



# Reporting Psychological Symptoms: Anonymity Matters

US Army Medical Research Unit - Europe  
Walter Reed Army Institute of Research

CPT Jeffrey Thomas (jeffrey.l.thomas@us.army.mil)  
Dr. Kathleen Wright (kathleen.wright@us.army.mil)  
Dr. Amy Adler (amy.adler@us.army.mil)  
LTC Paul Bliese (paul.bliese@us.army.mil)

---

The objective of psychological screening is to identify Soldiers who are experiencing psychological distress and link them with military mental health assets. One concern for the validity of psychological screening is the degree to which endorsing psychological distress may be perceived by Soldiers as stigmatizing and, thus, result in under-reporting on non-anonymous surveys. To address this concern, the US Army Medical Research Unit-Europe tested for significant differences between Soldiers' responses to an anonymous mental health survey and to a non-anonymous survey. Using a liberal alpha rate, significant differences were observed between the two survey types on four clinical symptom areas (anger, depression, alcohol and traumatic stress) with those in the anonymous group reporting more problems than those in the non-anonymous group.

---

## • Psychological Screening Background

Psychological screening has been a research focus of the U.S. Army Medical Research Unit-Europe (USAMRU-E) in Heidelberg since 1996 (see Wright, Huffman, Adler, & Castro, 2002, for a review). During that time, research has examined screening results across a range of operations (Adler, Wright, Huffman, Thomas, & Castro, 2002; Martinez, Huffman, Adler, & Castro, 2000). Subsequent studies have developed the groundwork for validating the primary screening instrument (Wright, Thomas, Adler, Ness, Hoge, & Castro, in press).

Building on recent screening research (Wright, et al., in press), five content areas have been identified as targets for screening: (1) traumatic stress, (2) depression, (3) relationship problems, (4) alcohol problems, and (5) anger problems.

In order to develop a valid screen that encompasses these five content areas, several steps in the validation process need to be accomplished. One of these steps is

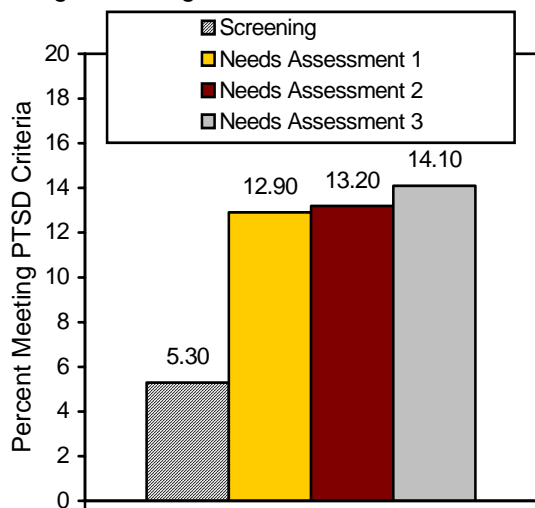
to assess the degree to which Soldiers under-report psychological symptoms. Previous research has found that Soldiers perceive psychological and physical screening as potentially stigmatizing (Britt, 1998), and this perception may result in under-reporting of symptoms. The degree to which under-reporting occurs raises questions about how to optimize the screening process, as well as questions about estimates of population rates of symptoms based on screening procedures.

## • Land Combat Study Needs Assessment Research

Researchers at the Division of Neurosciences, Walter Reed Army Institute of Research (WRAIR) have studied rates of psychological distress among Soldiers redeployed from Operation Iraqi Freedom as a part of a comprehensive needs assessment study (see Hoge, Castro, Messer, McGurk, Cotting, & Koffman, 2004). Needs assessment research has shown substantially higher rates of psychological distress among redeploying units than rates observed in screening settings with units

at post-deployment. For instance, Figure 1 shows rates of Post Traumatic Stress Disorder (PTSD) symptoms observed in a screening setting versus rates observed in a needs assessment setting. Notice in Figure 1 that PTSD symptom rates in the screening sample are dramatically lower than rates in the needs assessment samples.

