Author's response to reviews

Title: Association Between Baseline Psychological Attributes and Mental Health Outcomes After Soldiers Returned from Deployment

Authors:
Yu-Chu Shen (yshen@nps.edu)
Jeremy Arkes (arkes@nps.edu)
Paul Lester (paul.b.lester.mil@mail.mil)

Version: 1 Date: 04 Aug 2017

Author’s response to reviews:
Responses to Referee Comments

Reviewer reports:
Laura Palmer (Reviewer 1): Corrections:
(p.2, line 24-31) sentence clarity

We revise the sentence as the following:

We analyze whether a soldier screened positive for depression and posttraumatic stress disorder (PTSD) after returning from deployment using logistic regressions. Our key independent variables are 14 psychological attributes, and we control for relevant demographic and service characteristics.

(p.2, line 41) family capitalised and other GAT categories are not revised to be consistent.

(p.2, line 17) can examples of psychological health attributes be provided here (to help those quickly scanning abstracts)?
Example added
(p.2, line 24) acronym needed for Global Assessment Tool acronym added

(p.10, line 14) percent written as a word, inconsistent with % used elsewhere fixed
(p.10, line 19) post deployment not hyphenated like elsewhere fixed
(p.10) could refer the reader to the Appendices for Figure 1 and Table 2

We would like to keep Figure 1 and Table 2 in the main text as they show important information about our results, but refer readers to the Appendices for complete regression results.

(p.10, line 58) hyphen used for confidence interval rather than comma used elsewhere fixed

(p.12, line 46) GAT referred to as psychological and social attributes in the Discussion (capturing the breadth of variables well), but is referred to as 'psychological' elsewhere. Would there be a benefit in referring to psychological and social attributes throughout the paper (or psycho-social)?

For cleanliness of presentation and better readability, we decide to keep the shorter term “psychological attributes” instead of expanding it to be “psychological and social attributes”.

(p.13 line 6) interpretation query re. "dramatic" - Perhaps the OR= 8.16 effect, but the other results seem moderate yet notable

We removed the word “dramatic” from the revision.

(p.15, line 24) "were" not was

Fixed

(p.16, line 24) "are" not is

Fixed

Recommendations:

This reading could be relevant to your work.

- Warner et al. (2011) Effectiveness of mental health screening and coordination of in-theater care prior to deployment to Iraq: a cohort study

- Brewin et al (2000) Meta-Analysis of Risk Factors for Posttraumatic Stress Disorder in Trauma-Exposed Adults. This paper shows that post-deployment factors of life stress and lack of social support are more predictive than pretrauma factors. This might partly explain the increase in PTSD and depression encountered after 90-180 days post-deployment

- Rona et al (2017) Post-deployment screening for mental disorders and tailored advice about help-seeking in the UK military: a cluster randomised controlled trial. This paper shows immediate post-deployment screening does not appear to predict mental health outcomes 10-24 months later
We thank the reviewer for suggesting these additional articles. We have incorporated them in the intro and discussion sections.

Literature review:

- The paper reports a scarcity of research linking antecedent psychological factors and mental health outcomes in military populations, but are there any similar studies you can include in the introduction? The introduction is light on prior research in the field.

We have substantially revised the introduction to include additional relevant studies (pg 3-4).

- From lines 14-36, p.4 - The context of psychological screening might be better placed before the more specific aims of your study. This paragraph could be merged with the section following line 7, p.5.

Per reviewer suggestion, we moved down and combined these paragraphs.

- P.5 however reads a little too methodologically for an introduction or lit. review - can these elements be summarized here and expanded upon in the methods section?

Thanks for the suggestion, we moved the paragraphs on GAT measurements to the Methods section (now on pg 7-8).

Methods:

- Do you have response rates?

Because all soldiers are required to take GAT within the first few weeks of joining the Army and annually hereafter, the response rate is virtually 100%.

The post deployment health assessment is also a required assessment that need to be completed within 90 days after a soldier returned from the deployment. Of the soldiers we observe having had a deployment, all but 9.5% had completed the PDHA.

