Health Psychology
2002, Vol. 21, No. 4, 321–331

0278-6133/02/$5.00 DOI: 10.1037//0278-6133.21.4.321

Stress-Related Immune Changes in Middle-Aged and Older Women: 1-Year Consistency of Individual Differences

Mary H. Burleson
Arizona State University West

Kirsten M. Poehlmann
University of Houston

Louise C. Hawkley
University of Chicago

John M. Ernst
Illinois Wesleyan University

Gary G. Berntson, William B. Malarkey, Janice K. Kiecolt-Glaser, and Ronald Glaser
The Ohio State University

John T. Cacioppo
University of Chicago

This study reviews prior research and reports longer-term consistency of stress-related immune variables in middle-aged and older women who performed mental math and speech tasks 2 times 1 year apart. Leukocyte subsets, mitogen-induced lymphocyte proliferation, and natural killer cell activity were measured at baseline, after tasks, and after 30-min recovery. Epstein-Barr virus (EBV) antibody titers were assessed at baseline. Pearson coefficients and standardized maximum-likelihood estimates of year-to-year covariances for leukocyte subsets and EBV titers showed moderately high to high baseline and posttask consistency and lower recovery consistency; consistency for other functional immune assays and reactivity scores for all variables was moderate to low. Results support longitudinal study of psychosocial context effects on tonic immune function and posttask scores.

Key words: reproducibility, consistency, psychological stress, immune function, middle-aged women, older women

Changes in immune parameters resulting from both acute and chronic stress are well documented (e.g., Glaser & Kiecolt-Glaser, 1994b). Clinical implications for these alterations are not certain, but researchers have postulated that they may mediate the relations between stress and various kinds of illness (e.g., Baum & Posluszny, 1999; Glaser, Rabin, Chesney, Cohen, & Natelson, 1999; Maier & Watkins, 1998; McEwen, 1998). Like the reactivity hypothesis of cardiovascular risk, hypotheses proposing that stress-related immune changes may affect chronic disease often assume implicitly that the level and type of change are stable characteristics (i.e., traits) of the individuals being assessed. Because chronic diseases often take years to develop, it may take years of stress-induced alterations to affect substantially the severity or progression of these diseases. Nevertheless, only a few studies have tested the temporal consistency of individual differences in tonic immune function or in immune reactivity to stress in healthy individuals. Further, all of the human studies used relatively young participants and short retest intervals (see Table 1 for summary). Thus, little is known about longer term consistency of immune function in older populations. In this article, we report the 1-year temporal consistency of individual differences in stress-related immune parameters measured in a sample of middle-aged and older women, some of whom were experiencing the chronic stress of caring for a spouse with progressive dementia.

Mary H. Burleson, Department of Social and Behavioral Sciences, Arizona State University West; Kirsten M. Poehlmann, Department of Psychology, University of Houston; Louise C. Hawkley and John T. Cacioppo, Department of Psychology and Institute for Mind and Biology, University of Chicago; John M. Ernst, Department of Psychology, Illinois Wesleyan University; Gary G. Berntson, Department of Psychology, The Ohio State University; William B. Malarkey, Department of Medicine, Division of Endocrinology, Institute for Behavioral Medicine Research, Comprehensive Cancer Center, College of Medicine, The Ohio State University; Janice K. Kiecolt-Glaser, Department of Psychiatry, Institute for Behavioral Medicine Research, Comprehensive Cancer Center, College of Medicine, The Ohio State University; Ronald Glaser, Department of Molecular Virology, Immunology, and Medical Genetics, Institute for Behavioral Medicine Research, Comprehensive Cancer Center, College of Medicine, The Ohio State University.

This work was supported in part by National Institute of Mental Health Training Grant MH-18831 and Grants MH-42096 and MH-50538; National Institute on Aging Program Project Grant AG-11585; National Institute of Health Grant M01 RR00034 to the General Clinical Research Center; and The Ohio State University Comprehensive Cancer Center Core Grant CA 16058.

We thank Paul Wilkins, Jason Davis, Dan Litvack, David Lozano, Susan Moseley, Carolyn Cheney, Julianne Dorne, Catherine Bremer, and Tricia Rigel for their excellent technical support. We also thank the personnel of the General Clinical Research Center, including Tomasina Wall, Dana Ciccone, Bob Rice, Dave Phillips, and the nursing staff headed by Teresa Sampsel, for their excellent assistance and cooperation.

Correspondence concerning this article should be addressed to Mary H. Burleson, Department of Social and Behavioral Sciences, Arizona State University West, 4701 West Thunderbird, Phoenix, Arizona 85069-7100. E-mail: mary.burleson@asu.edu

Table 1
Temporal Stability of Immune Parameters Estimated From Peripheral Blood: Previous Research

		Sample			Measurement			
Study	Interval	*N*	Age	Sex	Base	Task	Delta	Resid
CD3+ cells (total T cells)								
Marsland et al. (1995)	2 wk	30	18–30	M	.58	.55		.50
Mills, Ziegler, et al. (1995)	6 wk	20	20–41	F	.47	.46	.09	
Mills, Haeri, et al. (1995)	6 wk	24	19–43	M	.45	.33	−.16	−.06
Cohen et al. (2000)	2 wk	115	18–30	M, F	.71	.60		.24
CD4+ cells (T helper cells)								
Marsland et al. (1995)	2 wk	30	18–30	M	.59	.56		.25
Mills, Ziegler, et al. (1995)	6 wk	20	20–41	F	.57	.59	.07	
Mills, Haeri, et al. (1995)	6 wk	24	19–43	M	.61	.55	−.17	−.14
Capitanio et al. (1998)[a]	3 mo–1 yr	36	5–8	M	.20–.46	.27		
Cohen et al. (2000)	2 wk	115	18–30	M, F	.75	.63		.04
CD8+ cells (T cytotoxic/suppressor cells)								
Marsland et al. (1995)	2 wk	30	18–30	M	.69	.82		.53
Mills, Ziegler, et al. (1995)	6 wk	20	20–41	F	.70	.65	.20	
Mills, Haeri, et al. (1995)	6 wk	24	19–43	M	.54	.32	.13	.14
Capitanio et al. (1998)[a]	3 mo–1 yr	36	5–8	M	.12–.34	.27		
Cohen et al. (2000)	2 wk	115	18–30	M, F	.64	.71		.50
Ratio of CD4+ cells to CD8+ cells								
Mills, Ziegler, et al. (1995)	6 wk	20	20–41	F	.92	.88	.55	
Mills, Haeri, et al. (1995)	6 wk	24	19–43	M	.90	.92	.48	.60
Capitanio et al. (1998)[a]	3 mo–1 yr	36	5–8	M	.47–.77	.81		
Natural killer cells								
Marsland et al. (1995)[b]	2 wk	30	18–30	M	.22	.72		.42
Mills, Ziegler, et al. (1995)[c]	6 wk	20	20–41	F	.53, .19, .52	.35, .37, .53	.20, .49, .36	
Mills, Haeri, et al. (1995)[c]	6 wk	24	19–43	M	.40, .45, .80	.49, .48, .60	.51, .40, .37	−.48, .41, .34
Cohen et al. (2000)[b]	2 wk	115	18–30	M, F	.63	.76		.69
Other immune parameters								
Fletcher et al. (1992)[d]	2 wk	15	NR	NR	.85, .92			
Marsland et al. (1995)[e]	2 wk	30	18–30	M	.75, .17, .01	.74, .47, .01		.31, .50, .04
Mills, Haeri, et al. (1995)[f]	6 wk	24	19–43	M	.82	.85	.52	.48
Capitanio et al. (1998)[a,g]	3 mo–1 yr	36	5–8	M	.22–.51, .07–.37	.63, .41		
Cohen et al. (2000)[h]	2 wk	115	18–30	M, F	.77, .56	.78, .67		.08, .52

Note. Marsland et al. (1995), Mills, Haeri et al. (1995), Mills, Ziegler et al. (1995), and Cohen et al. (2000) evaluated the number (rather than percentage) of cells, used two measurement occasions, measured between 7:00 and 9:30 a.m., used a speech task as the psychological stressor, and reported Pearson correlation coefficients. Base = baseline; Resid = residualized change score; wk = week(s); mo = month(s); yr = year(s); M = male; F = Female; NR = not reported.