Figure 1. Post-traumatic Stress Disorder (PTSD) Rates Among Screening and Needs Assessment Units



One key difference between screening studies and needs assessment studies is that needs assessment research uses anonymous surveys, while the screening research does not. Indeed, anonymity is not possible in screening where the purpose is to link individual Soldiers with mental health support. Nonetheless, the dramatic difference in psychological distress rates between the two lines of research needs to be explained. Therefore, in the present study we randomly assigned soldiers into one of two groups: Soldiers who completed an anonymous mental health survey and Soldiers who completed a non-anonymous mental health survey. The purpose of this research was to determine whether differences like those in Figure 1 can be attributed to anonymous versus non-anonymous surveys procedures.

• **Current Study: Sample and Procedure**

The current study examined the impact of anonymity on symptom reporting. The study is based upon responses from Soldiers collected 120-days following return from combat in Iraq. Soldiers were surveyed as part of an extended psychological screening and surveillance

procedure requested by senior leadership. In all, 1,556 Soldiers completed one of the versions of the mental health survey.

The procedure for psychological screening and surveillance at 120-days post-deployment consisted of a mental health surveillance survey (anonymous or non-anonymous). Survey administration was conducted as part of a medical readiness event on a Company by Company basis in blocks of approximately 100 Soldiers at a time. An envelope containing a consent form and either the anonymous or non-anonymous surveillance survey was placed on each chair. Anonymous and non-anonymous surveys were distributed in an alternating fashion in order to avoid confounding anonymity with Company membership. This is important because research consistently finds evidence that psychological symptoms cluster by unit (Bliese & Jex, 2002).

When a Company of Soldiers arrived they were seated and received an orientation PowerPoint brief. Soldiers were briefed about the two versions of the surveillance survey and informed it would be used only for research purposes. Soldiers were briefed that their fellow Soldiers to their left and right would have a different version of the survey.

An analysis comparing the two survey groups on demographics (e.g., rank, marital status, educational background, military experience, and deployment experience) revealed that there were no significant differences between the two survey groups as a result of the randomization procedure.

Soldiers were informed that non-anonymous surveys allowed for linking Soldier responses with data collected immediately after their return from Iraq. They were also informed that if they received an anonymous survey they should not write their name or Social Security Number on the form. The initial brief took approximately 10 minutes. After the brief was completed, Soldiers were instructed to begin the mental health survey. The mental health survey took approximately 20 minutes for Soldiers to complete. Once completed, Soldiers placed their survey back in the envelope and returned it to study personnel. At this point, Soldiers received a non-anonymous short form survey used for generating

referrals to mental health services. The short form took approximately 5 minutes for Soldiers to complete. Once they completed the short form survey, they were finished with the psychological screening portion of the medical readiness event and went to the next medical readiness station.

• **Setting a More Liberal Alpha to Test for Differences**

Our interest lies in examining the equivalence of responses between the anonymous and the non-anonymous groups' reports of symptoms. As such, we wanted to increase the probability that differences between the two conditions were detected. Statistically, we increased our probability of detecting differences by relaxing the alpha rate from the traditional value of 0.05 to a value of 0.20. That is, we accepted a 20% risk that any differences we detected were by chance alone rather a more typical 5% risk. Furthermore, we expected Soldiers completing the anonymous surveys would report more psychological distress than Soldiers completing the non-anonymous surveys, thus we adopted a one-tailed test. If no significant differences were found using the relaxed alpha rate of 0.20 (one-tailed), then it strengthens the argument that the two groups were not significantly different in their reporting of psychological distress.

• **Results**

On each psychological distress dimension, contrasts between the anonymous and non-anonymous surveys were made in two ways. First, to maximize the sensitivity of the contrasts, we computed t-tests on the summed scale scores. Second, because each scale is coded dichotomously as positive or negative in the screening process, we conducted Chi-square analyses of the dichotomously coded outcomes in all cases except for the alcohol dimension, which is under development and thus has no clear dichotomous cut-off value.

Traumatic Stress. The items used to assess traumatic stress were drawn from the 17-item Post-traumatic Stress Disorder Checklist (PCL, Weathers, Litz, Herman, Huska, & Keane, 1993). The PCL assesses three symptom clusters: re-experiencing, avoidance and hyperarousal. Cut-off scores based on whether the total

scores exceed 50, 44 and 30 have been recommended in the literature for determining whether or not individuals show symptoms of PTSD (Blanchard, Jones-Alexander, Buckley & Forneris, 1996; Weathers, et al. 1993).

