Supporting Information

Online-only material

Dependent variables

The data collections included many different psychometric measures some of which are irrelevant to the present analyses. The following criteria were used in selecting variables for the present analyses. For the stressor variables in Table 2, we include ones that are clearly related to deployment for which we have the most responses. Because our stressor measures are subjective assessments, we use the ones obtained at Deployment and Return in order to reduce retrospection as much as possible. For the pre-deployment measures, we included the ones that are theoretically and empirically most closely associated with PTSD. We left out variables that were either highly correlated with the ones included, or were obtained from military data files with many missing data points, or were theoretically less relevant for the present analyses.

PCL-C. Posttraumatic Stress Disorder Checklist, Civilian version (Blanchard et al., 2006) is a standardized checklist for PTSD symptoms, which has been used in previous studies of deployment-related PTSD. The scale consists of 17 items (equivalent to the 17 PTSD symptom categories in the DSM-IV) with a possible sum score ranging from 17 to 85 (Cronbach's Alpha = .88).

BDI-II. Beck's Depression Inventory (BDI-II; Beck et al., 1996) measures level of depression. The inventory consists of 21 items with a possible sum score ranging from 0 to 63 (Cronbach's Alpha = .85).

Neuroticism. The scale for Neuroticism is one of the five personality factors measured by the NEO-FFI (Costa & McCrae,1989). The neuroticism factor consists of 12 items with a possible sum score ranging from 12 to 60. (Cronbach's Alpha = .77)

TLEQ. Trauma Life Experience Questionnaire (TLEQ; Kubana et al., 2000) was here accommodated to a Danish population of soldiers. Here it consists of 20 questions addressing different types of traumas. The questionnaire measures how often (from never to more than 5 times) different traumas have been experienced with a possible sum score ranging from 0 to 100 (Cronbach's Alpha = .70).

Combat Exposure Scale (Keane, 1989) is a standard measure of exposure to combat such as episodes with hostile fire, firing at the enemy or fellow soldiers being injured or killed during combat. The scale consists of 7 items, possible sum score ranges from 7 to 35. (Cronbach's Alpha = .83).

The Danger/Injury Exposure Scale was constructed by the Danish military for the present and similar studies. It measures the degree of perceived war zone stress but not direct combat (e.g. aggressive locals, being threatened with a weapon, seeing dead or injured people). The scale consists of 10 items with a possible sum score range from 10 to 40.

Single Items used in the present study. Single item questions were included to supplement the psychometric scales without increasing overall length of the questionnaire. *Earlier Problems* addresses participants previously had received psychological or psychiatric treatment for any personal problems (yes/no answer), *Emotional Stress* refers to how often emotionally stressful situations have been experienced during deployment, range from 1 (never) to 4 (very often), *Times with Life Danger* refers to how often a soldier felt that his or her life was in direct danger or how often a soldier witnessed others' lives being in direct danger, range from 0 (never) to 6 (more than 5 times), *Wounded/Injured* addresses whether the soldier reported being wounded or injured during deployment (yes/no answer), *Did you Kill?* addresses whether the soldier reported having killed an enemy during the deployment (yes/no answer).

Patterns of Missing Data

A critical concern is the degree to which those soldiers included in the analysis sample are representative of the full sample. In order to address this concern, we conducted a series of statistical comparisons between subsets of the full sample based on their pattern of missingness. Here, because there was less theory to guide us, and because we wanted to err, if at all, in the direction of including more variables than necessary in statistical adjustments, comparisons were made on a larger number of variables than in our hypothesis testing, 56 in total, that ranged from demographic characteristics to prior experiences of trauma to basic personality. The results of these analyses both address concerns about systematic missingness and provide a basis for choosing auxiliary variables to be included in the latent class growth models for the purpose of ensuring that missingness is ignorable.

We first compared means or frequencies for the 366 individuals in the analysis sample and the 380 individuals excluded due to a lack of sufficient data. Of the 56 comparisons, 12 were statistically significant. When the twelve were entered simultaneously in an equation predicting whether they were in or out of the analysis sample the total R^2 was .10 and only the gender and extraversion variables remained predictive. Thus, the analysis sample differed only slightly from the sample of individuals excluded from the analysis, and the nature of the difference was that the analysis sample was slightly more female and slightly less extraverted.

We next focused on differences within the analysis sample between individuals with complete data and those for whom one or more values would be imputed in the statistical modeling. Although not a definitive analysis, the results of these analyses suggest whether one or more of the missing patterns were systematic or nearly random. The most basic comparison is between the 125 individuals with complete data and the 241 missing one or two scores. This comparison across the 56 variables yielded only four significant effects, of which none were large (e.g., the strongest effect, at $p < .01$, was a difference of 2 scale points on conscientiousness, which could range from 12 to 60; individuals with complete data were slightly more conscientious that individuals missing one or two PCL scores). Importantly, there were no differences on PCL scores. Satisfied that the pattern of missingness was not highly systematic, we next focused on determining which variables might be included in the latent class growth models as auxiliary variables. We sorted the sample into the nine possible patterns of missingness and compared means or frequencies on the 56 variables. These comparisons yielded 15 small but significant differences. These variables were included as auxiliary variables in the latent class growth models, increasing the likelihood that any systematic missingness could be treated as ignorable.