- Can you clarify if you're excluding National Guards? It might be an interesting inclusion/comparison considering known differences in mental health outcomes and potential preparedness for combat

We exclude National Guards, because the GAT is only implemented within the US Army.

Results:
- I see you adjusted for counselling, but did you adjust for pre-deployment PTSD and depression? It might be worth commenting upon if yes or no. One consideration is whether the worst 5% on psychological attributes are also the positive cases of pre-deployment PTSD and depression. If they map upon one another, this would indicate that pre-deployment mental health outcomes are associated with post-deployment mental health outcomes (which is expected).

If a soldier is screened positive for PTSD or other major mental health problems during their pre-deployment health assessment, they would not be deemed as “fit for deployment” and would be excluded from being deployed (our other works showed that the odds of being deployed is indeed a lot lower for those diagnosed with mental health problems). Our analysis is conditional on only those who were deployed. This way, we have a more uniform sample of soldiers in terms of their mental and physical capacity (i.e., they were all deemed to be fit for deployment). What we exploit is the variation in their post deployment mental health outcomes by soldiers’ baseline psychological attributes—attributes that were measured in the beginning of their military career before they were deployed or fully exposed to the military environment.

Discussion:

- It would be beneficial to see whether your findings fit with other studies looking at psychosocial antecedents and mental health outcomes in military samples (if any). If none, you could make the original contribution of this work in finding these associations clearer.

We are only aware of one other study that directly estimate the relationship between antecedent psychological factors and post deployment mental health outcomes. We added the following sentence in the discussion section (pg 13):

Our results are consistent with a recent study that showed that soldiers who scored high on psychological strength measures prior to being deployed are less likely to develop mental health problems post deployment [15].

- There is a strong thread throughout that this could be a strategy for pre-selecting a suitable workforce. In light of the sometimes delayed onset of PTSD post-deployment (years after exposure - see Rona et al. above), you could suggest that future studies should use at a longer follow-up period in order to conclude that this would be a viable pre-selection method - otherwise it is quite excluding.

We incorporated Rona et al and reviewer suggestion in the discussion as follows on pg 15:

For example, a recent study found that screening conducted 6-12 weeks after deployment did not predict mental health problems that occurred 10-24 months after deployment [55], so it would be critical for future studies to capture mental health problems from a longer follow-up period in order to validate our results.
- Regarding your supposition on line 6, p.14, you could address the difficulties more explicitly, i.e. issues surrounding the reliability of the data if personnel know that they may not be deployed as a result of their pre-deployment screening answers.

We agree with the reviewer that this is an important point to address. We added a new paragraph as part of the discussion (bottom of pg 16) to highlight the important fact that GAT in its current form is not designed to be used as a screening tool. Rather, we use GAT data in the current study to illustrate the potential value for psychological health screening in public safety and national defense occupations. We also caution that any future design of any screening tool needs to detect and minimize faking, since by then the personnel know that their career progression and chance of being deployed might depend on their pre-deployment screening answers.

Chantelle Whelan (Reviewer 2):

This is an interesting paper that requires some revisions before being suitable for publication.

There are some issues with the structure of this paper. The GAT is described in detail in the introduction section, this would be better placed in the methods section. It would also be beneficial for the reader to put the study aims at the end of the introduction, instead of half way through the section.

Per reviewer suggestion, we revised the text by moving the GAT details to the methods section and the study aims at the end of the introduction.

The study aims also need to be made more explicit. At present, the sentence 'In this study, we explore one potential strategy to achieve this goal' comes after discussing the development of strategies to recruit physically and mentally fit individuals, which may reduce the costs of mental health problems. This study does not address all of these issues. Physical fitness and the financial implications of pre-deployment screening are not measured. The specific aims of this study should be outlined in order to adequately review the methods used in this study.

We revised the last paragraph of introduction to be about the study aim and clearly state that we analyze the association between 14 baseline attributes and U.S. Army soldiers’ probability of screening positive for two costly mental health illness—depression and PTSD —following their first combat deployment.

More detail is required in the methods section. There was no mention of justification for the sample size used. I would have found it helpful to have the GAT questions as an appendix to get a better understanding of how the 14 attributes were measured.