[a] Subjects were adult rhesus monkeys (*Macaca mulatta*); reported are range of Spearman correlations among 4 baseline measurements (3–3:30 p.m.; 3–6 mo apart) and Spearman correlation between 2 measurements taken after 120 m of restraint stress (3 p.m.; 6–7 mo apart). [b] Natural killer (NK) cells include CD56+ cells. [c] NK cells include CD16+, CD56+, and CD57+ cells; correlations are reported for each subgroup. [d] Other parameters: cell proliferation induced by the mitogens phytohemagglutinen (PHA) and pokeweed, reported respectively. [e] Other parameters: CD19+ cells (B cells) and cell proliferation induced by the mitogens PHA and concanavalin A (ConA), reported respectively. [f] Other parameters: total white blood cells. [g] Other parameters: neutrophils and total lymphocytes, reported respectively. [h] Other parameters: CD19+ cells (B cells) and NK cell cytotoxicity, reported respectively.

Three previous studies have reported reproducibility of stress-related functional immune parameters (see Table 1). Using a laboratory stress paradigm, Marsland and colleagues studied the reproducibility of young adult men's brief stress responses over a 2-week interval (Marsland, Manuck, Fazzari, Steward, & Rabin, 1995). They reported moderate to high test–retest reliability for basal and stress levels of most leukocyte subsets measured in peripheral blood, with lower reliability for change scores. In addition, posttask and change levels of lymphocyte proliferative response to the mitogen phytohemagglutinin (PHA) showed moderately low stability. Lymphocyte proliferative responses to the mitogen concanavalin A (ConA) were not stable at any measurement period. In contrast, Fletcher and colleagues reported high generalizability of lymphocyte proliferation in response to PHA and pokeweed mitogen (PWM) over a 2-week interval; however, no information was given regarding time of day or context of

blood sampling (Fletcher, Klimas, Morgan, & Gjerset, 1992). In the largest study reported, Cohen and colleagues measured the 2-week consistency of leukocyte subsets and natural killer (NK) cell cytotoxicity before and after a speech task in 115 young adults (Cohen et al., 2000). They reported moderately high to high consistency for baseline and posttask scores of all of these measures. Consistency of reactivity was moderately high for T suppressor/cytotoxic cells, NK cell numbers, and NK cell cytotoxicity, but low for total T cells, T helper cells, and B cell numbers.

Along with two of the studies summarized above, three additional studies have reported reproducibility of leukocyte subsets (one using rhesus macaques; see Table 1). Mills and colleagues investigated the reliability of enumerative measures of leukocyte subsets in peripheral blood of young adult women (Mills, Ziegler, Dimsdale, & Parry, 1995) and young adult men (Mills, Haeri, & Dimsdale, 1995). The participants underwent two sessions of brief laboratory stressors approximately 6 weeks apart. The authors reported low to moderately high test–retest reliability for basal and stress levels of most lymphocyte subsets in both studies, with generally lower reliability for change scores. In a group of 36 adult male rhesus monkeys, Capitanio and colleagues studied individual differences in leukocyte subsets under baseline and physical-restraint stress conditions (Capitanio, Mendoza, & Lerche, 1998). Baselines were measured four times spanning a year-long interval; stress conditions were measured twice approximately 6 months apart. Test–retest Spearman correlations ranged from low to high, and the ratio of T helper to T cytotoxic cells was most consistent.

Older individuals are at higher risk for many diseases, both acute and chronic. Many factors likely contribute to this vulnerability, especially the cumulative effects of stress and the aging of physical systems. Individuals who characteristically react more strongly to stress (i.e., those whose reactions are consistently more intense and persist for a longer time) may accumulate greater stress-related effects over the course of their lives. This buildup may contribute to health detriments (McEwen, 1998). In addition, older individuals may be particularly likely to experience certain psychosocial stressors, such as the need to provide care for an infirm spouse. Spousal caregivers of dementia patients report high levels of stress, dysphoria, and social isolation (George & Gwyther, 1986; Haley, Levine, Brown, Berry, & Hughes, 1987; Kiecolt-Glaser, Dura, Speicher, Trask, & Glaser, 1991). This type of caregiving has been described as similar to exposure to multiple and severe long-term stressors (Schulz & Williamson, 1991) and is associated with detrimental effects on both cellular and humoral immune function (Esterling, Kiecolt-Glaser, Bodnar, & Glaser, 1994; Kiecolt-Glaser et al., 1991; Kiecolt-Glaser, Glaser, Gravenstein, Malarkey, & Sheridan, 1996). Finally, immune function is known to change with aging. Aging of the immune system (immunosenescence) causes reductions in both cellular and humoral immune function over time (National Institute on Aging & National Institute of Allergy and Infectious Diseases, 1996), probably contributing to some forms of illness. These changes in the immune system may also render an individual more vulnerable to the effects of stress.

By influencing immune function, these same factors may also reduce the temporal consistency of immune measures, making it more difficult to evaluate their effects over time. The convergence of physical aging with the potential buildup of stress effects makes it especially important to investigate the consistency of these measurements in older individuals. In the current study, we provide the first description, for humans, of 1-year consistency of individual differences in immune variables, including numbers and percentages of circulating leukocyte subsets and several measures of cellular immune function. We evaluate the temporal consistency of both tonic levels and reactions to acute stressors and extend previous work by examining an older population. We also take a preliminary look at the possible effects of age and chronic stress on the temporal consistency of these variables.

Method

Participants

We recruited our participants from an ongoing study of the stress of caring for a spouse with progressive dementia. This large initial sample was obtained from the community by advertisement and composed mostly of women because of the demographics of this caregiver population. For the first year of reported data, there were 45 participants; a subset of 35 participants returned for reassessment in the second year. Nineteen of our Year 1 participants were spousal caregivers, and 26 were age-matched comparison participants. The age of our initial sample ranged from 49 to 83 ($M = 67.8$, $SEM = 1.2$) at entry; 38 were Caucasian and 7 were African American. Mean body mass index (BMI; calculated as weight in kilograms divided by squared height in meters) at entry was 25.9 ($SEM = 0.7$). All women were postmenopausal, and 36% were receiving estrogen replacement therapy throughout the study. All participants met the following inclusion criteria for both assessments: (a) no chronic disease; (b) no current illness; (c) on average, less than 10 hr of exercise per week; (d) on average, less than 10 alcoholic beverages per week; (e) no math, speech, or needle phobia; and (f) consistent usage or nonusage of beta-adrenergic receptor blockers for the duration of the study.

The Year 1 characteristics of the retest sample of 35 women were compared with those of the women who were not retested in Year 2. The two groups did not differ in (a) caregiver status, race, hormone use, marital status, BMI, activity level, alcohol or caffeine consumption, sleep, fear of speech or math, or (b) scores on inventories to measure depression, optimism, coping, positive and negative affect, trait anxiety, social support, loneliness, perceived stress, and number of life events. However, the returnees were older than the dropouts, $F(1, 44) = 5.14$, $p = .03$, $M = 69.1$ ($SEM = 1.3$) and $M = 63.0$ ($SEM = 2.4$) years, respectively. They also reported less fear of needles and greater effort to complete the math task. In addition, they scored significantly higher on two items from the Accepting Responsibility subscale of the Ways of Coping Scale (described below, Folkman & Lazarus, 1985). Comparison of the two groups' Year 1 immune responses revealed no differences at baseline or after the tasks and no difference in reactivity. At recovery, however, the returnees had higher percentages of CD4+ lymphocytes, lower percentages and numbers of CD8+ lymphocytes, and a higher CD4+/CD8+ ratio. Thus, older participants who (a) were less fearful of needles, (b) thought the math task required more effort, (c) were more likely to criticize or blame themselves after a stressful event, (d) were more likely to believe they had brought the stressor on themselves, and (e) had more CD4+ and fewer CD8+ lymphocytes at recovery, returned for reassessment in Year 2 of the study.