Missing Data Estimation Strategy

We used state-of-the-science methods of managing missing data when estimating models. All models were estimated using full information maximum likelihood (FIML) as implemented in the M*plus* computer program. FIML produces unbiased parameter estimates when the missing data mechanism is missing at random (Rubin, 1976). Although the determination of whether the pattern of missingness is random cannot be determined with certainty, the likelihood is increased when potential reasons for missing data that are correlated with the outcome variables (PCL scores in our case) are accounted for in the model. As noted earlier, we examined a large number of variables and found small differences on a few. These variables were included as auxiliary variables in all estimated models. Auxiliary variables are those variables in the data set that either are correlated with missingness or correlated with variable on which there is missing data. To the extent these variables have been measured and are included in the model when FIML is applied, the missing data pattern can be made missing at random and therefore ignorable (Graham, 2003).

Latent Class Growth Modeling

The goal of our analyses was to model trajectories of PTSD symptoms from pre-deployment to six months after return from combat. We suspected significant variability in trajectories. Thus, our analytic strategy focused on estimating individual trajectories then probing for a set of distinct patterns that would allow for the classification of soldiers in terms of the trajectory of their PTSD symptoms across the time period.

We used latent class growth modeling, which includes elements of latent variable modeling and latent class analysis (Kreuter & Muthén, 2007). In terms of latent variables, variance in intercept (the first assessment in our models) and trajectory shape variables are modeled as latent influences on the pattern of PCL scores. We allowed for the possibility of nonlinearity by including both linear and quadratic latent variables in all models. In terms of latent classes, growth mixture modeling assumes significant and systematic heterogeneity in the latent growth variables that can be explained by an unmeasured class variable. A key concern in growth mixture modeling is determining the number of classes evident in the individual trajectories. Once the number of classes has been determined, individuals in the sample can be grouped according to latent class. These groups can then be characterized both in terms of the shape of their trajectory and in comparison to other groups on variables of interest.

We used full information maximum likelihood estimation for all models. All models included auxiliary variables as detailed earlier. We included linear and quadratic terms, and estimated models assuming from 1 to 10 classes. Fit indices associated with these models are shown in Table 1. There are no hard and fast rules for selecting an appropriate number of latent classes. As with, for example, exploratory factor analysis, selection involves a combination of factors, including statistical information, information about the data, and the substantive focus of the research. We consulted the fit information shown in Table S1 then, for models producing the best statistical fit, examine fitted growth curves and the distribution of respondents across class.

Simulation studies suggest that the sample-sized adjusted Bayesian information criterion (BIC_{SSA}) performs best in terms of model selection based purely on statistical information. As can be seen in Table 1, the values of BIC_{SSA} favor a model with nine classes. Examination of the fitted curves and distribution of respondents across classes for this model suggested against it, with some curves differing only slightly and membership in some classes very low. We examined more closely the results of models with fewer classes, focusing specifically on models with between five and eight classes. Ultimately, we selected a model with six classes, which balanced fit, interpretability, and distribution of respondents. The entropy value of this model was .91, suggesting good classification (i.e., probability near 1.0 for assignment to most probable class and near zero for assignment to other classes).[1]

The six fitted trajectories are displayed in Figure S1, overlaying the observed data. Class sizes ranged from N = 8 for the U-shaped trajectory (Strong Benefit Group) to N = 255 for the relatively straightly trajectory at the PCL minimum (Extremely Resilient Group). A class membership was created in the master data set, and the focal analyses involved comparing soldiers in the different trajectory classes on variables likely to differentiate them.

Supplementary regression analyses

A logistic regression analysis with New Onset versus Resilient group as the dependent measure, and years in the defense and amount of previous traumas as predictors showed a significant effect for amount of previous traumas (*Wald*(1)=7.22, B = .10, p<.01, odds ratio=1.10; [95% C.I.=1.03-1.18]) and years in the defense (*Wald*(1)=6.69, B = .21, p<.05, odds ratio=1.24; [95% C.I.=1.05-1.46]).We did not include earlier emotional problems as a predictor because of a high conceptual and empirical overlap between this categorical variable and the continuous measure of previous traumas and the few degrees of freedom in the analysis. If included both it and previous traumas were marginally significant (*p*s<.06).

