Measurement invariance of alcohol use motivations in junior military personnel at risk for depression or anxiety

Jason Williams⁎, Sarah B. Jones, Michael R. Pemberton, Robert M. Bray, Janice M. Brown, Russ Vandermaas-Peeler

Behavioral Health and Criminal Justice Research Division, RTI International, Hobbs Building, 3040 Cornwallis Road, Research Triangle Park, NC 27709, USA

ARTICLE INFO

Keywords:
Drinking motivations
Measurement
Alcohol
Military
Mediation

ABSTRACT

Measurement invariance is typically assumed when assessing drinking-related constructs across distinct groups of respondents. However, measurement properties of motivations related to mood maintenance and stress relief may differ in those experiencing symptoms of depression or anxiety. Invariance of social and coping drinking motives were explored with a sample of 4133 junior enlisted Air Force and Navy personnel. Measurement did not differ in those with depression symptoms. In contrast, those with anxiety symptoms differed in measurement of both motives. The impact of non-equivalence was demonstrated with a mediation model in which anxiety and depression predicted drinking motives, which in turn predicted heavy drinking. Incorporation of the partial invariance of the social motives factor attenuated the estimate of the mediated effect of social drinking motives by almost half compared to the estimate with invariance assumed. These results suggest that lack of measurement invariance could seriously bias or alter conclusions from tests of theoretical models and highlight the need for researchers to carefully consider the measurement properties of their constructs prior to model estimation.

© 2009 Elsevier Ltd. All rights reserved.

1. Introduction

Extensive literature has tied alcohol use disorders to mood and anxiety problems in both the general population (Rodgers, Korten, Jorm, Jacomb, Christensen and Henderson, 2005; Kessler, Nelson, McGonagle, Edlund, Frank and Leaf, 1996; Swendsen, Merikangas, Canino, Kessler, Rubio-Stipec and Angst, 1998) and clinical samples (Kushner, Sher and Beitman, 1990; Merikangas & Angst, 1995; Burns, Teesson and O’Neill, 2005). Rates of comorbidity range from 33% in the overall population to over 65% in treatment-seeking samples (Haver & Dahlgren, 1995; Burns et al., 2005). This co-occurrence has been explained with multiple theories which vary the order and magnitude of the relations between drinking and psychiatric problems. These theories fall into three overarching categories: (1) misuse of alcohol leads to secondary depression or anxiety (Brown & Schuckit, 1988; Davidson, 1995), (2) depression and anxiety lead to problematic alcohol use (Cheng, Gau, Chen, Chang, & Chang, 2004; Kessler et al., 1996; Hartka, Johnstone, Leino, Motoyoshi, Temple and Fillmore, 1991), and (3) alcohol, mood, and anxiety disorders are all manifestations of common underlying genetic or environmental liabilities (Kendler, Walters, Neale, Kessler, Heath and Eaves, 1995; Tambs, Harris and Magnus, 1997). Each of these explanations may be applicable to varying degrees across individuals, complicating treatment design and outcome (Rodgers et al., 2005; Burns et al., 2005).

The self-medication hypothesis (Duncan, 1974) is perhaps the most well-known theory in the second category of causal models above and has garnered considerable support (e.g., Tomlinson, Tate, Anderson, McCarthy, & Brown, 2006; Burns et al., 2005; Carrigan & Randall, 2003). It posits that those with anxiety or mood psychopathology attempt to alleviate their symptoms through use of non-prescription means such as drinking, often leading to alcohol abuse or dependence (Lynskey, 1998). A strong motivational mechanism is implicit in the theory, suggesting that motives to drink that are related to mood maintenance, stress relief, and removal or avoidance of negative affective states will be elevated in those experiencing distressing symptoms (Stewart & Zeitlin, 1995). These motivations lead to increased alcohol use in an attempt to manage negative affective experiences. Thus, at least part of the link between psychiatric symptoms and alcohol use is explained by the intervening, or mediating, role of increased self-medication or coping motives.