Procedures

Although this article reports only the results for immune parameters, a large panel of psychosocial, cardiovascular, and neuroendocrine variables was also measured. Because the measurement of these variables contributed to the participants' experience of the research protocol, we include some detail regarding the methods used. Results for these other variables will be reported elsewhere.

In preparation for their assessments, participants were asked (a) to reschedule their appointments if they became ill or experienced a major

negative life event; (b) to not consume any alcohol or take any nonprescription medication (e.g., antihistamines) the day before the study; (c) to refrain from exercise the day before the study; and (d) to refrain from eating, drinking anything besides water, or smoking from midnight until the time of their scheduled appointment the following morning.

All participants were tested beginning at approximately 8:00 a.m. Upon a participant's arrival, the tasks and measures were reviewed, any questions were answered, and informed consent was obtained. A strain-gauge respirometer was placed around the lower chest, and disposable band electrodes for impedance cardiography were placed around the participant's neck and chest. The participant was then asked to lie down, and a 20-gauge catheter was inserted into an antecubital vein. After catheter insertion, the participant's arm and hand were placed on a heating pad to arterialize the venous blood flow, and a wrist tonometer was placed over the radial artery of the other arm for blood pressure measurements. To adapt to the setting, the participant rested in a seated position for approximately 30 min while psychosocial questionnaires were verbally administered. After this adaptation period, participants were asked to relax quietly, and 6 min of cardiovascular and respiratory activity were recorded. Immediately following the recording, a blood sample was collected for assessment of baseline immune function.

Following the baseline measures, participants received instructions for the two stress tasks (described below), and any questions they had about the stressors were answered. The participants then carried out the stress tasks in close succession. To reduce order effects and habituation, task order was counterbalanced across participants and from Year 1 to Year 2. Measurement of cardiovascular and respiratory parameters continued during the entire 6-min speech and math stressor periods. Blood samples were drawn immediately after the first stressor for endocrine assays only and immediately after the second stressor for endocrine and immune assays. Participants rated the tasks and their responses to the tasks immediately after the second stressor. After the stress tasks, participants were asked to sit quietly and relax for 30 min. Four 3-min epochs of cardiovascular recording were evenly interspersed with three 6-min nonrecording epochs over the course of the 30-min recovery, and blood was drawn (for endocrine assays only) during the second 6-min nonrecording epoch of recovery. A final blood sample for immune and endocrine assays was drawn at the end.

Stress Tasks

Year 1. For the math stressor, participants performed six 1-min serial subtraction problems in immediate succession for 6 min. They were instructed that the experimenter would correct any errors they made and that they should continue from the correct number. The starting number from which the other number was subtracted for Minute 1 was 297, for Minute 2 was 688, for Minute 3 was 955, for Minute 4 was 593, for Minute 5 was 1,200, and for Minute 6 was 1,741. The number being subtracted in Minute 1 was 3. Results from research on mental arithmetic in older adults indicated that participants average approximately 10 serial subtractions per minute (Cacioppo et al., 1995; Uchino, Kiecolt-Glaser, & Cacioppo, 1992). To maintain maximal task involvement and moderate difficulty (i.e., about 10 correct answers per minute), we required that the subtrahend for each subsequent minute be contingent on the participant's performance in the preceding minute. In other words, better performance led to more difficult math problems, such as subtracting 7 or 13 rather than 2 or 3.

For the speech task, we used an adaptation of the speech stressor of Saab, Matthews, Stoney, and McDonald (1989). Each participant was asked to imagine that she had been taken to see a department store manager by a security guard who had falsely accused her of shoplifting a belt. She was instructed and given 3 min to prepare a 3-min speech to give to the store manager, covering the following points: (a) her side of the story (what actually happened), (b) what the security guard did that was wrong and why he or she might have suspected her of stealing the belt, (c) how she could prove she did not steal the belt, (d) what should happen to the security guard for the mistake, and (e) a summary of her points. Participants were told that their speeches would be recorded and compared with the speeches of others, and instructed to give "intelligent and well thought out" answers.

Year 2. The math task, the difficulty of which was contingent on the participant's performance, was the same as in Year 1. The speech task was structured in the same way as in Year 1, however, to reduce habituation, a different topic was used. The participant was asked to imagine that she was falsely accused of hitting another car with her car. She was given 3 min to prepare a 3-min talk to give to the police, covering a set of specific points equivalent to those used in Year 1. All other instructions were the same as in Year 1.

Immune Measures

For each year, immune assays were carried out on 3 blood samples per participant, which were taken at baseline, immediately after both stress tasks were completed and at the end of the recovery period. Identical procedures were used for Year 1 and Year 2. The Clinical Immunology Laboratory at the Ohio State University Hospital performed complete blood counts and differentials on each sample using 5 ml of EDTA-treated blood. Peripheral blood mononuclear cells (PBMCs) were obtained from heparinized blood, and the percentages of CD3+, CD4+, CD8+, and NK cells (CD56+) were determined with the use of monoclonal antibodies (Coulter; Beckman Coulter, Fullerton, CA) by fluorescence activated cell sorting analysis through routine procedures as described previously (Kiecolt-Glaser et al., 1991).

NK cell cytotoxicity was measured by incubation of various concentrations of PBMCs with ^{51}Cr-labelled K-562 target cells as previously described (Glaser, Rice, Speicher, Stout, & Kiecolt-Glaser, 1986). Briefly, PBMCs were prepared at 75:1, 37.5:1, and 18.75:1 effector-to-target (E:T) cell ratios and were seeded in triplicate in 96-well microtiter plates (Costar Corporation, Pleasanton, CA). Additional wells containing only labeled target cells (K-562) in medium or target cells in medium plus 5% sodium dodecyl sulfate were used to determine spontaneous and maximal release of radioactivity, respectively. Plates were incubated for 5 hr in a 5% CO_2 atmosphere at 37°C, and supernatants were harvested. Activity was determined by the release of ^{51}Cr into the supernatant, which was measured with a Beckman 9000 gamma counter (Beckman Coulter, Fullerton, CA). Analyses of three E:T ratios produced equivalent results; therefore only the 75:1 E:T ratio is reported.

Mitogen-stimulated activity in PBMCs was assessed using the Cell Titer 96 Aqueous Nonradioactive Cell Proliferation Assay (Promega, Madison, WI), which determines the number of viable proliferating cells by colorimetry (Cacioppo et al., 1998). Samples were set up in triplicate on 96-well plates, with ConA (Sigma; Sigma-Aldrich, St. Louis, MO) at a final concentration of 10.0 μg/ml, 5.0 μg/ml, and 2.5 μg/ml. Fifty microliters of sample cells from a stock solution at 1×10^6 cells/ml in RPMI-1640 medium, supplemented with 5% fetal bovine serum was added to 50 μl of each mitogen dilution and the media control. Plates were incubated in an atmosphere of 5% CO_2 at 37°C (with humidity), for approximately 68 hr. After incubation, 20 μl of a 20:1 solution of MTS/phenazine methosulfate was added to the plates. Plates were then incubated for 4 hr, after which the optical density was recorded directly from the plate using a Titertek (Huntsville, AL) Multiscan MCC plate reader. Plate background absorbance was removed using a reference wavelength of 650 nm, per the manufacturer's suggestion. Analyses of three ConA concentrations produced equivalent results; therefore, only the 5.0 μg/ml concentration is reported.