The average summed score on the PCL for the anonymous group was 26.44. For the non-anonymous group, the score was 25.11. Table 1a shows that the summed score on the PCL for the anonymous group was significantly higher than the value for the non-anonymous group. In addition to these overall findings, the PCL values were higher for the anonymous group than the non-anonymous group on each dimension of the PCL. The average summed score for the dimension of re-experiencing was 7.30 for the anonymous group and 6.90 for the non-anonymous group; likewise, the averaged summed score for the dimension of avoidance was 10.23 for the anonymous group and 9.78 for the non-anonymous group. Finally, for the dimension of hyper-arousal, the values were 8.93 and 8.46 for the anonymous and non-anonymous, respectively. Table 1a reveals that all of these differences were statistically significant.

*Table 1a. Differences between Anonymous and Non-Anonymous Groups on Summed PTSD symptoms*

Scale	Σ Diff.	t	d.f.	p
PTSD-Overall	1.33	-2.28	1531	.02
PTSD-Sub-scales:				
<i>Re-experiencing</i>	.40	-2.24	1516	.02
<i>Avoidance</i>	.45	-1.85	1522	.03
<i>Hyper-arousal</i>	.47	-2.14	1523	.02

Tables 1b, 1c, and 1d show the comparisons using the PCL cut-off values of 50, 44 and 30. Notice that when the value of 50 is used as the cut-off, there is no difference between the anonymous and non-anonymous groups. In contrast, a cut-off of 44 and 30 for the PCL shows that there are significant differences between the groups.

*Table 1b. Association between Survey Type and Scoring Negative or Positive Using PCL 50 (N = 1533)*

	Negative on Relationship Problems	Positive on Relationship Problems
Non-Anonymous	737	40
Anonymous	709	47

Pearson's Chi-Square = .82, Fisher's Exact 1-tailed  $p = .21$ , Non-significant

*Table 1c. Association between Survey Type and Scoring Negative or Positive Using PCL 44 (N = 1533)*

	Negative on Relationship Problems	Positive on Relationship Problems
Non-Anonymous	709	68
Anonymous	678	78

Pearson's Chi-Square = 1.09, Fisher's Exact 1-tailed  $p = .17$ , Significant

*Table 1d. Association between Survey Type and Scoring Negative or Positive Using PCL 30 (N = 1533)*

	Negative on Relationship Problems	Positive on Relationship Problems
Non-Anonymous	601	176
Anonymous	540	216

Pearson's Chi-Square = 7.06, Fisher's Exact 1-tailed  $p = .01$ , Significant

The results suggest that at least in this sample the overall score of 50 is relatively insensitive to the effects of anonymity versus non-anonymity and an overall score of either 44 or 30 is much more sensitive to the effects of anonymity versus non-anonymity. Nonetheless, it is clear that anonymity is related to Soldiers' responses on the PCL such that Soldiers with anonymous surveys are more likely to report symptoms across each dimension of the PCL.

#### • Relationship Problems

The six items used to assess relationship problems were adapted from Norton's marital satisfaction measure

(1983). Responses were scored on a scale from 1 to 5 where 1 = Strongly Disagree and 5 = Strongly Agree. Based on previous screening research, Soldiers who disagreed or strongly disagreed on three of the five marital satisfaction items were regarded as having exceeded criteria. This cut-off score has been validated in previous screening research (Ployhart, 2004).

Responses were analyzed only if Soldiers reported they were married or in a significant relationship ( $n = 700$ ). The average marital satisfaction score for Soldiers with anonymous surveys was 25.08, and the average marital satisfaction score for Soldiers with non-anonymous surveys was 24.46. Table 2a shows that this difference was statistically significant at  $p = .11$ , using a one-tailed t-test with an alpha of .20. The results, however, are in the opposite direction of what was hypothesized because Soldiers in the anonymous condition reported lower levels of relationship problems than those in the non-anonymous condition.