Tests of models such as the mediation of anxiety and depression on drinking through motivations assume that alcohol use motive items are measured the same across all persons. Measurement tools must “mean the same thing” to all respondents, or put another way, they must measure things the same way (Millsap & Kwok, 2004; McDonald, 1999). Individuals from two groups with the same value on the underlying construct yield the same observed scores when

⁎ Corresponding author. Tel.: +1 919 541 6734.
E-mail address: jawilliams@rti.org (J. Williams).
inequality holds [Meredith & Millsap, 1992]. Conversely, two individuals from different groups or populations may be equal on the construct of interest but may yield different observed values if the instrument violates measurement invariance. Measurement invariance is necessary for meaningful interpretation of the impact of anxiety or mood problems on alcohol motivations as well as the mediated effect, which is estimated as the product of the path from anxiety/mood to motivations and the path from motivation to alcohol use (MacKinnon, Lockwood, Hoffman, West and Sheets, 2002). However, it is not clear if this assumption is valid, nor what the consequences of its violation may be. For example, items that assess reasons to drink such as reducing worry and helping maintain mood may be more central to an underlying self-medication or coping motivational dimension in depressed or anxious persons. Failure to establish that motivational constructs are perceived and responded to the same way in those with and without psychiatric problems may unknowingly bias or distort conclusions in situations where invariance does not exist across groups.

1.1. The current study

Junior military personnel are a population at high risk for both mental health and substance use problems (Riddle, Smith, Smith, Corbeil, Engel, Wells et al., 2007; Bray, Hourani, Rae Olnsted, Witt, Brown, Pemberton et al., 2006). The 2005 DoD Survey of Health Related Behaviors, a representative survey of active duty personnel, found 33% of junior enlisted personnel (E1–E3) were in need of further screening for depression. Data from the Millennium Cohort study indicated that only alcohol abuse was more prevalent in junior personnel from 2001 to 2003. The high risk of problematic alcohol use and mental health problems in this population calls for rigorous study and explanatory models to probe the interrelationships of alcohol, motivations, and psychiatric symptoms. However, these models and studies are at risk for biased and erroneous conclusions if measurement of motivations to use alcohol coexisted with problems differ in those with depression or anxiety symptoms and those without such distress.

This study was undertaken to examine the degree of measurement invariance present in alcohol use motivations in a sample of junior Air Force and Navy personnel. Measures of coping/self-medication and social alcohol motivations were tested for measurement equivalence in personnel at risk for anxiety or depression. The mediation model of the hypothesized causal path from psychiatric symptoms to alcohol motivations to heavy alcohol use was tested and the consequences of measurement invariance on this model of heavy drinking as a result of self-medication were explored.

2. Methods

2.1. Sample

This study was a collaborative effort of RTI International, University of Kentucky, and two branches of the United States military: Air Force and Navy. Data were collected via a 120-item self-administered questionnaire assessing alcohol and tobacco use administered at one Naval training base in November 2004 and March 2005, and at four Air Force training bases in April and May 2005. All Naval students in advanced training and all Air Force students in technical training aged 18 years or older were eligible to participate.

A total of 9684 trainees attended informational briefings during which project staff explained the study and requested their participation. Of these attendees, 6298 persons (3438 Navy; 2860 Air Force) chose to participate for a response rate of 65%. Data for the current study were based on 5446 respondents (Navy, 2997; Air Force, 2449) aged 18–25. Age was restricted to this range as junior enlisted personnel were the primary focus of the larger project that this study was a part of. Approximately 92% of the total sample was within this range however, so loss of cases to this constraint was minimal. Data were collected in group sessions by civilian research teams who explained the procedures and the voluntary nature of participation. No military personnel were in the room (other than the respondents) while the questionnaires were completed. Race/ethnicity of the sample was 65.5% white non-Hispanic, 13.6% Hispanic, 11.0% African-American non-Hispanic, and 9.9% Other non-Hispanic. The majority of the sample (82.6%) was male and single (86.2%). Most respondents (60.3%) reported high school or less education, with 36.4% reporting having attended some college and only 3.3 having a college degree. Over half (55.0%) of respondents were from the Navy. The average length of time from the end of basic training until participation in the survey for personnel aged 18–25 was 119.7 days (median = 57.0 days).

2.2. Measures

2.2.1. Drinking motivations

Respondents answered 19 items derived from the National Alcohol Survey (Greenfield, Midanik and Rogers, 2000) that inquired about reasons to use and not to use alcohol. Items assessed agreement with statements such as “drinking helps me relax,” “drinking can make me feel sick,” and “drinking increases my self-confidence.” Responses were “strongly disagree,” “disagree,” “don’t know/no opinion,” “agree,” or “strongly agree.” The nine items assessing motivations to use alcohol (as opposed to inhibitory motives) were included in this study and are shown in Table 1.