Antibody levels of Epstein-Barr virus (EBV) VCA IgG were determined using the indirect immunofluorescence (IF) test, as described elsewhere (Glaser et al., 1987). Briefly, acetone-fixed HR-1 cells (a Burkitt lymphoma cell line latently infected with EBV) prepared on glass cover slips were adsorbed with twofold dilutions of plasma in phosphate-buffered

saline in order to determine endpoint antibody titers. The highest dilution of serum that resulted in the detection of at least 1% IF-positive cells was considered the endpoint. Dilution values were transformed using base 10 logarithms to improve distribution.

Psychosocial Measures

Individual differences. Prior to the stress tasks and physiological data collection, participants completed the short form of the Beck Depression Inventory (Beck & Beck, 1972), the COPE (Carver, Scheier, & Weintraub, 1989), the Life Orientation Test (Scheier & Carver, 1985), the Positive and Negative Affect Scale (Watson, Clark, & Tellegen, 1988), an adaptation of the Differential Emotions Scale (DES; Cacioppo, Martzke, Petty, & Tassinary, 1988), the Spielberger Trait Anxiety Inventory (Spielberger, Gorsuch, Lushene, Vaag, & Jacobs, 1983), the Interpersonal Support Evaluation List (Mermelstein, Cohen, Lichtenstein, Kamarck, & Baer, 1986), the Social Network Interview (Cohen & Hoberman, 1983), the New York University (NYU) Loneliness Scale (Rubenstein & Shaver, 1982), the Perceived Stress Scale (Cohen, Kamarck, & Mermelstein, 1983), and an inventory of life events. Selected items from the Ways of Coping Scale (Folkman & Lazarus, 1985; Folkman, Lazarus, Dunkel-Schetter, DeLongis, & Gruen, 1986) were also given, with the following instructions: "When you are dealing with the hassles and stresses of everyday life, how often do you use the following coping strategies?" The items included three from the Accepting Responsibility subscale: "You criticize or lecture yourself," "You think you brought the problem on yourself," and "You promise yourself that things will be different next time"; two items from the Self-Controlling subscale: "You tried to keep your feelings to yourself" and "You kept others from knowing how bad things were"; and one item from the Escape–Avoidance subscale: "You avoided being with people in general." All scales were administered verbally.

Posttask ratings. After completing the stress tasks, the participants rated how pleasant versus unpleasant and how relaxed versus aroused they felt during each task, how mentally effortful each task was, their uncertainty about their ability to complete the tasks, their uncertainty about what was expected, and their feelings of helplessness and control during the tasks. They completed another DES regarding their feelings during the tasks, and they also rated their anxiety, tension, worry, upset, relaxation, and contentment during the tasks.

Data Analysis

Reactivity values were delta scores derived by subtracting the baseline value from the posttask value. We evaluated changes in mean baseline, posttask, recovery, and reactivity scores from year to year using within-subjects analysis of variance (ANOVA).

Missing data. Temporal consistency is often evaluated using Pearson correlations between the same variables measured at different occasions, with pairwise deletion for observations with missing values. As described above, 10 of the original 45 participants did not return for Year 2. In addition, problems with blood draws and immune assays caused other data to be missing (see Tables 2–5 for exact *n*s). Simulation studies have shown that conventional methods of handling missing data, such as pairwise case deletion, introduce bias to parameter estimates unless data are missing completely at random. On the other hand, maximum likelihood estimates reduce bias even in the worst case, when the data are neither missing completely at random nor missing at random, a weaker assumption about the causes of missingness (Arbuckle, 1996). Thus, for the current report, we present both year-to-year Pearson coefficients derived using pairwise deletion, and standardized maximum-likelihood estimates of year-to-year covariances derived using the full dataset. When data are complete, the maximum-likelihood method yields estimates identical to Pearson correlations.

Following the recommendations of Graham and colleagues (Graham, Hofer, Donaldson, MacKinnon, & Schafer, 1997), estimates were obtained with the program Amos 4.0 (Arbuckle & Wothke, 1999), which compensates for missing data by maximizing the likelihood function at the level of the individual. Completely saturated covariance models were used (i.e., no constraints were placed on the covariance estimates). Full information maximum-likelihood estimation uses all available data to improve parameter estimates involving variables with incomplete data; variables that are highly correlated with the incomplete data are especially useful in this regard (Arbuckle, 1996). Because of the relatively small sample size, a maximum of only six variables could be included in a given model, so a separate model was used for each immune variable. Because they tend to be correlated with each other and with Year 2 values, baseline, posttask, and recovery values at Year 1 were included in each model for estimating year-to-year consistency of baseline, posttask, or recovery values. For example, for estimating the year-to-year covariance for posttask CD4+/CD8+ ratios, the model included Year 1 baseline, posttask, and recovery values of CD4+/CD8+ ratios, along with the Year 2 value of the posttask CD4+/CD8+ ratio. For estimating year-to-year covariances of delta scores, the models included Year 1 values of delta and recovery, along with the Year 2 value of delta (again because of their predictive value). In addition, because age and fear of needles in Year 1 were related to missingness for the 10 dropouts, these two variables were included in all models (Graham et al., 1997). As expected for fully saturated models, all models reached convergence without problems.

Effects of age and caregiver stress. To evaluate the possible relation between age and temporal consistency, we dichotomized the sample at the median (68 years) into younger ($M = 61.50$, $SEM = 1.10$, $n = 23$) and older ($M = 74.40$, $SEM = 0.90$, $n = 24$) subsamples. Using Amos, models (of the structure described above) were estimated in which the year-to-year covariance between Year 1 and Year 2 values of the variable of interest was constrained to be equal in the two subsamples. Poor fit of the constrained model (compared with the saturated model) provides evidence that the covariances differ in the two age groups, implying an effect of age on the covariance. We used a similar strategy to examine the effects of caregiver stress (caregivers and comparison participants did not differ in age). All models converged without problems.

Age and stress effects on changes in the mean baseline, posttask, recovery, and reactivity scores from year to year were evaluated using mixed ANOVA.

Results

Temporal Consistency Across Years

Year-to-year Pearson correlations, along with standardized maximum-likelihood estimates of year-to-year covariances, are shown for baseline, posttask, recovery, and reactivity scores in Table 2.

Baseline, posttask, and recovery values. For T cell subset numbers and percentages, Pearson *r* values were consistently moderately high to high for baseline and posttask values. Recovery values were less consistent, ranging from low to high. Pearson coefficients for the CD4+/CD8+ ratios were high or moderately high at all three time points. Very similar patterns were seen in the standardized covariance estimates. NK cell number and percentage had moderate to moderately high Pearson correlations, and again, standardized covariance estimates based on the full sample showed a similar pattern, although the values appeared somewhat higher than for the Pearson coefficients. Both Pearson correlations and standardized covariance estimates for ConA-induced proliferation were low. For percentage of monocytes and NK cell cytotoxicity, Pearson *r* values ranged from low to moderately low. Standardized covariance estimates were similar, although slightly higher. Because EBV VCA IgG antibody titers will not change within a 3-hr

Table 2
Pearson Correlations and Standardized Maximum-Likelihood Covariance Estimates for Year 1 With Year 2 Immune Variables

	Pearson correlation								Standardized maximum-likelihood covariance estimate (n = 45)			
	Baseline		Posttask		Recovery		Delta					
Measure	n	r	n	r	n	r	n	r	Baseline	Posttask	Recovery	Delta
Lymphocyte subsets												
No. total T cells	26	.72**	27	.81**	26	.55**	24	.06	.67	.78	.51	.09
No. CD4+ cells	26	.82**	27	.88**	26	.67**	24	.43*	.82	.88	.68	.44
No. CD8+ cells	26	.73**	27	.78**	26	.60**	24	.50*	.70	.80	.61	.60
% total T cells	28	.83**	29	.79**	27	.27	27	.20	.85	.83	.26	.34
% CD4+ cells	28	.85**	29	.89**	27	.55**	27	.37†	.90	.94	.65	.53
% CD8+ cells	29	.77**	30	.85**	28	.79**	28	.41*	.82	.89	.83	.36
CD4+/CD8+ ratio	30	.93**	31	.96**	28	.84**	29	.54*	.92	.96	.78	.83
Other immune variables												
No. NK cells	25	.76**	26	.64**	25	.37†	23	.32	.89	.72	.56	.18
% NK cells	29	.73**	30	.56**	27	.35†	28	.07	.84	.58	.42	.21
% monocytes	30	.41*	31	.46**	28	.33†	29	−.14	.59	.63	.38	−.07
ConA proliferation	32	.22	32	.02	31	.42*	32	−.05	.21	−.01	.40	−.08
NK cell cytotoxicity	25	.38†	23	.31	21	.38†	23	−.02	.37	.24	.33	−.05
EBV antibody titers	31	.91**	—	—	—	—	—	—	.92	—	—	—

