects of neuropsychological functioning that did not differentiate DV and NDV. The study thus demonstrates the importance of comparing DV and NDV, as well as examining subgroups within DV, and points to the need for better documentation of exposures to toxic substances and other war-zone hazards for future veterans.

ACKNOWLEDGMENTS

This study was supported by the Cooperative Studies Program of the Department of Veterans Affairs, Office of Research and Development. Nine authors of the study were employed by the VA at the time of the study, while one author (MJL) was a paid consultant for the study. The information in this manuscript and the manuscript itself has never been published either electronically or in print.

We acknowledge the Participating Investigators, Nurse Coordinators, and Research Assistants at the 16 participating Veterans Affairs Medical Centers (Albuquerque, Baltimore, Birmingham, Boston, Cincinnati, Hines, Houston, Miami, Minneapolis, New Orleans, New York-Manhattan, Portland, Richmond, St. Louis, Salt Lake City, San Diego); the staff at the Hines Cooperative Studies Program Coordinating Center, Hines, Illinois and the staff at the Cooperative Studies Program, Department of Veterans Affairs Central Office, Washington, DC; the members of the executive committee and the Data and Safety Monitoring Board; and the study Project Manager (Tammy Nydam, MA, Hines VAMC) and Coordinators (Joyce Hunter, RN, MSN, St. Louis VAMC; Leila Jackson, Dr.PH., VA Central Office).

We acknowledge the staff at the Brockton Psychometrics Lab, Brockton, Massachusetts: Lisa James and Zachary Warren, Research Assistants; Frank Weathers, Ph.D., and Marie Walbridge, Ph.D., Consultants.

REFERENCES

American Psychiatric Association. (1994). *Diagnostic and Statistical Manual of Mental Disorders (DSM-IV).* Washington, DC: American Psychiatric Association.

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

Army Individual Test Battery. (1944). *Manual of directions and scoring*. Washington, DC: War Department, Adjutant General's Office.

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

Beck, A.T., & Steer, R.A. (1993). *Beck Anxiety Inventory Manual*. San Antonio, TX: The Psychological Corporation, Harcourt Brace & Company.

Beck, A.T., Steer, R.A., & Brown, G.K. (1996). *Beck Depression Inventory-II Manual* (2nd ed.). San Antonio, TX: The Psychological Corporation, Harcourt Brace & Company.

Beckham, J.C., Crawford, A.L., & Feldman, M.E. (1998). Trail Making Test performance in Vietnam combat veterans with and without posttraumatic stress disorder. *Journal of Traumatic Stress*, *11*(4), 811–819.

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

Blanchard, E.B., Jones-Alexander, J., Buckley, T.C., & Forneris, C.A. (1996). Psychometric properties of the PTSD Checklist (PCL). *Behavior Research and Therapy*, *34*(8), 669–673.

Blanchard, M.S., Eisen, S.A., Alpern, R., Karlinsky, J., Toomey, R., Reda, D.J., et al. (2006). Chronic multisymptom illness complex in Gulf War I veterans 10 years later. *American Journal of Epidemiology*, *163*(1), 66–75.

David, A.S., Farrin, L., Hull, L., Unwin, C., Wessely, S., & Wykes, T. (2002). Cognitive functioning and disturbances of mood in UK veterans of the Persian Gulf War: A comparative study. *Psychological Medicine*, *32*, 1357–1370.

Davis, L.E., Eisen, S.A., Murphy, F.M., Alpern, R., Parks, B.J., Blanchard, M., et al. (2004). Clinical and laboratory assessment of distal peripheral nerves in Gulf War veterans and spouses. *Neurology*, *63*, 1070–1077.

Davies, D.R., & Parasuraman, R. (1982). *The psychology of vigilance*. Orlando, FL: Academic Press.

Delis, D.C., Kramer, J.H., Kaplan, E., & Ober, BA. (1987). California Verbal Learning Test: Adult Version. San Antonio, TX: The Psychological Corporation.

Egeland, J., Rund, B.R., Sundet, K., Landrø, N.I., Asbjørnsen, A., Lund, A., Roness, A., Stordal, K.I., & Hugdahl, K. (2003). Attention profile in schizophrenia compared with depression: Differential effects of processing speed, selective attention and vigilance. *Acta Psychiatrica Scandinavica*, *108*(4), 276–284.

Eisen, S.A., Kang, H.K., Murphy, F.M., Blanchard, M.S., Reda, D.J., Henderson, W.G., et al. (2005). Gulf War veterans' health: Medical evaluation of a U.S. cohort. *Annals of Internal Medicine*, *142*(11), 881–890.

Fukuda, K., Nisenbaum, R., Stewart, G., Thompson, W.W., Robin, L., Washko, R.M., et al. (1998). Chronic multisystem illness affecting Air Force veterans of the Gulf War. *Journal of the American Medical Association*, *280*, 981–988.