2.2.2. Depression

Probable depression was assessed using the five-item Center for Epidemiologic Studies Depression (CESD) scale (Shrout & Yager, 1989). Respondents were asked about how often in the past week they “could not shake off the blues even with help from family or friends,” “felt depressed,” “felt lonely,” “had crying spells,” or “felt sad.” Five response categories ranged from 0 (rarely or none of the time/less than 1 day) to 4 (most or all of the time/5–7 days). Items were summed and a cutoff score of 5 or more was used to create a binary indicator of presence of depressive symptoms. Cronbach’s alpha for the five depression items was 0.89. The mean depression score was 2.4 (sd = 3.5). Research on shortened versions of the full 20-item CESD instrument has indicated that the five-item version has adequate psychometric properties and nearly comparable performance as a screen for depression (Shrout & Yager, 1989; Zaudig, & Graham, 2009; Furukawa, Anraku, Hiroe, Takahashi, Kitamura, Hirai et al., 1997).

2.2.3. Anxiety

Anxiety was measured using the Anxiety Sensitivity Index (ASI; Reiss, Peterson, Gursky and McNally, 1986; Blais, Otto, Zucker, McNally, Schmidt, Fava et al., 2001), a 16-item question asking respondents how much they believed statements applied to them such as, “it scares me when my heart beats rapidly,” “unusual body sensations scare me,” and “it scares me when I am nervous.” Possible response items were “very much,” “much,” “some,” “a little,” or “very little” for each statement, with a value of 0 assigned to “very little” and 3 assigned to “very much.” The 16 items were summed and a cutoff of 25 or greater was used to indicate the presence of anxiety symptoms (Peterson & Reiss, 1992). The mean sum score of the ASI was 15.2 (sd = 13.7). Internal reliability for the anxiety items was high, with a Cronbach's alpha of 0.95.

2.2.4. Heavy drinking

The primary outcome measure was a binary indicator of heavy drinker status. Participants were classified as a heavy drinker if they
reported 4 or more binge episodes in the past month (or an average of one or more binge episodes a week). Binge episodes were defined as 5 (4 for females) or more drinks on one occasion.

2.3. Analysis

A combination of exploratory and confirmatory factor analysis (EFA, CFA) models were used to establish the factor structure of the 9 alcohol use motivation items. First, because the motivation items did not originate from an instrument with an established factor structure, EFA was used to suggest the optimal factor-item configuration. This model was then estimated with CFA to obtain various fit indices to compare it to several competing models. These indices were the standardized root-mean-square residuals (SRMR), comparative fit index (CFI), and root-mean-square error of approximation (RMSEA). Generally acceptable cutoffs are 0.08 or lower, 0.95 and higher, and 0.06 or lower, respectively (Hu & Bentler, 1999). The EFA-derived model was compared to a unidimensional motivation model and a model in which items were a priori assigned to factors based on Cooper and colleagues’ models of alcohol use motivation (Cooper, 1994; Cooper, Russel, Skinner, & Windle, 1992). Cooper’s Drinking Motives Questionnaire (DMQ) was derived from Cox and Klinger’s (1988) perspective of drinking as a rationale behavior that can be used to achieve desired outcomes such as relief of stress or conformity with social norms. The DMQ has 20 items and generally decomposes motivations to equate parameters across groups and test the degree of invariance (Steenkamp & Baumgartner, 1998; Vandenberg & Lance, 2000). Configural invariance indicates that each group perceives the underlying constructs or dimensions (i.e., factors) in the same way, with the same pattern of factor loadings (Meredith, 1993; Riordan & Vandenberg, 1994). Metric invariance is obtained when factor loadings for each item are the same across groups and is generally held as a prerequisite for meaningful across-group comparisons based on composites or scales (Bollen, 1989). Scalar invariance occurs when item intercepts can be equated across groups and indicates that observed scores are the same for the identical underlying factor score (Cheung & Rensvold, 2002; Meredith, 1993). Meaningful comparisons of latent means require scalar invariance. Strict invariance, in which residual errors or uniquenesses are equal, is obtained when the latent factors are measured with the same degree of error by the observed items in each group, but is not generally a requirement for group comparisons (Cheung & Rensvold, 2002). To meaningfully compare differences in measures of drinking motivation between those with anxiety or depression and those without, invariance at the scalar (for latent variables) level is required. Tests of group differences in drinking motivations and the mediation of the impact of psychopathology on drinking by motivational constructs may be biased or produce erroneous conclusions if measurement invariance at any of these three levels – configural, metric, and scalar – is not met.