*Table 2a. Differences between Anonymous and Non-Anonymous Groups on Summed Relationship Problems*

Scale	$\Sigma$ Diff.	t	d.f.	p
Relationship Problems	.62	-1.21	699	.11

*Table 2b. Association between Survey Type and Scoring Negative or Positive for Relationship Problems (N = 707)*

	Negative on Relationship Problems	Positive on Relationship Problems
Non-Anonymous	316	43
Anonymous	315	33

Pearson's Chi-Square = 1.15, Fisher's Exact 1-tailed  $p = .17$ , Significant

The un-anticipated direction of the finding was also reflected in the analyses based on cut-off values. Table 2b shows that 12% of the Soldiers (43 of 359) in the non-anonymous group met the criteria for relationship problems while only 7% of the soldiers in the anonymous group met the criteria. Thus anonymity may be associated with reporting higher marital satisfaction, but any definitive statements regarding this finding need to

be replicated given the weak and unexpected directional nature of the results.

### • Anger Problems

The third dimension assessed was anger problems. The five items used were modified from anger items published in the open literature. Soldiers rated the items on a five-point scale ranging from 0 = Not At All to 4 = Very Often. In order to determine the cut-off value for anger problems, there were two considerations. First, Soldiers responding “often” or “very often” to two or more of the five items were regarded as having exceeded criteria for anger. Second, soldiers responding often or very often to the item, “sometimes felt that you could not control your urge to harm others such as a unit member or friend” were also regarded as having exceeded criteria.

The specific items were:

- 1) Felt angry and irritated for no reason at all.
- 2) Had flare ups or temper outbursts.
- 3) Was on the verge of losing control of your anger.
- 4) Was verbally abusive to someone.
- 5) Sometimes felt that you could not control your urge to harm others such as a unit member or friend.

The average summed score on the anger scale for Soldiers completing the anonymous survey was 8.38, while for Soldiers completing the non-anonymous survey the average score was 7.97. Table 3a and 3b show these differences were significant both in terms of summed scores and in terms of exceeding criteria. Specifically, Soldiers who completed the anonymous survey reported more anger problems than Soldiers who completed the non-anonymous survey.

*Table 3a. Differences between Anonymous and Non-Anonymous Groups on Summed Anger Problems*

Scale	Σ Diff.	t	d.f.	p
Anger Problems	.41	-2.00	1554	.02

*Table 3b. Association between Survey Type and Scoring Negative or Positive for Anger Problems (N = 1,556)*

	Negative on Anger Problems	Positive on Anger Problems
Non-Anonymous	728	61
Anonymous	692	75

Pearson’s Chi-Square = 2.04, Fisher’s Exact 1-tailed  $p = .09$ , Significant

### • Depression

Depression was assessed using the depression subscale from the Patient Health Questionnaire (PHQ; Spitzer, Kroenke, & Williams, 1999). The PHQ-Depression subscale is a 9-item scale on which respondents report on a 4-point scale (0= not all, 1 = few or several days, 2 = more than half the days, or 3 = nearly every day) how often they experience symptoms (e.g., feeling tired or having little energy). The cut-off score followed the algorithm recommended by Spitzer et al. (1999).

The average summed score on the depression items for the Soldiers in the anonymous group was 12.49, while the score for the Soldiers in the non-anonymous group was 12.25. Table 4a reveals that this difference in the summed score was statistically significant; Soldiers who completed the anonymous survey reported more symptoms of depression than did Soldiers who completed the non-anonymous survey.

Interestingly, when the PHQ-Depression items were coded according to whether or not the Soldier met the criteria for depression, the differences between the two groups were no longer evident. Notice in Table 4b that the rates of those exceeding criteria for depression in the anonymous group were not significantly different than the rates in the non-anonymous group.

*Table 4a. Differences between Anonymous and Non-Anonymous Groups on Summed PHQ Depression Scale*

Scale	Σ Diff.	t	d.f.	p
Depression	.24	-.96	1534	.17

*Table 4b. Association between Survey Type and Scoring Negative or Positive for PHQ Depression (No Functional Impairment) (N = 1,556)*

	Negative on PHQ Depression	Positive on PHQ Depression
Non-Anonymous	743	46
Anonymous	726	41

Pearson's Chi-Square = .17, Fisher's Exact 1-tailed  $p = .38$ , Non-significant

### • Alcohol Problems

In order to assess alcohol problems, we relied on a composite six- item alcohol scale developed by the US Army Medical Research Unit-Europe in collaboration with the WRAIR Division of Neurosciences. The first two items were taken from Brown, Leonard, Saunders, & Papasouliotis (2001). The remaining four items were taken from items used by WRAIR Division of Neurosciences. These items are behaviorally-anchored indicators of drinking problems. Soldiers were asking to provide responses concerning their drinking in the last four weeks. The items were scored using a Yes/No response format. The specific items were:

- 1) Have you used alcohol more than you meant to?
- 2) Have you felt that you needed or wanted to cut down on your drinking?
- 3) Has your drinking resulted in your blacking out?
- 4) Have you driven when drunk?
- 5) Has your drinking resulted in serious injury to you or someone else?
- 6) Has your drinking resulted in repeated, intense conflict with others?

