Summary

As noted in the 2006 Quadrennial Defense Review, international partnerships are increasingly important to the security and welfare of the United States and its allies. Without strong international cooperation, global terrorist threats cannot be effectively countered. Military medicine has long been in the forefront of international bridge-building in research, education and practice. Recently, force health monitoring and assessment have become central to prevention efforts, and many of these assessments rely on language-based tests and screening tools. In order for meaningful cross-national comparisons to be made, these tools must be standardized and validated across cultures. Too often, measures are simply translated into the new language, and differences in meaning go undetected while validity is diminished. This paper reports on a cross-national research project that aims to create a more valid tool for assessing stress resilience in military forces from multiple nations. An improved Norwegian hardiness (Dispositional Resilience Scale-DRS) measure is being developed, applying: (1) the principles of testing and international standards for test adaptation; (2) knowledge of the hardness – resilience construct; (3) knowledge of Norwegian language and culture. In addition, information gained regarding the conceptualization and measurement of hardness – resilience will be used to improve the original English language hardness scale - DRS. This project also serves as a model for other cross-national test adaptation efforts, providing better tools and procedures for psychological assessment of military personnel before, during and after deployment.

Importance

Military operations expose troops to a variety of stressors, both psychological and physical. The most extreme stressors are related to combat, including death and serious injury. But other operations also cause stress. In recent years the number and range of military operations have greatly increased, including peacekeeping, peacemaking, humanitarian assistance, counter-insurgency and stability operations. Troops involved in these operations encounter psychological stressors including boredom, ambiguity, and powerlessness (Bartone, Adler & Vahttkus, 1998). At the same time, many military organizations are downsizing due to shrinking budgets and changing national priorities. In the U.S., reduction and restructuring of the military has led to an increase in the number and duration of deployments for both active duty and reserve troops. More deployments also mean more training activities and exercises, planning and preparation, all of which increases the pace of operations. Thus an important question for military medical practitioners and policy makers concerns how to build up stress resilience in troops so as to reduce psychiatric casualties during and after military operations.

Hardiness: Concept and Measurement

“Hardiness” is a personality or cognitive style linked to continued good health and performance under stress (Kobasa, 1979). It includes three facets, Commitment, Control and Challenge. A 45-item hardiness measure (the Dispositional Resilience Scale-DRS) was developed by Bartone (1989), and later reduced to 30- and 15-item versions (Bartone, 1995). The DRS has been used extensively in U.S. military and non-military samples, with excellent results (e.g., Bartone, Ursano, Wright & Ingraham, 1989; Britt, Adler & Bartone, 2000). In a review of hardness theory and research, Funk (1992) recommended the DRS as the best available hardiness measure. Also using the DRS, Sinclair & Tetrick (2000) confirmed a factor structure of three facets, commitment control and challenge, nested under a more general hardness construct. The short DRS-15 scale was translated into Norwegian in 1998, and has since been used in multiple studies in Norway (Johnsen, Eid & Bartone, 2004; Bartone et al., 2002). Results have been positive, but there is evidence that the measure is in need of psychometric improvement in order to capture the concept of hardiness more effectively in Norwegian and other non-American groups.

Results (over)

Conclusion

Resilience under stress is a matter of high consequence in many occupations, but especially so in the military. The costs of unhealthy responses to stress can be devastating not only for individuals, but also for military units. Indeed, mission success can hinge upon how well the troops respond to operational stress. The personality style of hardiness offers one important explanation for how some people remain healthy while others degrade in health and performance when exposed to stressful conditions. More research is needed to clarify how hardiness and other possible resiliency factors operate. The present study makes an essential contribution by providing an improved, valid and reliable instrument for assessing resilience - hardiness in Norwegian, while also making improvements to the English language version. This process has already led to an intermediate corrected Norwegian scale, which is now being used with military units deployed to Lebanon and Afghanistan. Also, an improved hardness scale was translated into Swedish for use by Swedish naval forces deploying to Lebanon with the new United Nations peacekeeping mission there. In addition to more accurate research data for cross-national comparisons, results and insights from this work will also allow practitioners to design more effective training programs for building-up the kinds of healthy cognitions and behaviors that hardy persons display in response to stress.

References

Results

An important first step in the hardness test adaptation process was a review of previous studies with the Norwegian scale, paying close attention to the psychometric properties of this scale. Reliability coefficients for the Norwegian hardness scale and the three facets were somewhat low, suggesting possible translation problems. For example, overall scale reliability (Cronbach’s alpha) in a sample of Norwegian Navy cadets was only .61 (year 1998 cadets, as reported in Bartone et al., 2002), and reliability coefficients for the three facets were also somewhat low. Scale reliability was subsequently examined in three available Norwegian samples and three U.S. samples. Table 1 shows the Cronbach’s alpha coefficients for the hardness facets of commitment, control and challenge, and for the total hardness scale. The Cronbach’s alpha coefficient gives an estimate of scale reliability, and indicates the extent to which items in the scale intercorrelate with each other to form a coherent grouping. In this table, any coefficient lower than .60 is highlighted in bold to indicate unacceptably low alpha coefficients.