Constraints were added to the measurement models of drinking motivations to equate parameters across groups and test the degree of measurement equivalence in each construct. Increasing levels of parameter constraints imposed configural, metric, scalar, and strict invariance. Constrained models were nested within each previous model and tested by comparing the change in the $\chi^2$, $\Delta \chi^2$, to the $\chi^2$ distribution with $df$ equal to the change in free parameters. All models were estimated using Mplus version 5 (Muthén & Muthén, 1998–2007). Invariance across depression/no depression and anxiety/no anxiety groups was evaluated independently as it was conceivable that one type of psychopathology would show invariance whereas the other did not depending on how the motivational assessment items were perceived. If either metric or scalar invariance were not met, partial invariance tests were conducted (Byrne, Shavelson and Muthen, 1989; Millsap & Kwok, 2004). This involved a series of iterative models in which parameters (loadings for metric invariance, intercepts for scalar) were freed for individual items. Through this process, separate items could be determined to be invariant or to differ across groups. Finally, an exploration of the impact of lack of invariance was conducted using the multiple mediator model and added parameters in the measurement model of the drinking motive constructs to include the non-equivalent measurement across groups.
Mplus programs for testing invariance, partial invariance, and mediation are available from the first author upon request.

3. Results

3.1. Factor structure of motivations to drink

Exploratory factor analysis on the nine alcohol use motivation items suggested two factors. The first included social motivations to drink. The second factor captured motivations directed primarily at mood maintenance or self-medication (i.e., coping). The CFI and SRMR indices indicated good model fit (CFI = 0.96, SRMR = 0.03), although RMSEA indicated a somewhat more marginal fit (RMSEA = 0.09). The upper limit of the RMSEA 90% confidence interval did not exceed 1.0 which would have indicated poor fit on this index and so the model was retained without further modification. The two dimensions of coping and social were highly correlated however, with a factor correlation of 0.83. A single-factor CFA indicated that both dimensions were needed as the one factor model had considerably worse fit, below accepted cutoffs for some of the three indices (CFI = 0.92, RMSEA = 0.14, SRMR = 0.04). In addition to the superior model fit, the 2-factor model was preferable as it would allow differentiation of the impact of anxiety and depression on coping motivations to use alcohol and more general social motives. The 2-factor model based on probable factor-item relationships derived from Cooper (1994) fits better than the single-factor model, but was inferior to the empirically derived factor structure, most likely due to the lack of strictly parallel motivation items in the two studies. Despite this difference in fit, the majority of items loaded on the same factor across the two models, and indeed exhibited very similar loadings. Two items – drinking to relax and drinking to increase self-confidence – loaded on the social factor instead of the coping/self-medication factor as would be expected under the Cooper model. The correlation of the two factors was slightly higher in the Cooper model (r = 0.85). The 1- and 2-factor models are summarized in Table 1.

3.2. Prevalence of mental health risk and heavy drinking

Out of the total pool of 5446 respondents in the targeted age range (18–25), 1404 (25.8%) did not respond to any of the anxiety questions and 1450 (26.6%) did not respond to the depression items. Exploratory analyses did not find a difference by service in this lack of response and responders did not differ from non-responders on a variety of alcohol use measures. Because indicators of anxiety or depression symptoms were necessary for the MCFA analysis of measurement invariance, the sample was subsequently limited to those with valid scores for at least one of these mental health indices (N = 4133). Approximately 24.1% of the remaining personnel indicated significant experience of anxiety symptoms. A somewhat lower number, 21.4%, reported depression symptoms that placed them above the cutoff score on the CESD-5. About 6% reported symptoms of both anxiety and depression.

Preliminary analysis demonstrated that both risk for anxiety and risk for depression were predictive of significantly (both p < 0.001) elevated heavy drinking, controlling for age, minority status, gender, service, and weeks since completion of basic training. Prevalence rates of heavy drinking adjusted for these covariates are presented in Table 2. The rate of heavy drinking in the at-risk groups was approximately twice that of the non-risk groups, both for raw prevalence estimates as well as for those adjusted for demographic and service related covariates. The overall unadjusted rate of heavy drinking was 9.1%.

3.3. Mediation by motivations to use alcohol

Motivations to drink alcohol that include mood maintenance or relief of stress in social situations are hypothesized to serve as mediators of the impact of depression or anxiety on heavy alcohol use. In other words, the significantly higher rate of heavy alcohol use in those with these mental health risks may in part be due to increased motivation to use alcohol to maintain mood and relieve stress. Increasing these motivational constructs then leads to increased alcohol use. As shown in Table 3, all paths from mental health groups to the two mediators and from the mediators were significant and positive. Thus, those with anxiety or depression symptoms showed elevated motivations to use alcohol for coping and social facilitation. Greater levels of these motivations were in turn associated with significantly more heavy drinking, and the mediated effects for both motivations were significant for both risk groups. However, these results may not be entirely valid if the measurement properties of the motivation items vary by mental health risk.