Garfield, R., & Leu, C. (2000). A Multivariate method for estimating mortality rates among children under 5 years from health and social indicators in Iraq. *International Epidemiological Association*, *29*, 510–515.

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

Griffin, S.L., Rivera Mindt, M., Rankin, E.J., Ritchie, A.J., & Scott, J.G. (2002). Estimating premorbid intelligence: Comparison of traditional and contemporary methods across the intelligence continuum. *Archives of Clinical Neuropsychology*, *17*(5), 497–507.

Gronwall, D.M.A. (1977). Paced Auditory Serial-Addition Task: A measure of recovery from concussion. *Perceptual Motor Skills*, *44*, 367–373.

Haapea, M., Miettunen, J., Läärä, E., Joukamaa, M.I., Järvelin, M., Isohanni, M.K., Veijola, J.M. (2008). Non-participation in a field survey with respect to psychiatric disorders. *Scandinavian Journal of Public Health*, *36*, 728–736.

Haley, R.W., Hom, J., Roland, P.S., Bryan, W.W., Van Ness, P.C., Bonte, F.J., et al. (1997b). Evaluation of neurologic function in Gulf War veterans. A blinded case-control study. *Journal of the American Medical Association*, *277*(3), 223–230.

Haley, R.W., & Kurt, T.L. (1997c). Self-reported exposure to neurotoxic chemical combinations in the Gulf War. A cross-sectional epidemiologic study. *Journal of the American Medical Association*, *277*(3), 231–237.

Haley, R.W., Kurt, T.L., & Hom, J. (1997a). Is there a Gulf War Syndrome? Searching for syndromes by factor analysis of symptoms. *Journal of the American Medical Association*, *277*(3), 215–222.

Halstead, W.C. (1947). *Brain and intelligence*. Chicago: University of Chicago Press.

Heaton, R.K. (1981). *Wisconsin Card Sorting Test manual*. Odessa, FL: Psychological Assessment Resources.

Heaton, K.J., Palumbo, C.L., Proctor, S.P., Killiany, R.J., Yurgelun-Todd, D.A., & White, R.F. (2007). Quantitative magnetic resonance brain imaging in US army veterans of the 1991 Gulf War potentially exposed to sarin and cyclosarin. *Neurotoxicology*, *28*, 761–769.

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 illnesses and health status among Gulf War veterans: A population-based study. *Journal of the American Medical Association*, *277*, 238–245.

Ishoy, T., Suadicani, P., Guldager, B., Appleyard, M., Hein, H., & Gyntelberg, F. (1999). State of health after deployment in the Persian Gulf: The Danish Gulf War study. *Danish Medical Bulletin*, *46*, 416–419.

Jastak, S., & Wilkinson, G.S. (1997). *Wide Range Achievement Test–III: Administration manual*. Wilmington, DE: Jastak Associates.

Jenkins, M.A., Langlais, P.J., Delis, D., & Cohen, R.A. (2000). Attentional dysfunction associated with posttraumatic stress disorder among rape survivors. *The Clinical Neuropsychologist*, *14*, 7–12.

Johnstone, B., Callahan, C.D., Kapila, C.J., & Bouman, D.E. (1996). The comparability of the WRAT-R reading test and NAART as estimates of premorbid intelligence in neurologically impaired patients. *Archives of Clinical Neuropsychology*, *11*(6), 513–519.

Kang, H.K., & Bullman, T.A. (2001). Mortality among US veterans of the Persian Gulf War: 7 year follow-up. *American Journal of Epidemiology*, *154*, 399–405.

Kang, H.K., Mahan, C.M., Lee, K.Y., Magee, C.A., & Murphy, F.M. (2000). Illnesses among United States veterans of the Gulf War: A population-based survey of 30,000 veterans. *Journal of Occupational and Environmental Medicine*, *42*, 491–501.

Lange, G., Tiersky, L.A., DeLuca, J., Scharer, J.B., Policastro, T., Fiedler, N., et al. (2001). Cognitive functioning in Gulf War illness. *Journal of Clinical and Experimental Neuropsychology*, *23*(2), 240–249.

Letz, R. (1991). *NES2 user's manual (version 4.4)*. Winchester, MA: Neurobehavioral Systems, Inc.

Lezak, M.D., Howieson, D.B., & Loring, D.W. (2004). *Neuropsychological assessment* (4th ed.). New York: Oxford University Press.

Lindem, K., Proctor, S.P., Heeren, T., Krengel, M., Vasterling, J., Sutker, P.B., et al. (2003). Neuropsychological performance in Gulf War era veterans: Neuropsychological symptom reporting. *Journal of Psychopathology and Behavioral Assessment*, *25*(2), 121–127

McDiarmid, M.A., Keogh, J.P., Hooper, F.J., McPhaul, K., Squibbs, K., Kane, R., et al. (2000). Health effects of depleted uranium on exposed Gulf War veterans. *Environmental Research*, *82*, 168–180.