No cut-off values were used for the alcohol problems scale because the measure is still under development. However, for each item, we compare anonymous versus non-anonymous responses.

Overall, there were no differences between the groups on the endorsement of alcohol problems. The summed responses on the items for the anonymous group were 11.55; for the non-anonymous group, the summed responses were 11.55 (see Table 5a).

Tables 5b-5g show the association between survey group and endorsing each of the alcohol problem items (Yes/No). There were no significant differences from the categorical analyses of survey type and alcohol problems with the exception of endorsing the items, "driving while drunk" and "wanting/needing to cut down". More Soldiers in the anonymous survey group reported driving while drunk and wanting to cut down than Soldiers in the non-anonymous survey group.

*Table 5a. Differences between Anonymous and Non-Anonymous Groups on Summed Alcohol Problems*

Scale	Σ Diff.	t	d.f.	p
Alcohol Problems	.00	-0.26	1545	.49

*Table 5b. Association between Survey Type and Indicating Yes or No for Using Alcohol More Than You Meant To (N = 1,555)*

	Use Alcohol More Than Meant To	Do Not Use Alcohol More Than Meant To
Non-Anonymous	119	670
Anonymous	119	647

Pearson's Chi-Square = .06, Fisher's Exact 1-tailed  $p = .43$ , Non-significant

*Table 5c. Association between Survey Type and Indicating Yes or No for Wanting/Needing to Cut Down on Drinking (N = 1,554)*

	Want or Need to Cut Down on Drinking	Do Not Want or Need to Cut Down on Drinking
Non-Anonymous	89	700
Anonymous	75	690

Pearson's Chi-Square = .89, Fisher's Exact 1-tailed  $p = .19$ , Significant

*Table 5d. Association between Survey Type and Indicating Yes or No for Reporting That Drinking Has or Has Not Resulted in Blacking Out (N = 1, 550)*

	Drinking Has Resulted in Blacking Out	Drinking Has Not Resulted in Blacking Out
Non-Anonymous	77	712
Anonymous	67	694

Pearson's Chi-Square = .42, Fisher's Exact 1-tailed  $p = .29$ , Non-significant

*Table 5e. Association between Survey Type and Indicating Yes or No for Driving While Drunk (N = 1,552)*

	Has Driven While Drunk	Has Not Driven While Drunk
Non-Anonymous	37	751
Anonymous	49	715

Pearson's Chi-Square = 2.18, Fisher's Exact 1-tailed  $p = .09$ , Significant

*Table 5f. Association between Survey Type and Indicating Yes or No for Drinking Resulting in Injury to Self or Other (N = 1,553)*

	Drinking Has Resulted in Injury to Self/Other	Drinking Has Not Resulted in Injury to Self/Other
Non-Anonymous	10	779
Anonymous	9	755

Pearson's Chi-Square = .26, Fisher's Exact 1-tailed  $p = .53$ , Non-significant

*Table 5g. Association between Survey Type and Indicating Yes or No for Drinking Resulting in Conflict with Others (N = 1,551)*

	Drinking Has Resulted in Conflict with Others	Drinking Has Not Resulted in Conflict with Others
Non-Anonymous	22	767
Anonymous	26	736

Pearson's Chi-Square = .50, Fisher's Exact 1-tailed  $p = .29$ , Non-significant

## • Discussion

The purpose of these analyses was to determine if there were differences in Soldiers' reports of psychological symptoms as a function of whether or not they completed an anonymous or non-anonymous survey. In essence, it was a validity check for whether or not data from psychological screening could be used to estimate population prevalence rates. The research was driven by the observation that reported symptom rates were much higher in anonymous needs assessment research than in screening research (recall Figure 1).