3.4. Measurement invariance

3.4.1. Depression

Motivations to use alcohol displayed a high degree of measurement invariance across the no depression and depression groups. Equating the factor loadings (metric invariance) and observed item intercepts (scalar invariance) did not significantly affect the overall model fit, with $\Delta\chi^2 = 9.39$ (df = 7, ns) and $\Delta\chi^2 = 11.92$ (df = 7, ns), respectively. The strict invariance test, in which the residual variances and the means were constrained to be equal, produced different conclusions about the measurement equivalence across these groups. It was highly significant however ($\Delta\chi^2 = 124.07$, df = 9, p < 0.001). Table 4 summarizes the incremental change in fit and the significance of these changes across the increasingly invariant measurement models.

3.4.2. Anxiety

Models comparing the measurement of drinking motivations in those at risk for anxiety to motivations in the no anxiety group produced different conclusions about the measurement equivalence across groups. Constraining the factor loadings and item intercepts to equality across groups each produced significant changes in the $\chi^2$ ($\Delta\chi^2 = 18.52$, df = 7, p < 0.01; $\Delta\chi^2 = 76.11$, df = 7, p < 0.001). Given these failures of invariance, strict invariance was not examined. Following the finding of invariance for both the factor loadings and intercepts, partial invariance was explored. This was particularly relevant because the source of the variant measurement properties could lie in one of the alcohol use motivation factors and not the other, a situation suggesting measurement invariance of one construct but not both. First, all the loadings of the social factor were constrained to be equal. This produced a $\Delta\chi^2$ of 8.15 with 4 degrees of freedom, a non-significant decrease in fit. Constraining all the intercepts of this factor to invariance yielded a large difference, $\Delta\chi^2 = 66.33$, p < 0.001, indicating that the intercepts were not the same across groups. A series of iterative models was then estimated that tested the invariance of each intercept. These models revealed that the source of invariance lay with two items: (1) drinking increases sexual opportunities, and (2) drinking increases my self-confidence. Those in the anxiety risk group reported higher levels for these two items, an increase that was beyond that accounted for by increases in the underlying factor. Exploration of invariance for the coping factor by itself found a lack of invariance beyond configural. Iterative models were again used to determine which loadings were invariant, if any, since measurement was not equivalent at the metric level. These

Table 2

<table>
<thead>
<tr>
<th>Raw</th>
<th>Adjusted</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Anxiety</td>
</tr>
<tr>
<td>No</td>
<td>0.0731</td>
</tr>
<tr>
<td>Yes</td>
<td>0.1376</td>
</tr>
</tbody>
</table>

Note: Adjusted prevalence rates adjust for gender, age, service, minority status, and weeks since completing basic training.
models revealed that all loadings were invariant across group except for the item “drinking helps me forget my worries.” This item’s loading was greater for the anxiety risk group, indicating it was a more central item in assessing this factor for this group than it was for the no anxiety risk group. At the scalar level, it was found that all items on the coping factor were invariant except for the “forget my worries” item.

3.5. Group differences in alcohol motivation

As mentioned previously, for the mediation model to yield valid results that can be used to explain (at least in part) how mental health risks are related to increased heavy drinking, it is critical that the constructs being measured are perceived the same way by each group of respondents. Measurement invariance was found at the scalar (factor loading and intercept) level for the depression/no depression groups. Thus any group differences in drinking motivations by membership in these groups can be reasonably expected to be valid and not an artifact of differential measurement properties.

Anxiety models cannot be interpreted as readily however. The MCFA models indicated failure of metric and scalar invariance for the overall model. This carried over to the partial invariance exploration, which indicated that coping lacked anything beyond configural invariance. This complicated interpretation of the group effects on this factor in the mediation model. Although the impact of anxiety risk compared to no risk was highly significant ($\beta=0.26$, SE = 0.04, $p<0.001$), it is unclear whether this difference is wholly or partially due to variation in how the items assessing this factor were perceived and responded to by those in the anxiety risk group. The significant mediated effect for the path from anxiety to heavy drinking through coping is similarly ambiguous.

