Multidimensional Inventory of Hypochondriacal Traits: An Examination of a Bifactor Model and Measurement Invariance Between Those With and Without a Self-Reported Medical Condition

Joseph R. Bardeen¹ and Thomas A. Fergus²

Abstract
The Multidimensional Inventory of Hypochondriacal Traits (MIHT) is a self-report measure that assesses four interrelated domains of health anxiety (i.e., Cognitive, Behavioral, Perceptual, Affective). Prior research has supported a correlated four-factor model, as well as a hierarchical model, in which each of the four factors load onto the higher order health anxiety construct. However, a bifactor modeling approach has yet to be used to examine the factor structure of the MIHT. Results supported a bifactor model of the MIHT in three different samples (i.e., unselected based on current medical status [n = 824], and those with [n = 348] and without [n = 354] a self-reported medical condition). The MIHT appears to be strongly multidimensional, with three of the four subscales providing substantive value. Confirmatory factor analysis supported the configural and metric/scalar invariance of the bifactor model between those with and without a self-reported medical condition. Results provide support for a bifactor conceptualization of the MIHT and the invariance of that model across levels of current health status.

Keywords
Multidimensional Inventory of Hypochondriacal Traits, MIHT, health anxiety, bifactor, measurement invariance, confirmatory factor analysis

The term health anxiety is commonly used to refer to a wide variety of health-related worries (Taylor & Asmundson, 2004). Taxometric analyses suggest that health anxiety is a dimensional construct, a matter of a difference in quantity, not kind (Ferguson, 2009; Longley et al., 2010). For many individuals, health anxiety persists despite reassurance that a medical condition does not exist (Warwick & Salkovskis, 1990). However, health anxiety also exists among individuals who have current medical conditions (e.g., Alberts, Sharpe, Kehler, & Hadjistavropoulos, 2011; DeMarinis, Barsky, Antin, & Chang, 2009; Kehler & Hadjistavropoulos, 2009; Tang et al., 2009). Individuals with, versus without a current medical condition report significantly greater levels of health anxiety (e.g., Alberts et al., 2011; DeMarinis et al., 2009; Kehler & Hadjistavropoulos, 2009; Tang et al., 2009). Moreover, health anxiety relates to greater health preoccupation, poorer physical health, increased disability, greater psychological symptom severity, and increased avoidant coping among individuals with medical conditions (Abramowitz & Braddock, 2008; Fink, Ørnøl, & Christensen, 2010; Taylor & Asmundson, 2004). Individuals with health anxiety often exhibit maladaptive behaviors in an attempt to mitigate health-related worries, such as excessive reassurance seeking from friends or family members (Noyes et al., 2003), overutilization of medical services (Fink et al., 2010), and excessive online searching for health-related information (Fergus, 2013). In contrast to avoidance through care-seeking behaviors (e.g., reassurance), some individuals with extreme health anxiety avoid seeking out medical care for fear of receiving confirmation that they are seriously ill (i.e., “care-avoidant”; Starcevic, 2013). The personal and economic burden of health anxiety has led some to call for increased screening for, and treatment of, health anxiety (e.g., Fink et al., 1999; Tyrer et al., 2011).

To heed those calls, it is of course necessary to have psychometrically sound, comprehensive, measures of health anxiety to aid in screening, case conceptualization, treatment

¹Auburn University, Auburn, AL, USA
²Baylor University, Waco, TX, USA

Corresponding Author:
Joseph R. Bardeen, Department of Psychology, Auburn University, Auburn, AL 36849, USA.
Email: jbardeen@auburn.edu
planning, and outcome research (Stewart, Sherry, Watt, Grant, & Hadjistavropoulos, 2008). Among existing psychological interventions, cognitive–behavioral therapy seems particularly efficacious in the treatment of health anxiety (Olatunji et al., 2014). Four interrelated factors are central to the cognitive–behavioral model of health anxiety, including affective (health-related worry and fear), cognitive (dysfunctional beliefs, especially regarding disease conviction despite contrary evidence), perceptual (hypervigilance for physical sensations), and behavioral (avoidance behaviors, especially reassurance seeking; Longley, Watson, & Noyes, 2005; Warwick & Salkovskis, 1990). Although a number of measures of health anxiety exist, there is only one, to our knowledge, that has been developed to be consistent with this model. Specifically, the Multidimensional Inventory of Hypochondriacal Traits (MIHT; Longley et al., 2005) was developed to provide an assessment of these four domains via self-report and address limitations of other self-report measures of health anxiety (see Witthöft, Weck, & Gropalis, 2015). Although the content of some of the MIHT scales has been met with critiques (e.g., Fergus & Valentiner, 2011; Olatunji, 2008), the MIHT remains a useful measure for assessing health anxiety (Witthöft et al., 2015).