Note. For Pearson correlations, number of participants varied because of technical problems with data acquisition or assay. All standardized covariance estimates were derived on the basis of the number of participants in the initial sample. Dashes indicate data were not collected. NK = natural killer; ConA = concanavalin A; EBV = Epstein-Barr virus.
† $p < .10$. * $p < .05$. ** $p < .01$.

study (the half life of IgG is approximately 20 days), they were measured only at baseline. Year-to-year temporal consistency was high as indicated by both Pearson coefficients and standardized covariance estimates.

Reactivity scores. Year-to-year Pearson correlations and standardized covariance estimates for change in the number and percentage of CD3+, CD4+, and CD8+ T cells, and the CD4+/CD8+ ratios ranged from low to moderate. Pearson correlations for change in both number and percentage of NK cells were low. Pearson correlations for change in ConA-induced proliferation, monocyte percentage, and NK cell cytotoxicity were very low. Standardized maximum-likelihood covariance estimates showed the same patterns.

Effects of age. Results suggesting differences in temporal consistency between younger and older subsamples were found for only 3 of the 49 variables tested (see Table 3): baseline and delta CD4+/CD8+ ratios, and delta percentage of CD3+ lymphocytes. Interestingly, the older subsample had higher temporal consistency in all three cases; however, because the sample sizes of our subgroups were so small, and the age range was reduced by attrition, results of these analyses cannot be considered conclusive.

Effects of caregiver stress. Results suggesting differences in temporal consistency between caregivers and comparison participants were found for seven variables (see Table 3). For five of these, including posttask and recovery values of percentage of CD3+ and CD4+ lymphocytes, and baseline values of NK cell number, caregivers had higher temporal consistency than comparison participants. For the remaining two variables, recovery NK cell cytotoxicity and posttask percentage of monocytes, comparison participants had higher temporal consistency.

Change in Mean Responses Across Years

Means and standard errors for all variables at Year 1 and Year 2 are shown in Table 4. CD4+ cell percentages, the CD4+/CD8+ ratios, and NK cell percentage were significantly higher in Year 2

Table 3
Standardized Maximum-Likelihood Covariance Estimates That Differed for Age and Caregiver Subsamples

Age effects		
Measure	Younger (n = 23)	Older (n = 22)
Baseline CD4+/CD8+ ratios	.84	.97
Delta CD4+/CD8+ ratios	.38	.94
Delta % total T cells	−.54	.85
Caregiver stress effects		
	Caregiver (n = 19)	Comparison (n = 26)
Baseline number NK cells	.94	.65
Posttask % total T cells	.98	.60
Posttask % CD4+ cells	.98	.85
Posttask % monocytes	−.30	.75
Recovery % total T cells	.86	.02
Recovery % CD4+ cells	.82	.33
Recovery % NK cells	−.01	.57

Note. All standardized covariance estimates were derived on the basis of the number of participants in the initial sample. NK = natural killer.

Table 4
Mean (and SEM) Baseline, Posttask, Recovery, and Delta Change Levels of Immune Activity as a Function of Year in Study

	Baseline value					Posttask value					Recovery value					Delta change value				
		Year 1		Year 2			Year 1		Year 2			Year 1		Year 2			Year 1		Year 2	
Measure	*n*	*M*	*SEM*	*M*	*SEM*	*n*	*M*	*SEM*	*M*	*SEM*	*n*	*M*	*SEM*	*M*	*SEM*	*n*	*M*	*SEM*	*M*	*SEM*
Lymphocyte subsets																				
No. total T cells	26	$1{,}203.0_a$	90.0	$1{,}191.7_a$	84.9	27	$1{,}290.1_a$	95.8	$1{,}263.5_a$	97.0	26	$1{,}464.6_a$	105.1	$1{,}423.3_a$	100.7	24	156.4_a	31.3	126.2_a	40.4
No. CD4+ cells	26	870.1_a	70.7	871.8_a	66.0	27	894.3_a	77.0	918.1_a	77.3	26	$1{,}077.6_a$	83.6	$1{,}043.3_a$	75.3	24	84.5_a	21.4	93.5_a	33.4
No. CD8+ cells	26	325.0_a	30.6	324.0_a	37.9	27	389.9_a	36.6	380.9_a	38.9	26	370.4_a	32.9	380.6_a	42.3	24	82.9_a	15.2	70.4_a	14.39
% total T cells	28	75.5_a	1.3	75.3_a	1.2	29	73.1_a	1.3	71.6_a	1.5	27	76.6_a	1.2	75.7_a	1.8	27	-2.5_a	0.7	-2.9_a	0.6
% CD4+ cells	28	54.4_a	1.6	56.5_b	1.5	29	51.1_a	1.6	53.2_b	1.4	27	56.5_a	1.6	56.8_a	1.7	27	-2.8_a	0.8	-2.5_a	0.7
% CD8+ cells	29	20.8_a	1.4	19.3_a	1.2	30	21.5_a	1.1	20.8_a	1.1	28	19.4_a	1.3	19.3_a	1.3	28	0.6_a	0.9	1.5_a	0.5
CD4+/CD8+ ratio	30	2.9_a	0.3	3.2_b	0.3	31	2.6_a	0.2	2.8_b	0.3	28	3.2_a	0.3	3.5_a	0.4	26	-0.2_a	0.1	-0.4_a	0.1
Other immune variables																				
No. NK cells	25	221.0_a	19.2	250.7_a	24.3	26	298.8_a	36.5	358.7_a	43.9	25	222.7_a	21.9	260.7_a	28.1	23	89.1_a	27.3	122.9_a	27.9
% NK cells	29	11.0_a	0.7	12.7_b	1.0	30	13.9_a	1.2	16.6_b	1.3	27	10.3_a	0.9	11.7_a	0.8	28	2.9_a	0.8	4.2_a	0.8
% monocytes	30	13.9_a	0.7	14.6_a	1.0	31	13.0_a	0.6	13.4_a	0.9	28	13.2_a	0.9	12.7_a	0.9	29	-0.8_a	0.5	-1.0_a	0.5
ConA proliferation	32	0.15_a	0.03	0.14_a	0.02	32	0.12_a	0.02	0.13_a	0.02	31	0.15_a	0.02	0.15_a	0.02	22	-0.03_a	0.03	-0.02_a	0.02
NK cell cytotoxicity	25	59.7_a	4.0	56.5_a	3.1	23	66.4_a	4.4	58.4_a	3.1	21	58.5_a	3.7	47.7_b	2.7	23	6.9_a	2.3	1.8_a	3.2
EBV antibody titer	31	2.2_a	0.1	2.5_b	0.1	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—

Note. Means with different subscripts are significantly different at $p < .05$. Number of participants varied because of technical problems with data acquisition or assay. Dashes indicate data were not collected. NK = natural killer; ConA = concanavalin A; EBV = Epstein-Barr virus.

at baseline and posttask. NK cell cytotoxicity at recovery was significantly lower in Year 2. EBV antibody titers measured at baseline was significantly higher in Year 2. Mean reactivity scores showed no significant changes from Year 1 to Year 2.