We did not provide additional justification for the sample size, because our analytical sample captures 100% of the US Army population who completed their first GAT anytime between October 2009 and March 2013 and who had a valid PDHA after their first GAT date.
We now included the GAT questions and how they are grouped into the 14 attributes as appendix Table 1.

I found the sentence 'We found similar results using different aggregations of the GAT questions' unclear. Perhaps having the questions would have helped in my understanding of this.

That sentence is removed from the current text. In the past, other researchers using GAT have used a broader categorization (4 dimensions [emotional fitness, social fitness, family fitness, spiritual fitness], instead of 14 subcategories) for other purposes. In this revision, we removed the sentence to avoid confusion since the broader dimensions are harder to comprehend and have weaker theoretical foundations for the purpose of this analysis.

The statistical methods section also requires further information. A description of the statistical software used to carry out the analysis should be reported, as well as details of the statistics reported in the results section (Odds ratios and Confidence Intervals). There is also no mention of the statistical significance level used.

We now added the statistical software used in on pg 9.

We added clarification in the results section whether the attributes were significant at the 1% or 5% significance level. We also revised the text to report all 95% CIs associated with the odds ratios estimates.

There is some demographic data missing from Table One - 'Males', 'Caucasian' and 'Single' are not included in the table.

We added those information in Table 1.

The discussion is currently quite limited, focusing predominantly on the financial implications of pre-deployment screening. A discussion of the 14 attributes and how they related to post deployment depression and PTSD could broaden the scope of this section.

We discuss how the 14 attributes were designed to be predictive of mental health outcomes within the context of military settings on pg 8 (first paragraph), when we first introduce these psychological attributes. We felt that it is a better place to provide such a discussion early on rather than to wait until the discussion section.

James Naifeh (Reviewer 3):

This study examined whether the baseline scores of new soldiers on the Global Assessment Tool (GAT) can predict postdeployment screening results for PTSD and depression. The findings indicate that soldiers with GAT scores in the bottom 5% may be at elevated risk for postdeployment mental health problems. The manuscript is well written and addresses the important issue of identifying soldiers who may be vulnerable to significant distress. Longitudinal analysis of soldiers who have just entered the Army is critical for understanding
adverse mental health outcomes in this population. Below I have suggested several areas for improvement. Of greatest concern are the seemingly contradictory statements about whether the GAT should be used as a screening tool to exclude high-risk soldiers, as this has important policy implications.

1) Page 2, Line 43: To avoid confusion, the Abstract results for postdeployment PTSD should specify that it is associated with "baseline depression," rather than just "depression."

Fixed

2) Page 3, Lines 50-58: The manuscript states "One important hypothesis is that while the increasing trend in mental health service needs may be influenced by the stressors involved with the protracted war on terrorism and military life in general, the upward trend in mental health service needs may also be reflective of poor psychological health of soldiers entering the military." Although pre-military mental health is likely associated with risk for post-deployment outcomes, it does not explain upward trends in psychiatric morbidity over time unless the proportion of new soldiers with pre-military mental health problems has increased. Perhaps the authors can clarify the point they are trying to make.

We agree this statement is confusing and has revised it as follows on page 4:

One important hypothesis that has not been fully explored in the literature is that some soldiers might enter the military with poor psychological health such that they have much higher demand for mental health service when they are exposed to the stressors involved with the protracted war on terrorism, relative to others who are psychologically fit for the military life.

3) There appear to be conflicting statements about potential uses of the GAT. On one hand, the manuscript states "Knowing how well these psychological attributes can predict future mental health outcomes can potentially aid the DoD in identifying a workforce that is better suited for the stresses associated with its unique environment..." (Page 4, Lines 14-29). But then a couple paragraphs later the manuscript states "The GAT is not designed to be used as a screening tool and to do so in high stakes settings where employment decisions are made would be a mistake" (Page 5, Line 9-14). Then in the Discussion, the manuscript states that a composite risk score could be "useful in screening recruits" (Page 12, Line 6-9). The authors should clarify their stance on how the GAT should be used. Is it a screening instrument, and if so, what is the purpose of the screening (e.g., to intervene with high risk recruits vs. exclude them from service)?