Table 1: Cronbach’s alpha coefficients for hardness scale in U.S. and Norwegian samples

<table>
<thead>
<tr>
<th></th>
<th>USA</th>
<th>Norwegian</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N=391</td>
<td>N=1379</td>
</tr>
<tr>
<td></td>
<td>Army N=2842</td>
<td>Army N=482</td>
</tr>
<tr>
<td></td>
<td>Army officers</td>
<td>Navy cadets</td>
</tr>
<tr>
<td></td>
<td>USMA CL ’98</td>
<td>Kosovo 2001</td>
</tr>
<tr>
<td>Commitment</td>
<td>.74</td>
<td>.60</td>
</tr>
<tr>
<td>Control</td>
<td>.58</td>
<td>.53</td>
</tr>
<tr>
<td>Challenge</td>
<td>.76</td>
<td>.73</td>
</tr>
<tr>
<td>Hardiness</td>
<td>.76</td>
<td>.72</td>
</tr>
</tbody>
</table>

DIF – Differential Item Functioning analysis was performed in order to identify items that may not be operating the same way in the two different languages and cultures (Allalouf, Hambleton & Sireci, 1999; Holland & Wainer, 1993). The DIF – Delta Plot method first elaborated by Angoff (1972) is still an extremely useful tool for identifying potential problem items (Angoff, 1993). Two closely comparable samples of U.S. and Norwegian military cadets were used for a DIF analysis of hardness items. The U.S. cadets (N = 436) completed the short hardness scale - DRS-15 during their senior year at the U.S. Military Academy – West Point. Norwegian Navy cadets (N=297) were near the end of their first year at the Royal Norwegian Navy Academy. The Norwegian cadets have military experience before entering the Naval Academy, and are close in age (Mean=23.1 years, s.d.=2.6) to the West Point cadets (Mean=22.1, s.d.=1.1). Both samples were mostly male, with similar percentages (8-13%) of women. Mean overall hardness levels were slightly higher in the American sample (Mean=30.76, s.d.=4.65) compared to the Norwegian (Mean=30.03, s.d.=4.42).

DIF analysis by the delta-plot technique (also known as “transformed item difficulty” or TID) followed these steps. Item responses were collapsed into binary categories, with codes 0 or 1 – 2 or 3 – 2. Next, the proportion of 2 (positive) responses was calculated for each item, within each sample. These scores were standardized within samples and converted to the standard metric used by ETS (Educational Testing Service), with Mean=13 and s.d.=4, as recommended for delta plots (Dorans & Holland, 1993). Finally, delta scores for each group were plotted, with the American scores on the x – horizontal axis, and Norwegian scores on the y – vertical axis. Each data point represents the delta scores for both groups on a particular hardness scale item. A line is drawn diagonally approximating the regression line that minimizes the distances between points. This line shows the area of perfect agreement. If both samples answered items exactly the same way, all points would fall directly on the line.

These data show a clear pattern of low reliability coefficients in the Norwegian samples for the Control facet of hardness, and marginal reliability coefficients for the Commitment facet. It is noteworthy that reliability coefficients for the Commitment facet are also somewhat low in the U.S. sample, suggesting that the original English version could also benefit from improvements.

DIF – Delta Plot Analysis of Hardness Scale American vs. Norwegian Military Cadets

In the DIF plot (Figure 1), points that stray far from the diagonal line are suggestive of DIF – Differential Item Functioning. Agreement between the Norwegian and English language versions is overall quite good, with 10 of 15 items showing little or no DIF (correlation between delta scores is .74). Five items are noticeably distant from the center line, suggesting DIF. Norwegians responded more favorably on items 12 and 7, while the Americans responded more favorably on items 1, 15 and 2. The next step is to examine these items closely using several techniques in order to identify causes of DIF. A detailed conceptual analysis of Norwegian and English items by bilingual experts is now underway. Early results suggest that some of the observed DIF is due to idiomatic terms that carry a different meaning when translated to Norwegian. Also, both language forms of the test are being administered to groups of bilingual Norwegian students, in what is known as a bilingual research design. These conceptual and quantitative analytic techniques will continue in an iterative process that will result in a final, highly reliable and valid Norwegian hardness scale.

*Thanks are due to the U.S. – Norwegian Fulbright Foundation for supporting this research.

Thanks also to the National Defense University, the Industrial College of the Armed Forces, the University of Bergen, Norway, the U.S. Army Medical Department, and the U.S. Army Medical Service Corps.

For further information contact: paul.bartone@psysp.uib.no bartonep@ndu.edu