Effects of age and caregiver stress were as follows. Means and standard errors for variables that differed from year to year in the different age and chronic stress groups are shown in Table 5. In the older group, recovery values of NK cell numbers went down from Year 1 to Year 2, whereas they increased in the younger group. EBV VCA IgG antibody titers increased more from Year 1 to Year 2 in caregivers. Posttask value of the CD4+/CD8+ ratios increased in comparison participants but did not change in caregivers. Recovery values of monocyte percentage decreased from Year 1 to 2 in caregivers and increased slightly in the comparison group. Finally, CD3+ cell number at recovery, along with both number and percentage of CD8+ cells at recovery, decreased from Year 1 to Year 2 in the comparison participants and increased in the caregivers.

Discussion

The current study provides the first report, for humans, of 1-year temporal consistency of individual differences in a set of immune variables that often are studied in the context of psychological stress. As summarized above, these results replicate and extend prior research on the reproducibility of immune parameters. We measured baseline, posttask, recovery, and stress-induced change for levels of circulating leukocyte subsets and indicators of cellular immune function in a group of older women, with a retest interval of 1 year. To our knowledge, only four previous human studies have reported temporal consistency for the enumerative measures of leukocyte subsets. All of these were carried out on young to middle-aged adults, 6 weeks was the longest retest interval, and recovery from stress was not measured. Only three previous studies have reported retest reliability of cell proliferation induced by mitogens or NK cell cytotoxicity, and none has reported temporal consistency data for the percent of monocytes in peripheral blood or a measure for stress and viral latency.

The current report uses individual-level maximum-likelihood estimation to account for missing data, allowing a separate set of consistency estimates that is based on a larger sample size. For the most part, these estimates were very similar to those derived using Pearson correlations and pairwise deletion. In addition, through comparison of alternative models, this estimation method also allowed us to examine possible effects of age and caregiver stress on the temporal consistency of the immune variables. Analyses revealed few significant effects on temporal stability. The sample sizes in these analyses were small, however, and the age range was reduced by attrition, so these results should be interpreted cautiously.

Baseline Measures

Baseline values of the leukocyte subsets demonstrated moderately high to high year-to-year Pearson correlations and standardized covariance estimates, except for monocyte percent, which was only moderately consistent. Although mean CD4+ and NK cell percentages and CD4+/CD8+ ratios increased from year to year, the high consistency estimates suggest that individual differences in tonic levels of these variables are stable across time. In general, despite the older age of the group and the chronic stress experienced by a subset of the participants, temporal consistency for these variables was equal to or higher than reported in previous research and shown in Table 1. In fact, results from the maximum-likelihood models suggest that the consistency of only two baseline variables differed between age and caregiver groups: lower consistency of baseline CD4+/CD8+ ratios was found in the younger subsample, and caregivers had higher consistency of baseline NK cell numbers.

In contrast to the leukocyte subsets, baseline temporal consistency for the lymphocyte proliferation and cytotoxicity measures was much lower and more variable than for the enumerative

Table 5
Year 1 and Year 2 Means and Standard Errors That Differed for Age and Caregiver Subsamples

		Year 1		Year 2			Year 1		Year 2	
Measure	*n*	*M*	*SEM*	*M*	*SEM*	*n*	*M*	*SEM*	*M*	*SEM*
	Age effects									
	Younger					Older				
Recovery no. NK cells	14	224.4	29.9	313.6	34.6	11	220.7	33.8	193.5	39.0
	Caregiver stress effects									
	Caregiver					Comparison				
EBV antibody titers	11	2.4	0.2	2.9	0.2	20	2.1	0.2	2.3	0.2
Posttask CD4+/CD8+ ratio	12	2.4	0.4	2.4	0.4	19	2.7	0.3	3.0	0.3
Recovery % monocytes	11	14.1	1.5	10.8	1.5	17	12.6	1.2	13.9	1.2
Recovery no. total T cells	11	1,442.5	164.7	1,632.2	147.7	15	1,480.8	141.1	1,270.0	126.5
Recovery no. CD8+ cells	11	371.7	51.6	487.6	59.9	15	369.5	44.2	302.2	51.3
Recovery % CD8+ cells	11	20.1	2.1	22.1	2.0	17	18.9	1.6	17.4	1.6

Note. Number of participants varied because of technical problems with data acquisition or assay. NK = natural killer; EBV = Epstein-Barr virus.

measures. This low reproducibility was similar to results from Marsland et al. (1995), and may be due, in part, to technical issues. It is impossible to control completely for all laboratory conditions (Schleifer, Eckholdt, Cohen, & Keller, 1993), particularly when waves of data are analyzed separately. Longitudinal studies of functional immune measures might improve consistency by assaying all waves of data at the same time. However, this would likely require use of frozen cells, which may, in itself, produce variability through differential effects of freezing/thawing on different cell populations. Separation and incubation of mononuclear cells from whole blood may also produce variability, and whole-blood assays may be preferable when feasible. Indeed, both previous studies in which researchers used whole-blood methods for functional assays reported higher reproducibility (Cohen et al., 2000; Fletcher et al., 1992).

Year-to-year consistency for baseline EBV antibody titers, as a measure of the control of viral latency, was high. Although generally constant, EBV antibody titers do increase gradually with age (Glaser et al., 1985), consistent with our finding that the sample mean EBV antibody titer increased from Year 1 to Year 2. In addition, immune suppression caused by psychological stress can result in reactivation of latent EBV, with a concomitant increase in EBV antibody titers (Glaser & Kiecolt-Glaser, 1994a). This is consistent with our finding that mean EBV antibody titers increased more from Year 1 to Year 2 in the caregivers than in the comparison group. Nevertheless, individual differences in EBV antibody titers were found to be relatively stable across a 1-year interval.

Posttask Measures

For the most part, previous researchers have found posttask temporal consistency of leukocyte subsets in the same range as for baseline measures (see Table 1). Our results are consistent with previous findings; for the leukocyte subsets, year-to-year temporal consistency of posttask values tended to be similar to or slightly higher than what was reported for baseline, although both number and percentage of NK cells were slightly lower. Changes in means from year to year were also similar to those for baseline values.

The few previous reports of reproducibility for functional immune measures have not shown a consistent pattern. In Marsland et al. (1995), stability of PHA-induced proliferation was low for baseline and moderately low for posttask, whereas stability of ConA-induced proliferation was low at both time points. On the other hand, Fletcher et al. (1992) found high 2-week reliability using both ConA and pokeweed mitogen. Cohen et al. (2000) found moderately high consistency for NK cell cytotoxicity at both baseline and posttask. Results from the current study are most similar to those of Marsland et al. (1995). Consistency of posttask values was lower than baseline for ConA-induced proliferation but similar to baseline for NK cell cytotoxicity.

The maximum-likelihood models suggested that caregiver stress was related to temporal consistency for three posttask measures. Percentage of monocytes was more consistent in the comparison group. On the other hand, both percentage of total T cells and percentage of CD4+ T cells had higher year-to-year consistency in caregivers than in comparison participants. If chronic stress does cause long-term changes in lymphocyte subsets, the caregivers may have reached an asymptotic level, leading to reduced variability. Only one mean posttask value was related to chronic stress: the yearly mean CD4+/CD8+ ratio did not change in the caregivers but increased in the comparison group. Age within the sample was not related to any of the posttask measures. However, because our age and caregiver subsamples were small, replication of these findings is important.

Recovery Measures

Because different physiological pathways may "recover" from stress at different rates, simultaneous measurement of recovery of different variables can be difficult. Selection of the recovery interval can influence whether recovery values appear reliable or not. The study of recovery from stress is nonetheless important, particularly for researchers in aging. For example, reduction in resiliency of physiological systems characterizes one type of stress effect that may contribute to the aging process (McEwen, 1998).