The social motivation factor for the anxiety group did show metric invariance, although the intercepts were significantly different. This level of invariance permitted exploration of the impact on the group to social motivation effect through the use of MIMIC (Multiple Indicator, Multiple Causes) structural equation models. MIMIC models expand basic CFA measurement models by adding paths from a predictor, usually a group indicator, to the latent factor and one or more of the observed indicators (Johnson, Morgan-Lopez, Breslau, Hatsuksi & Bierut, 2008; Chen & Anthony, 2003). This formulation parcels the impact of a grouping variable on observed responses into a measurement invariant part — the effect transmitted through the group’s impact on the latent variable (which then impacts all items), and a non-invariant part captured by the individual direct path from the group item to an observed variable. Significant direct paths, as they circumvent the latent factor and thus have idiosyncratic impact on each observed variable, can capture sources of scalar variance. The model in Fig. 1 was modified by adding paths from the anxiety group indicator directly to the observed indicators of the social factor until all significant paths were determined. Those in the anxiety group reported elevated endorsement of two items: (1) drinking increases sexual opportunities, and (2) drinking increases my self-confidence. These paths indicated that these higher reports were not directly attributable to measurement of the overall social factor, and that by ignoring these paths one would estimate higher levels of social drinking motivation than would be found with measurement invariance at the scalar level. Such model misspecification could influence conclusions from subsequent analyses. Indeed, in this instance, inclusion of these paths attenuated the path from anxiety group to social motivation (from $\beta=0.20$, SE = 0.04, $p<0.001$ to $\beta=0.11$, SE = 0.05, $p<0.05$). This in turn attenuated the estimate of the mediated effect, reducing its estimate by 48%.

4. Discussion

Measurement of motivations to drink was not found to differ for those with depression symptoms and those without. Both the social and coping factors met the criteria for strong (Meredith, 1993) or

<table>
<thead>
<tr>
<th>Mediator</th>
<th>Depression</th>
<th>Mediation effect on heavy alcohol use (estimate, SE)</th>
<th>Anxiety</th>
<th>Mediation effect on heavy alcohol use (estimate, SE)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coping</td>
<td>Mediator effect</td>
<td></td>
<td>Mediation effect</td>
<td></td>
</tr>
<tr>
<td>Social</td>
<td>impact</td>
<td>0.31 (0.04)**</td>
<td>0.33 (0.03)**</td>
<td>0.10 (0.02)**</td>
</tr>
<tr>
<td></td>
<td>(estimate, SE)</td>
<td>0.32 (0.03)**</td>
<td>0.10 (0.02)**</td>
<td>0.20 (0.04)**</td>
</tr>
</tbody>
</table>

Note: ** = $p < 0.001$.

### Table 3

Mediated effect estimates.

### Table 4

Change in fit indices for sequential invariance models.
scalar (Mullen, 1995) invariance. Thus, those with this mental health risk have the same operational definition and zero points for social and coping motivations to drink as those who do not have this risk. Consequently, estimates of the impact of depression on motivations to drink and the mediated effect of depression on heavy drinking through motives were unaffected by differential measurement.

However, results from the MCFA invariance tests suggest that there are different measurement properties for those who have anxiety symptoms and those who do not. The anxiety group displayed both different factor loadings and observed intercepts, violating scalar and metric invariance. Any observed differences in the motivations to use alcohol measures may be due to differences in how the anxiety group perceived or interpreted items of the instrument. The exploration of partial invariance by motivational factor suggested that much of the failure in the metric component was in the coping domain, whereas for both social and coping the item intercepts were quite different across groups. Invariance of the coping factor was primarily violated by the item assessing drinking to forget worries. Social motivation to drink items were similar in relationship (factor loadings) to the underlying factor across anxiety and no anxiety risk groups, but two items were responded to at an elevated level by those in the anxiety group, producing violations of scalar invariance. This lack of invariance may have important implications for models of drinking behavior that included psychiatric problems and motivations to use alcohol. This is illustrated by the almost 50% attenuation of the drinking behavior that included psychiatric problems and motivations to drink and the mediated effect of depression on heavy drinking as those who do not have this risk. However, the ΔCFI test is not a clear panacea for testing invariance in large samples. Although Cheung and Rensvold's simulations indicated the superior performance of ΔCFI in large samples, there was a non-negligible impact on the estimated mediated effect, almost a 50% difference, could be used to represent these motivations however. Given the high response extremity has been found when items are particularly salient or personally meaningful (Warr & Coffman, 1970; Greenleaf, 1992). Perceiving items to be more central or important may also explain differences in the metric portion of the measurement models, in that factor loadings were variant in the anxiety group because items were perceived to be more centrally tied to the latent dimension (reflected in a significantly larger factor loading). In contrast, depression has not been associated with response extremity and in this study motivations to use alcohol were invariant at the metric and scalar levels.