Overall, results generally supported the hypothesis that anonymity is associated with higher reports of psychological symptoms. In particular, Soldiers completing anonymous surveys reported higher PTSD rates, anger rates, depression rates, and higher rates on some aspects of alcohol use. Interestingly, however, Soldiers completing the anonymous surveys also reported higher rates of marital/relationship satisfaction than soldiers completing the non-anonymous surveys. This latter finding is inconsistent with the notion that anonymity is related to higher levels of reported symptoms.

It is important, however, to keep in mind that the magnitude of effects observed was relatively weak as evidenced by the p-values and absolute group differences. Furthermore, differences were more likely to be detected in analyses of summed scores than in analyses based on cut-off values. Based on these results, we would expect Soldiers completing anonymous surveys to report higher symptom rates than Soldiers completing non-anonymous surveys. The magnitudes of the effects, though, explain only a small portion of the large differences in symptom rates we initially observed between screening and needs assessment studies. Nonetheless, it is clear that it will generally be problematic to compare results from anonymous surveys to results from non-anonymous surveys even if assurances are given that data will be kept confidential in screening settings. In conclusion, symptom rate estimates from non-anonymous psychological screening will almost certainly underestimate true population prevalence rates.

## • References

- Adler, A.B., Wright, K.M., Huffman, A.H., Thomas, J.L. & Castro, C.A. (2002). Deployment cycle effects on the psychological screening of Soldiers. *U.S. Army Medical Department Journal*, 4/5/6, pp. 31-37.
- Blanchard, E. B., Jones-Alexander, J, Buckley, T. C. & Forneris, C. A. (1996). Psychometric properties of the PTSD Checklist (PCL). *Behavioral Research and Therapy*, 8, 669-673.
- Bliese, P. D. & Jex, S. M. (2002). Incorporating a multilevel perspective into occupational stress research: Theoretical, methodological, and practical implications. *Journal of Occupational Health Psychology*, 7, 265-276.
- Britt, T. W. (2000). The stigma of psychological problems in a work environment: Evidence from the screening of service members returning from Bosnia. *Journal of Applied Social Psychology*, 30, 1599-1618.
- Brown, R.L., Leonard, T., Saunders, L.A., & Papanicolaou, O. (2001). A two-item conjoint screen for alcohol and other drug problems. *Journal of American Board of Family Practice*, 14, 95-106.
- Hoge, C. W., Castro, C. A., Messer, S. C., McGurk, D., Cotting, D., & Koffman, R. L., (2004). Combat duty in Iraq and Afghanistan, mental health problems, and barriers to care. *New England Journal of Medicine*, 351, 1-10.
- Martinez, J.A., Huffman, A.H., Adler, A.B., & Castro, C.A. (2000). Assessing psychological readiness in U.S. Soldiers following NATO operations. *International Review of the Armed Forces Medical Services*, 73, 139-142.
- Norton, R. (1983). Measuring marital quality: A critical look at the dependent variable. *Journal of Marriage and the Family*, 45, 141-151.
- Ployhart, R.E. (2004). *Examination of the pre-deployment rotation 4B psychological screening survey. Final report: Review and recommendations* (U.S. Army Medical Research Unit-Europe Research Report 2004-003). Heidelberg, Germany: USAMRU-E.
- Spitzer, R.I., Kroenke, K., & Williams, J.B. (1999). Validation and utility of a self-report version of PRIME-MD: The PHQ primary care study. *Journal of American Medical Association*, 282, 1737-1734.
- Weathers, F.W., Litz, B.T., Herman, D.S., Huska, J.A., & Keane, T.M. (1993). The PTSD Checklist (PCL): *Reliability, validity, and diagnostic utility*. Paper presented at the annual meeting of the International Society for Traumatic Stress Studies, San Antonio.
- Wright, K.M., Huffman, A.H., Adler, A.B., & Castro, C.A. (2002, October). *Psychological screening program overview*. *Military Medicine*, 167, 853-861.
- Wright, K.M., Thomas, J.L., Adler, A.B., Ness, J.W., Hoge, C.W., & Castro, C.A. (in press). *Psychological screening procedures for deploying U.S. Forces*. Manuscript accepted for publication in *Military Medicine*.