In general, recovery values in this study demonstrated somewhat lower and less consistent year-to-year correlations and standardized covariance estimates than did baseline and posttask values. For the lymphocyte subsets, most values were moderate, although percentage of CD8+ cells and the CD4+/CD8+ ratios showed high consistency for recovery values, just as they did for baseline and posttask values. On the other hand, consistencies for percentage of T cells, number and percentage of NK cells, and percentage of monocytes were low.

As seen for baseline and posttask values, year-to-year consistency of recovery values for enumerative assays was generally higher than for lymphocyte proliferation and NK cell cytotoxicity assays. Interestingly, recovery levels of mitogen-induced proliferation were more consistent than either baseline or posttask levels. Very few mean recovery values differed from year to year. Mean NK cell cytotoxicity at recovery was lower in Year 2 than Year 1. In addition, caregivers had higher mean CD3+ and CD8+ cells at recovery. Maximum-likelihood models suggested that consistency of NK cell cytotoxicity was related to caregiver status, with comparison participants demonstrating higher consistency than caregivers.

Measures of Stress-Induced Change

In the four previous reports of the reliability of stress-related changes in immune variables, consistency coefficients ranged from low to moderate (see Table 1). Results of the current study are generally within the range of previous findings. Because of the compounding of measurement error inherent in the computation of change scores (Strube, 1990), we expected year-to-year consistency of the reactivity scores to be lower than those for baseline, posttask, or recovery levels. This was true for all of the variables measured. In fact, change scores demonstrated low year-to-year Pearson correlations for all but two of the measured variables (CD8+ cell number and CD4+/CD8+ ratios, which had moderate r values). Standardized covariance estimates, based on the entire sample, were somewhat higher, most notably for CD4+/CD8+ ratios. Only this particular coefficient reached a conventionally high level (i.e., $\geq .80$).

Conclusions

In practical terms, what does this mean for researchers studying stress effects on immune function? At minimum, it indicates that

special care must be taken to reduce measurement error in longitudinal studies of stress-induced change in immune function, at least in middle-aged and older women. Aggregation to reduce measurement error is particularly useful with change scores; hence, it may be important to design studies with multiple measures of the same variables whenever possible. For functional immune measures, simultaneous assay of all waves of samples may also be desirable. Because different protocols can invoke different types and degrees of stress response, the apparent reliability of physiological measures of stress-induced change depends to a great extent on the protocol used to elicit them. The nature and timing of the stress tasks can influence the participants' interpretations and performance and their range of responses. This may be particularly true with older individuals, whose experience and perspective may lead them to view laboratory experiments differently from younger participants. Thus, it may be especially important with older participants to use laboratory stressors designed to maximize shared interpretation when assessing reactivity (e.g., see Cacioppo et al., 1995).

A lack of temporal stability in reactivity measures could have implications for the effects on health of short-term immune responses to minor stress. Some disease processes (e.g., cancer) appear to require years to develop. In these cases, low temporal stability would suggest that immunological reactivity to mild stressors is not likely to be a major determinant of individual differences in health outcomes. On the other hand, for some illnesses (e.g., upper respiratory tract infections), a relatively short-term change in immune function may be all that is required to increase susceptibility. In addition, laboratory stressors allow one to gauge individual differences in reactivity but not differential exposure to stressors. A second pathway by which individual differences may contribute to immunological compromise and disease is in terms of differences in the type, frequency, intensity, and duration of exposure to stressors during daily life. The development of experience sampling methodologies now makes it possible to examine this alternative pathway.

In contrast to the change scores, many of the baseline, posttask, and recovery values (with the general exception of the lymphocyte proliferation and NK cell cytotoxicity assays) were highly consistent over the 1-year interval between measurements. These findings support the usefulness of longer term studies of individual differences in tonic levels of immune function and of the effects of social and psychological contexts on these variables across years. They also suggest that researchers interested in short-term responses to stress might profit by further exploring the predictive value of posttask scores, in addition to change scores, on health-related outcomes.

Several potential limitations of the current study deserve mention. First, we note that our final sample differed in some ways from the original group. Because the nonreturnees were younger, the age range of our final sample was reduced, which may have biased or reduced the sensitivity of our tests of age effects on consistency of immune function. Further, the returnees were less distressed by venipuncture than the nonreturnees, and their mean levels of circulating lymphocyte subsets 30 min after the stress tasks were consistent with greater recovery from the stressors. These differences could be interpreted to suggest that the returnees were less stressed by the first session than those who left the study. If so, systematic bias could be introduced. On the other hand, the only other difference in reported psychological responses to the procedures was that the returnees thought the math task required more effort than did those who dropped out, and there were no differences between dropouts and returnees in immune reactivity to the stressors. Finally, measures of a number of psychosocial characteristics revealed that the groups differed on only two coping items, in which the returnees reported themselves more likely to accept responsibility for stressors in their daily lives. Thus, it is difficult to reach a conclusion regarding the overall meaning of the reported differences between our final sample and the original group. The current findings nevertheless provide the first information about longer term consistency of immune parameters in middle-aged and older women.

Second, and related to the first issue, is the fact that our overall sample size was small ($N = 35$), and that technical problems during data collection and assay reduced the number of participants used for some Pearson correlations to the low 20s. Although these sample sizes are well within the range of most previous studies (see Table 1), we acknowledge that small sample sizes may reduce the reliability of parameter estimates. On the other hand, the maximum-likelihood covariance estimates of temporal consistency, which were based on a sample size of 45, were very similar to the Pearson correlations. This close correspondence between the two types of estimates increases our confidence in the results.

Finally, these results cannot strictly be interpreted as test–retest reliability estimates, because the laboratory stress protocols were changed slightly from year to year to reduce the effects of habituation and to more closely approximate the variety of real-life everyday stressors. Our interest in this study was to examine individual differences in reactivity to daily stressors, with brief laboratory stressors serving as a surrogate for everyday stressors. The laboratory stressors varied across the two sessions to approximate the variations in daily stressors to which people are exposed over time. If temporal stability in immunological responses to laboratory stressors requires that the context and stressors be identical across time, then such stability estimates may contribute little to our understanding of the mechanisms underlying the effects of stress on health.

References

Arbuckle, J. (1996). Full information estimation in the presence of incomplete data. In G. Marcoulides & R. Schumacker (Eds.), *Advanced structural equation modeling: Issues and techniques* (pp. 243–277). Mahwah, NJ: Erlbaum.

Arbuckle, J., & Wothke, W. (1999). *Amos 4.0 user's guide*. Chicago: SmallWaters Corporation.

Baum, A., & Posluszny, D. (1999). Health psychology: Mapping biobehavioral contributions to health and illness. *Annual Review of Psychology, 50,* 137–163.

Beck, A. T., & Beck, R. W. (1972). Screening depressed patients in family practice: A rapid technique. *Postgraduate Medicine, 52,* 81–85.

Cacioppo, J. T., Malarkey, W. B., Kiecolt-Glaser, J. K., Uchino, B. N., Sgoutas-Emch, S. A., Sheridan, J. F., et al. (1995). Heterogeneity in neuroendocrine and immune responses to brief psychological stressors as a function of autonomic cardiac activation. *Psychosomatic Medicine, 57,* 154–164.

Cacioppo, J. T., Martzke, J. S., Petty, R. E., & Tassinary, L. G. (1988). Specific forms of facial EMG response index emotions during an interview: From Darwin to the continuous flow hypothesis of affect-laden information processing. *Journal of Personality and Social Psychology, 54,* 592–604.

Cacioppo, J. T., Poehlmann, K. M., Kiecolt-Glaser, J. K., Malarkey, W. B.,

Burleson, M. H., Berntson, G. G., & Glaser, R. (1998). Cellular immune responses to acute stress in female caregivers of dementia patients and matched controls. *Health Psychology, 17,* 182–189.

Capitanio, J. P., Mendoza, S. P., & Lerche, N. W. (1998). Individual differences in peripheral blood immunological and hormonal measures in adult male rhesus macaques (*Macaca mulatta*): Evidence for temporal and situational consistency. *American Journal of Primatology, 44,* 29–41.