We now devoted a whole paragraph (copy below, bottom of page 16) in the discussion section to carefully clarify our stance on the GAT as a potential screening tool. We believe this paragraph is consistent with the objective that we lay out in the revised introduction.
Lastly, it is important to recognize that the GAT in its current form is not designed to be used as a screening tool and to do so in high stakes settings where employment decisions are made would be a mistake. Rather, we use GAT data in the current study to illustrate the potential value for psychological health screening in public safety and national defense occupations. When taken together, the data gathered from the GAT offers us a unique opportunity to quantify the psychological health and resilience of soldiers prior to full immersion into the military and deployment to combat zones. It would be important for any future design of any screening tool to detect and minimize faking, since by then the personnel know that their career progression and chance of being deployed might depend on their pre-deployment screening answers.

4) Page 7, Lines 7-9: The manuscript states, "In an alternate analysis exercise, we included all soldiers." The manuscript should explain why this alternate analysis was conducted. This is revealed in the Results section but it would be helpful to explain the rationale up front.

We now expanded that paragraph on pg 6 to include the following rationale:

Comparing results from both the restricted sample and the whole sample allow us to investigate whether the relationship between these baseline psychological attributes and post-deployment health conditions differ whether or not a person experienced military life before taking the GAT.

5) Page 7, Line 48: It should be specified that that these PC-PTSD questions are based on the DSM-IV version of PTSD. Additionally, DSM-IV PTSD officially has only 3 symptom clusters, not 4 clusters as reported in the manuscript -- the avoidance and detachment questions in the PC-PTSD are assessing two aspects of the same cluster (Criterion B). Although confirmatory factor analytic studies support 4-factor models of DSM-IV PTSD, the official DSM-IV version is still 3 clusters. DSM-5 PTSD does have 4 clusters, but this is assessed using a revised version of the PC-PTSD that has 5 questions (https://www.ptsd.va.gov/professional/assessment/screens/pc-ptsd.asp).

We thank the reviewer for providing these clinical clarifications. We revised the text on pg 7 to reflect that the questions were based on DSM-IV version and represents 3 symptom clusters.

6) Page 8, Lines 6-12: The manuscript appears to be saying that the GAT has both binary and Likert-type responses. If some of the GAT responses are binary, how were they standardized to a scale of 1-5? Perhaps I have misunderstood the description of the GAT?

We now clarify on pg 7 that for binary responses, we converted the no and yes responses to 1 and 5 point, respectively.
7) Page 8, Line 48: The manuscript should clarify whether time-varying demographics (e.g., age, rank) are based on their value at the time of the GAT assessment or the post-deployment assessment.

We clarify on pg 9 that “All time-varying variables (such as age, rank) are based on their value at the time of the post-deployment assessment.”

8) Did the multivariate analyses of each GAT scale also adjust for the other GAT scales, or was each scale examined separately?

The multivariate analysis includes all 14 attributes at once. We also estimated separate analysis where we include each scale separately (i.e., 14 separate regressions for each outcome). Not surprisingly, the odds ratio is much higher for all attributes when we do such bivariate model, since the attributes are highly correlated with each other.

9) Page 11, Lines 36-44: Were ROC curves calculated for the full sample or the restricted sample? The results are described following the results for the full sample, so I assume that is the sample that was used. Is there a reason these analyses were not conducted on the restricted sample, which seems to be the primary analytic sample in the study?

For completeness, we now show the ROC curves for both samples in Figure 2. They are very similar to each other.

10) Page 11, Lines 46-58: Similar to the previous comment, it is not clear if the interactions were tested in using the restricted sample, the full sample, or both. It would be helpful if the manuscript clearly specified which sample is being used for the ROC and interaction analyses. If only the full sample is being used then the manuscript should explain why the focus was switched to the full sample.

Please see our detailed response below for the reason we remove the interaction paragraph from this revised version.