Despite support for significantly different measurement in those with anxiety, the practical meaningfulness of these findings is not unequivocal. The impact of sample size on χ² and the Δϕ test in MCFA is well known (Brannick, 1995; Kelloway, 1995) and so the apparent differences in measurement for the anxiety group may simply be an artifact of this study's large sample. The relationship of χ² and sample size has lead some researchers to propose alternative tests of nested models using other, more sample agnostic, goodness of fit indices. After a large simulation study of alternative goodness of fit criteria for invariance testing, Cheung and Rensvold (2002) recommended the difference in CFI (ΔCFI) as a method superior to Δϕ in large samples. Using ΔCFI in the present study produced non-significant changes for the metric and scalar invariance for the depression and anxiety groups, suggesting no difference in measurement properties for those with either mental health risk. However, the ΔCFI test is not a clear panacea for testing invariance in large samples. Although Cheung and Rensvold's simulations indicated the superior performance of ΔCFI, follow-up simulations (French & Finch, 2006) found the test to perform more erratically than Δϕ and supported the use of Δϕ even in large samples. In the present study, the ΔCFI indicated no difference in measurement, but there was a non-negligible impact on the substantive mediation model when the invariance indicated by the Δϕ test was incorporated using a MIMIC model. The large difference in the estimated mediated effect, almost a 50% difference, could potentially influence conclusions about significant mediators and pathways of causation, and suggests that the Δϕ test yielded important information about how those with anxiety differed in their responses from those without symptoms.

A cautionary note must be made about the model of drinking motivations itself. In this study, there were two distinct yet highly correlated factors of motivations to drink alcohol. Other models may be used to represent these motivations however. Given the high response extremity has been found when items are particularly salient or personally meaningful (Warr & Coffman, 1970; Greenleaf, 1992). Perceiving items to be more central or important may also explain differences in the metric portion of the measurement models, in that factor loadings were variant in the anxiety group because items were perceived to be more centrally tied to the latent dimension (reflected in a significantly larger factor loading). In contrast, depression has not been associated with response extremity and in this study motivations to use alcohol were invariant at the metric and scalar levels.

Despite support for significantly different measurement in those with anxiety, the practical meaningfulness of these findings is not unequivocal. The impact of sample size on χ² and the Δϕ test in MCFA is well known (Brannick, 1995; Kelloway, 1995) and so the apparent differences in measurement for the anxiety group may simply be an artifact of this study’s large sample. The relationship of χ² and sample size has lead some researchers to propose alternative tests of nested models using other, more sample agnostic, goodness of fit indices. After a large simulation study of alternative goodness of fit criteria for invariance testing, Cheung and Rensvold (2002) recommended the difference in CFI (ΔCFI) as a method superior to Δϕ in large samples. Using ΔCFI in the present study produced non-significant changes for the metric and scalar invariance for the depression and anxiety groups, suggesting no difference in measurement properties for those with either mental health risk. However, the ΔCFI test is not a clear panacea for testing invariance in large samples. Although Cheung and Rensvold’s simulations indicated the superior performance of ΔCFI, follow-up simulations (French & Finch, 2006) found the test to perform more erratically than Δϕ and supported the use of Δϕ even in large samples. In the present study, the ΔCFI indicated no difference in measurement, but there was a non-negligible impact on the substantive mediation model when the invariance indicated by the Δϕ test was incorporated using a MIMIC model. The large difference in the estimated mediated effect, almost a 50% difference, could potentially influence conclusions about significant mediators and pathways of causation, and suggests that the Δϕ test yielded important information about how those with anxiety differed in their responses from those without symptoms.

A cautionary note must be made about the model of drinking motivations itself. In this study, there were two distinct yet highly correlated factors of motivations to drink alcohol. Other models may be used to represent these motivations however. Given the high response extremity has been found when items are particularly salient or personally meaningful (Warr & Coffman, 1970; Greenleaf, 1992). Perceiving items to be more central or important may also explain differences in the metric portion of the measurement models, in that factor loadings were variant in the anxiety group because items were perceived to be more centrally tied to the latent dimension (reflected in a significantly larger factor loading). In contrast, depression has not been associated with response extremity and in this study motivations to use alcohol were invariant at the metric and scalar levels.

Despite support for significantly different measurement in those with anxiety, the practical meaningfulness of these findings is not unequivocal. The impact of sample size on χ² and the Δϕ test in MCFA is well known (Brannick, 1995; Kelloway, 1995) and so the apparent differences in measurement for the anxiety group may simply be an artifact of this study’s large sample. The relationship of χ² and sample size has lead some researchers to propose alternative tests of nested models using other, more sample agnostic, goodness of fit indices. After a large simulation study of alternative goodness of fit criteria for invariance testing, Cheung and Rensvold (2002) recommended the difference in CFI (ΔCFI) as a method superior to Δϕ in large samples. Using ΔCFI in the present study produced non-significant changes for the metric and scalar invariance for the depression and anxiety groups, suggesting no difference in measurement properties for those with either mental health risk. However, the ΔCFI test is not a clear panacea for testing invariance in large samples. Although Cheung and Rensvold’s simulations indicated the superior performance of ΔCFI, follow-up simulations (French & Finch, 2006) found the test to perform more erratically than Δϕ and supported the use of Δϕ even in large samples. In the present study, the ΔCFI indicated no difference in measurement, but there was a non-negligible impact on the substantive mediation model when the invariance indicated by the Δϕ test was incorporated using a MIMIC model. The large difference in the estimated mediated effect, almost a 50% difference, could potentially influence conclusions about significant mediators and pathways of causation, and suggests that the Δϕ test yielded important information about how those with anxiety differed in their responses from those without symptoms.