Carver, C. S., Scheier, M. F., & Weintraub, J. K. (1989). Assessing coping strategies: A theoretically based approach. *Journal of Personality and Social Psychology, 56,* 267–283.

Cohen, S., Hamrick, N., Rodriguez, M. S., Feldman, P. J., Rabin, B. S., & Manuck, S. B. (2000). The stability of and intercorrelations among cardiovascular, immune, endocrine, and psychological reactivity. *Annals of Behavioral Medicine, 22,* 171–179.

Cohen, S., & Hoberman, H. M. (1983). Positive events and social supports as buffers of life change stress. *Journal of Applied Social Psychology, 13,* 99–125.

Cohen, S., Kamarck, R., & Mermelstein, R. (1983). A global measure of perceived stress. *Journal of Health and Social Behavior, 24,* 385–396.

Esterling, B. A., Kiecolt-Glaser, J. K., Bodnar, J. C., & Glaser, R. (1994). Chronic stress, social support, and persistent alterations in the natural killer cell response to cytokines in older adults. *Health Psychology, 13,* 291–298.

Fletcher, M., Klimas, N., Morgan, R., & Gjerset, G. (1992). Lymphocyte proliferation. In N. Rose, E. DeMacario, J. Fahey, H. Friedman, & G. Penn (Eds.), *Manual of clinical laboratory immunology* (4th ed. pp. 213–219). Washington, DC: American Society for Microbiology.

Folkman, S., & Lazarus, R. S. (1985). If it changes, it must be a process: Study of emotion and coping during three stages of a college examination. *Journal of Personality and Social Psychology, 48,* 150–170.

Folkman, S., Lazarus, R. S., Dunkel-Schetter, C., DeLongis, A., & Gruen, R. J. (1986). Dynamics of a stressful encounter: Cognitive appraisal, coping, and encounter outcomes. *Journal of Personality and Social Psychology, 50,* 992–1003.

George, L. K., & Gwyther, L. P. (1986). Caregiver well-being: A multidimensional examination of family caregivers of demented adults. *The Gerontologist, 26,* 253–259.

Glaser, R., & Kiecolt-Glaser, J. (1994a). Stress-associated immune modulation and its implications for reactivation of latent herpesviruses. In R. Glaser & J. Jones (Eds.), *Human herpesvirus infections* (pp. 245–270). New York: Dekker.

Glaser, R., & Kiecolt-Glaser, J. K. (1994b). *Handbook of human stress and immunity.* San Diego, CA: Academic Press.

Glaser, R., Rabin, B., Chesney, M., Cohen, S., & Natelson, B. (1999). Stress-induced immunomodulation: Implications for infectious diseases? *Journal of the American Medical Association, 281,* 2268–2270.

Glaser, R., Rice, J., Sheridan, J., Fertel, R., Stout, J., Speicher, C. E., et al. (1987). Stress-related immune suppression: Health implications. *Brain, Behavior, and Immunity, 1,* 7–20.

Glaser, R., Rice, J., Speicher, C. E., Stout, J. C., & Kiecolt-Glaser, J. K. (1986). Stress depresses interferon production by leukocytes concomitant with a decrease in natural killer cell activity. *Behavioral Neuroscience, 100,* 675–678.

Glaser, R., Strain, E., Tarr, K., Holliday, J., Donnerberg, R., & Kiecolt-Glaser, J. (1985). Changes in Epstein-Barr virus antibody titers associated with aging. *Proceedings of the Society for Experimental and Biological Medicine, 179,* 352–355.

Graham, J., Hofer, S., Donaldson, S., MacKinnon, D., & Schafer, J. (1997). Analysis with missing data in prevention research. In K. Bryant, M. Windle, & S. West (Eds.), *The science of prevention: Methodological advances from alcohol and substance abuse research* (pp. 325–366). Washington, DC: American Psychological Association.

Haley, W. E., Levine, E. G., Brown, S. L., Berry, J. W., & Hughes, G. H. (1987). Psychological, social, and health consequences of caring for a relative with senile dementia. *Journal of the American Geriatrics Society, 35,* 405–411.

Kiecolt-Glaser, J. K., Dura, J. R., Speicher, C. E., Trask, O. J., & Glaser, R. G. (1991). Spousal caregivers of dementia victims: Longitudinal changes in immunity and health. *Psychosomatic Medicine, 53,* 345–362.

Kiecolt-Glaser, J. K., Glaser, R., Gravenstein, S., Malarkey, W. B., & Sheridan, J. (1996). Chronic stress alters the immune response to influenza virus vaccine in older adults. *Proceedings of the National Academy of Sciences, USA, 93,* 3043–3047.

Maier, S., & Watkins, L. (1998). Cytokines for psychologists: Implications of bidirectional immune-to-brain communication for understanding behavior, mood, and cognition. *Psychological Review, 105,* 83–107.

Marsland, A. L., Manuck, S. B., Fazzari, T. V., Steward, C. J., & Rabin, B. S. (1995). Stability of individual differences in cellular immune responses to acute psychological stress. *Psychosomatic Medicine, 57,* 295–298.

McEwen, B. S. (1998). Protective and damaging effects of stress mediators. *The New England Journal of Medicine, 338,* 171–179.

Mermelstein, R., Cohen, S., Lichtenstein, E., Kamarck, T., & Baer, J. (1986). Social support and smoking cessation maintenance. *Journal of Consulting and Clinical Psychology, 54,* 447–453.

Mills, P. J., Haeri, S. L., & Dimsdale, J. E. (1995). Temporal stability of acute stressor-induced changes in cellular immunity. *International Journal of Psychophysiology, 19,* 287–290.

Mills, P. J., Ziegler, M. G., Dimsdale, J. E., & Parry, B. L. (1995). Enumerative immune changes following acute stress: Effect of the menstrual cycle. *Brain, Behavior, and Immunity, 9,* 190–195.

National Institute on Aging & National Institute of Allergy and Infectious Diseases. (1996). *Report of the Task Force on Immunology and Aging* (NIH Publication No. 96–4018). Washington, DC: U.S. Department of Health and Human Services.

Rubenstein, C., & Shaver, P. (1982). The experience of loneliness. In L. A. Peplau & D. Perlman (Eds.), *Loneliness: A sourcebook of current theory, research, and therapy.* New York: Wiley.

Saab, P. G., Matthews, K. A., Stoney, C. M., & McDonald, R. J. (1989). Premenopausal and postmenopausal women differ in their cardiovascular and neuroendocrine responses to behavioral stressors. *Psychophysiology, 26,* 270–280.

Scheier, M. F., & Carver, C. S. (1985). Optimism, coping, and health: Assessment and implications of generalized outcome expectancies. *Health Psychology, 4,* 219–247.

Schleifer, S., Eckholdt, H., Cohen, J., & Keller, S. (1993). Analysis of partial variance (APV) as a statistical approach to control day to day variation in immune assays. *Brain, Behavior, and Immunity, 7,* 243–252.

Schulz, R., & Williamson, G. M. (1991). A 2-year longitudinal study of depression among Alzheimer's caregivers. *Psychology and Aging, 6,* 569–578.

Spielberger, C. D., Gorsuch, R. L., Lushene, R. E., Vaag, P. R., & Jacobs, G. A. (1983). *Manual for the state-anxiety inventory (Form Y) (Self-Evaluation Questionnaire).* Palo Alto, CA: Consulting Psychologists Press.

Strube, M. J. (1990). Psychometric principles: From physiological data to psychological constructs. In J. T. Cacioppo & L. G. Tassinary (Eds.), *Principles of psychophysiology: Physical, social, and inferential elements* (pp. 34–57). New York: Cambridge University Press.

Uchino, B. N., Kiecolt-Glaser, J. K., & Cacioppo, J. T. (1992). Age-related changes in cardiovascular response as a function of a chronic stressor and social support. *Journal of Personality and Social Psychology, 63,* 839–846.

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology, 54,* 1063–1070.