11) Page 11, Lines 50-52: The manuscripts states that "in general" the interactions between GAT variables and combat exposure/MOS were not significant, with the results available upon request. The language is a little vague. Does this mean that some of interactions were significant? If any of the interactions were significant, the authors might consider reporting them to enhance understanding of how GAT scores influence post deployment risk. In addition, the conclusion that "the relationship between the baseline psychological attributes and post-deployment health conditions do not appear to be modified by the stress level of the work environment" may be too broad given the limited assessment of work-related stressors in this study. The 3 deployment stressors are on the more severe end, capturing direct exposure to combat and death (only 18-31% of the sample had experienced these stressors). The authors
might consider using more specific language to describe these stressors rather than suggesting they capture work-related stress, which has many dimensions in addition to combat and death/injury.

After careful consideration, we remove this paragraph altogether in the revision for the following reason. In an earlier draft of the paper, we used a much broader aggregation of GAT responses (lumping the 14 attributes into 4 dimensions) that have much weaker theoretical foundation and harder to explain. In that version, we included 12 interaction terms (4 dimensions x 3 combat exposure variables) and the paragraph in pg 11 referred to this much earlier model. However, in our current version we have 14 attributes and 3 combat exposure variables, which means we will need to include 42 interaction terms—such model is impossible to interpret and distract from the main take-away points of our paper. Given that we now discuss the PDHRA results in the methods and results section (see our response to #14 below), and the rest of the discussion and conclusion are derived from our main models and risk profile results, we believe the paper is better without this paragraph.

12) Page 12, Lines 4-34: Examining concentration of risk was a particularly useful step. I'm curious how the model would have performed if only the GAT was included. In other words, how useful is the GAT, on its own, in identifying concentration of risk? This is important if the GAT is going to be used as an early indicator of soldiers who may have subsequent problems.

In our concentration of risk calculation, we purposely excluded all service characteristics that cannot be observed by the recruiter. But we do think the model should include demographics and AFQT scores that were available and collected by the recruiters along with the GAT scores, since all those information can be used in creating a composite risk profile by the recruiter.

We experimented with computing the concentration of risk using only GAT information and included the graph at the end of the referee report. The graph shows that the concentration at the top (which is what we cared about for this exercise) is actually identical to those shown in Figure 3: 31% of those who screened positive for depression and 27% of those who screened positive for PTSD were concentrated among the top 5 percent high risk population when the model included only the GAT attributes. Another thing to note is that ¾ of the soldiers are in the bottom 1 or 2 ventiles, because of the model set-up (i.e., identify soldiers if they are in the bottom 5 percentile of each of the GAT attributes). In other words, variations in the outcomes among the top 5 ventiles in our Figure 3 are really driven by GAT attributes, while the variations in outcomes among the lower ventiles are driven by other demographic and AFQT information that can be observed by the recruiter. If the reviewer feels strongly that Figure R at the end of this response document should replace Figure 3, we are happy to do so.

13) Page 13, Lines 9-14: This is a very strong conclusion given the limitations of the current study. The brief screening scales in the PDHA (2 depression items, 4 PTSD items) provide a limited assessment of postdeployment mental health. It is impossible to know how soldiers who screened positive would have been classified based on a more thorough assessment that incudes all symptoms and, importantly, an assessment of clinical distress and/or functional impairment.
Further, this again raises the issue of how the GAT should be used. In the Introduction, the manuscript states that the GAT should not be used for employment screening, and yet that seems to the implication of describing low-scoring soldiers as a “poor fit for military service.”

We have toned down this statement on pg 13, and offered as a speculation that “Those soldiers who scored worst might be more susceptible to developing debilitating mental health disorders when they are later exposed to combat environments.”

14) Page 15, Lines 7-19: The PDHRA analyses are a nice addition, but they should be included, and more thoroughly described in the Method and Results section. The Discussion is not the place to introduce new major analyses.

We now include the description of the exploratory analysis using PDHRA as part of methods (pg 9, last paragraph before Results section) and results section (bottom of pg 11).

15) Page 15, Line 58: A citation should be included for the TAPAS.
Citation added

16) Please clarify if the tables and figures are using the full or restricted sample.

All table and figure titles are revised to indicate the restricted or full sample.

Figure R. Replicating Figure 3 but only include GAT attributes in the model