A cautionary note must be made about the model of drinking motivations itself. In this study, there were two distinct yet highly correlated factors of motivations to drink alcohol. Other models may be used to represent these motivations however. Given the high...
degree of interrelatedness, it may be that alcohol use motivations may be better modeled as a General–Specific (G–S) model (Gustafsson & Balke, 1993). These models are often used for intelligence and ability or tendency measures (e.g., Gignac, 2005, 2006) and would reflect a general motivation to use alcohol, as well as specific facets of use such as mood maintenance and relief of negative affective states. Another alternative structure, proposed by Grant, Stewart, O’Connor, Blackwell, & Conrad, (2007), explicitly posits separate coping motivations for anxiety and depression. This five factor instrument has shown promising psychometric properties in several samples that have supported these anxiety- and depression-specific drinking motives and some differences in drinking behavior has been associated with these factors. However, the extent to which these factors differentiate between those experiencing depression or anxiety on each of these drinking motives is unknown at this time.

Although these findings suggest non-equivalence for those with anxiety, several caveats may limit the generalizability of this finding. First, the military sample is somewhat unique. The strong military drinking culture and greater stress of respondents (and hence more prevalent symptoms of depression and anxiety) made this an ideal high risk sample with which to study the measurement of motivations to drink and the relationships of mental health, motivations, and alcohol use. However, it may also limit how general these findings are for the general population. Similarly, care should be exercised in that the depressed and anxious groups were not based on formal diagnoses, and were instead based on DSM-type screens of symptoms. Replication of these findings with non-military respondents and groups based on formal diagnoses should be conducted to extend the confidence of how general these effects are. Another limitation is that the drinking motivation items used in this study were very similar to Cooper’s (1994), but not strictly parallel. Consequently, the results reported here may not generalize to investigations using Cooper’s measures of coping, social, environmental, and conformity alcohol motives.

Other limitations include the lack of longitudinal data and a significant amount of missing data. Although exploratory analyses suggested that the missing anxiety and depression data did not vary systematically by a variety of factors such as service and alcohol use, it is still concerning that these items appear to be subject to respondent self-censoring. At the very least, the amount of data missing negatively impact precision and power of the analyses presented. Lack of longitudinal data may limit the substantive conclusions about the mediation model used to illustrate the impact of the lack of measurement invariance. Mediation hypotheses include a strong temporal requirement that some causal variable first changes the mediator, and these changes in the mediator are then related to changes in the outcome. This has lead some (e.g., Cole and Maxwell, 2003) to argue that mediation is legitimately tested only with three or more assessments, one for each variable in the mediated causal chain. This requirement is equivocal however, and the coping mediation model tested here with cross-sectional data makes the assumption that the temporal ordering of the constructs comports to the theoretical ordering of this theory of the drinking–psychiatric relationship. Although this assumption may not be true, the model still functions as an illustration of how lack of measurement invariance may impact models of interest.

The results of this study highlight the need of researchers to consider the constructs under study and their measurement properties. Such consideration of measurement is often overlooked, with analyses focused solely on evaluating the hypothesis of greatest interest. As shown, ignoring possible group differences in how those with a risk for anxiety compared to those without such risk can influence the magnitude of findings if not outright alter conclusions about model results.

Role of Funding Source
This study was supported by the USAMARDC Fort Detrick Peer Reviewed Medical Research Program grant number DAMD17-00-1-0581. The views, opinions, and findings contained in this report are those of the authors and should not be construed as an official Department of Army position, policy, or decision, unless so designated by other official documentation. The Department of Army reviewed and approved the study design and implementation but had no part in this research report.

Contributors
J. Williams: analysis, literature review, and manuscript writing; S.B. Jones and R. Vandermaas-Peeler: literature review, project management; R.M. Bray, Pemberton, M.R., and J.M. Brown: study design and implementation. All authors contributed to the writing of the manuscript and have reviewed the final version.

Conflict of Interest
All authors declare that they have no conflicts of interest.

Acknowledgements
The authors thank Antonio Morgan-Lopez and Scott Novak for their helpful comments on previous versions of this manuscript.

References