A logistic regression analysis with Benefit versus Resilient group as the dependent measure and pre-deployment depression, neuroticism and previous traumas as predictor variables (but not pre-deployment PTSD symptoms and earlier emotional problems for reasons given above) showed significant

[1] Entropy values for the 7- and 8-class models were .92 and .91, respectively. In the 7-class model the group in the small U-shaped trajectory class in the 6-class solution split into two smaller classes that differed primarily with reference to the severity of the U shape. In the 8-class solution, another small class emerged that differed only in intercept from the class characterized by a slight, primarily linear decrease over time.

effects of all three variables: Neuroticism (*Wald*(1)=11.00, $B = .24$, $p<.005$, odds ratio=1.27; [95% C.I.=1.10-1.46]), depression (*Wald*(1)=9.02, $B = .19$, $p<.005$, odds ratio=1.21; [95% C.I.=1.07-1.36]) and previous traumas (*Wald*(1)=8.01, $B = .11$, $p<.01$, odds ratio=1.12; [95% C.I.=1.04-1.21]).

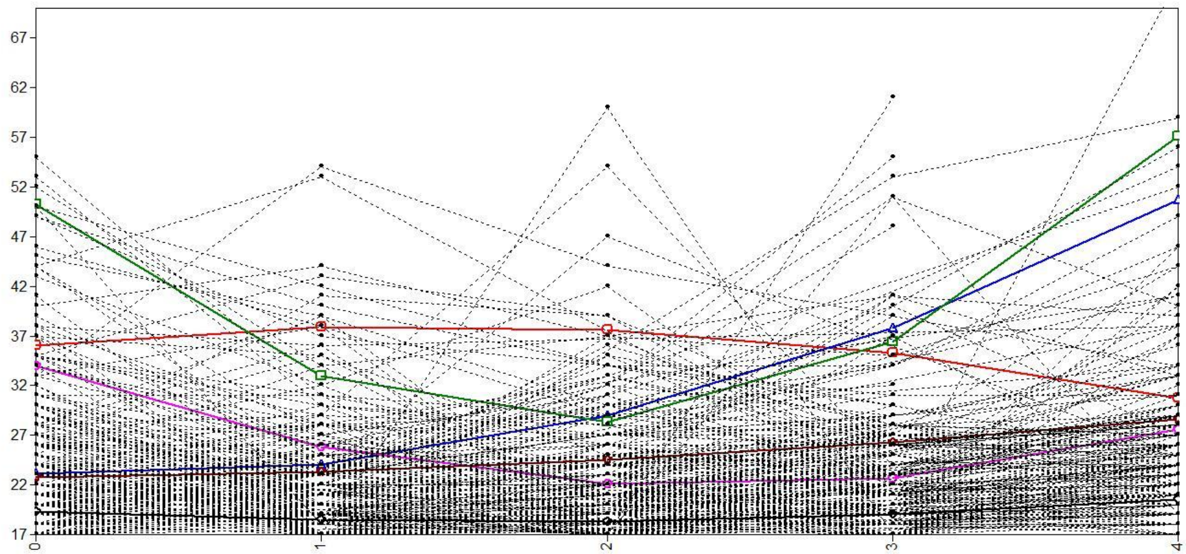

Figure S1. Raw PCL scores overlaid by fitted curves for the six latent classes.

References

Beck, A.T., Steer, R.A., & Brown, G.K. (1996). Beck Depression Inventory Manual (2nd ed.). San Antonio: The Psychological Corporation.

Blanchard, E. B., Jones-Alezander, J., Buckley, T. C., & Forneris, C. A. (1996). Psychometric properties of the PTSD checklist (PCL). *Behaviour Research and Therapy, 34,* 669-673.

Costa, P. T., & McCrae, R. R. (1989). The NEO-PI/NEO-FFI Manual supplement. Odessa, FL: Psychological Assessment Resources.

Graham, J.W. (2003). Adding missing-data-relevant variables to FIML-based structural equation models. *Structural Equation Modeling, 10,* 80-100.

Keane, T., Fairbank, J., Caddell, J., Zimering, R., Taylor, K., & Mora, C. (1989). *Clinical evaluation of a measure to assess combat exposure*. Psychological Assessment, *1,* 53-55.

Kreuter, F., & Muthén, B. (2007). Longitudinal modeling of population heterogeneity: Methodological challenges to the analysis of empirically derived criminal trajectory profiles. In G. R. Hancock &

K. M. Samuelsen (Eds.), *Advances in latent variable mixture models* (pp. 53-75). Charlotte, NC: Information Age Publishing.

Kubany, E. S., Haynes, S. N., Leisen, M. B., Owens, J. A., Kaplan, A. S., Watson, S. B., & Burns, K. (2000). Development and preliminary validation of a brief broad-spectrum measure of trauma exposure: The Traumatic Life Events Questionnaire. *Psychological Assessment, 12,* 210-224.

Rubin, D.B. (1976) Inference with missing data. *Biometrika*, *63*, 581-592.