Osterrieth, P.A. (1944). Le test de copie d'une figure complex: Contribution a l'etude de la perception et de la memoire. *Archives of Psychology*, *30*, 286–356.

Partington, J.E., & Leiter, R.G. (1949). Partington's Pathway Test. *The Psychological Service Center Bulletin*, *1*, 9–20.

Presidential Advisory Committee on Gulf War Veterans' Illnesses. (1996). *Final Report*. Washington, DC: U.S. Government Printing Office; ISBN 0-16048942-3.

Proctor, S.P., Heaton, K.J., Heeren, T., & White, R.F. (2006). Effects of sarin and cyclosarin exposure during the 1991 Gulf War on neurobehavioral functioning in US army veterans. *Neurotoxicology*, *27*, 931–939.

Purdue Research Foundation. (1948). *Examiner's manual for the Purdue Pegboard*. Chicago: Science Research Associates.

Research Advisory Committee on Gulf War Veterans' Illnesses. (2008). *Gulf War illness and the health of Gulf War veterans: Scientific findings and recommendations*. Washington, DC: U.S. Government Printing Office.

Rey, A. (1941). L'examen psychologique dans les cas d'encephalopathie traumatique. *Archives of Psychology*, *28*, 286–340.

Sillanpaa, 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.

Storzbach, D., Campbell, K.A., Binder, L.M., McCauley, L., Anger, W.K., Rohlman, D.S., & Kovera, C.A. (2000). Psychological differences between veterans with and without Gulf War unexplained symptoms. *Psychosomatic Medicine*, *62*, 726–735.

Strauss, E., Sherman, E.S., & Spreen, O. (2006). *A compendium of neuropsychological tests.* New York: Oxford University Press.

SUDAAN, Release 8.0. (2001). *Research Triangle Park*, NC: Research Triangle Institute.

Tiffin, J. (1968). *Purdue Pegboard examiner's manual*. Rosemont, IL: London House.

Tombaugh, T.N. (1996). *Test of Memory and Malingering*. North Tonawanda, NY: Multi-Health Systems, Inc.

Toomey, R., Kang, H.K., Karlinsky, K., Baker, D.G., Vasterling, J.J., Alpern, R., Reda, D.J., Henderson, W.G., Murphy, F.M., Eisen, S.A. (2007). Mental health of U.S. Gulf War veterans: A psychological evaluation 10 years after the war. *British Journal of Psychiatry*, *190*, 385–393.

Unwin, C., Blatchley, N., Coker, W., Ferry, S., Hotopf, M., Hull, L., et al. (1999). Health of UK servicemen who served in the Persian Gulf War. *Lancet*, *353*, 169–178.

van der Meere, J., Börger, N., & van Os, T. (2007). Sustained attention in major unipolar depression. *Perceptual Motor Skills*, *104*(3 Pt. 2), 1350–1354.

Vasterling, J.J., & Brailey, K. (2005). Neuropsychological findings in adults with PTSD. In J. Vasterling & C.R. Brewin (Eds.), *Neuropsychology of PTSD: Biological, cognitive, and clinical perspectives*. New York: Guilford.

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

Vasterling, J.J., Brailey, K., Tomlin, H., Rice, J., & Sutker, P.B. (2003). Olfactory functioning in Gulf War-era veterans: Relationships to war-zone duty, self-reported hazards exposures, and psychological distress. *Journal of the International Neuropsychological Society*, *9*, 407–418.

Vasterling, J.J., & Bremner, J.D. (2006). The impact of the 1991 Gulf War on the mind and brain: Findings from neuropsychological and neuroimaging research. *Philosophical Transactions of the Royal Society and Biological Sciences*, *361*(1468), 593–604.

Vasterling, J.J., Proctor, S.P., Amoroso, P., Kane, R., Heeren, T., & White, R.F. (2006). Neuropsychological outcomes of Army personnel following deployment to the Iraq war. *Journal of the American Medical Association*, *296*, 519–529.

Waber, D.P., & Holmes, J.M. (1986). Assessing children's memory productions of the Rey-Osterrieth Complex Figure. *Journal of Clinical and Experimental Neuropsychology*, *8*, 565–580.

Wechsler, D. (1997). *Manual for the Wechsler Adult Intelligence Scale–III*. San Antonio, TX: The Psychological Corporation.

White, R.F., Proctor, S.P., Heeren, T., Wolfe, J., Krengel, M., Vasterling, J., et al. (2001). Neuropsychological function in Gulf War veterans: Relationships to self-reported toxicant exposures. *American Journal of Industrial Medicine*, *40*, 42–52.

Winkenwerder, W. (2002). *U.S. demolition operations at Khamisiyah. Final Report*. Washington, DC: Department of Defense.