The original validation study of the MIHT provided psychometric support for scores on the 31-item measure in both undergraduate and medical patient samples, including evidence in support of internal consistency, test–retest reliability over an 8-week interval, and convergent and discriminant validity (Longley et al., 2005). Moreover, results from confirmatory factor analysis (CFA) supported the hypothesized correlated four-factor model. The correlated four-factor model has also been shown to have adequate fit among patients with hypochondriasis (n = 178; Witthöft et al., 2015). Longley et al. (2005) concluded that their findings suggest that “each scale assesses a related, but separable, trait,” or “different aspects of the same syndrome” (p. 12). However, a correlated trait model does not provide evidence that the four factors represent the same overarching construct (Brown, 2015). Examination of a hierarchical model, in which each of the four factors load significantly onto the higher order health anxiety construct, would help support such a conclusion. This examination is especially important because researchers and clinicians commonly sum all of the items of the MIHT as an indicator of health anxiety. As noted by Clark and Watson (1995), it is essential when using a measure in this manner “to establish that all of the items—regardless of how they are placed in the various subscales—define a single general factor” (p. 318).

Based on this rationale, Stewart et al. (2008) used CFA to examine the factor structure of the MIHT in an undergraduate sample (n = 535). Both the hierarchical and correlated four-factor models provided adequate fit to the data. An examination of the model comparison statistics indicated that neither model evidenced superior fit. Stewart et al. (2008) suggested retention of the hierarchical model due to its “superior parsimony.” Of note, three of the four domain-specific factors evidenced loadings on the higher order health anxiety construct that supported the hypothesis that these factors are related, but distinct (i.e., Perceptual = .49, Cognitive = .64, Behavioral = .71), whereas the Affective domain exhibited a factor loading (i.e., .91) that may indicate redundancy with the higher order factor. That is, the Affective factor may not capture unique information beyond that of the higher order construct. Adequate fit of the hierarchical model, as well as invariance of this model by gender, was also observed by MacSwain et al. (2009). Moreover, in their undergraduate sample (N = 950), MacSwain et al. (2009) found factor loadings for the lower order factors on the health anxiety construct (i.e., Perceptual = .51, Cognitive = .67, Behavioral = .70, Affective = .89) that were consistent with Stewart et al. (2008).

As described, the MIHT items are summed to create subscales and a total scale score. Two assumptions are central to these approaches for operationalizing the MIHT. First, use of the total score assumes that the subscales represent the same overarching health anxiety general factor. Second, use of the subscales assumes that the factors represented by these scales provide unique information beyond that of the higher order health anxiety construct. Whereas a hierarchical model can be useful in examining the first of these two assumptions, it is unable to address the second. The structure of the MIHT must be examined using a bifactor modeling approach to determine the value of the domain-specific (commonly referred to as a lower order) factors beyond the general factor (Reise, 2012).

Similar to testing a hierarchical model, a bifactor model can indicate the presence of a general factor. However, in contrast to a hierarchical approach, direct effects from the general factor to the indicators are modeled in bifactor analysis (Reise, 2012). Because the general factor is modeled independently from the domain-specific factors, the unique contributions of both the domain-specific and general factors can be isolated, thus allowing for the simultaneous investigation of the overarching construct and the degree to which the lower order constructs are distinct from this general factor. For example, the degree to which the Affective factor of the MIHT, or any other domain-specific factor for that matter, is redundant with the general health anxiety factor can be determined through this bifactor approach. Subscales, derived from domain-specific factors, that exhibit redundancy with the general factor are questionable for use (Reise, 2012). Similarly, the degree to which the domain-specific factors all appear redundant with a general factor can provide evidence for or against the use of a total score (Reise, 2012). To our knowledge, an examination of a bifactor model of the MIHT has yet to be reported in the extant literature.
Another unaddressed question regarding the structure of the MIHT relates to whether it is similar across individuals with (vs. without) a current medical condition. Although Longley et al. (2005) included a medical sample in their initial validation study of the MIHT, structural analyses were not completed using data from that sample. As noted, health anxiety appears to be an important construct among individuals who currently have medical conditions (e.g., Alberts et al., 2011; DeMarinis et al., 2009; Kehler & Hadjistavropoulos, 2009; Tang et al., 2009). As described by Sue (1999), our discipline has “not followed good scientific principles in assuming that findings from research on one population can be generalized to other populations” (p. 1073). As such, the lack of research examining whether the factor structure of the MIHT is similar between those with and without a current medical condition represents a substantive gap in the extant literature. Using an alternative measure of health anxiety, Alberts et al. (2011) found that scores on items of that measure functioned somewhat differently between those with (vs. without) a current medical condition. Evidence of measurement invariance in the MIHT factor structure among individuals with (vs. without) a current medical condition would support its use for assessing the four core factors implicated in cognitive–behavioral models of health anxiety irrespective of the current medical status of respondents.

In the present study, we examined the structure of the MIHT in a large sample of general population adults (Sample 1). We recruited two additional samples (i.e., those with [Sample 2] and without [Sample 3] a self-reported medical condition) to confirm the structural findings from Sample 1 and examine measurement invariance based on current health status. Specifically, consistent with previous research examining measurement invariance of bifactor models, configurational and metric/scalar invariance of the MIHT was examined in the present study (e.g., Ebesutani, McLeish, Luberto, Young, & Maack, 2014; Olatunji, Ebesutani, & Abramowitz, 2017). We predicted that a bifactor model most accurately represents the structure of the MIHT. This prediction is consistent with cognitive–behavioral theory suggesting that health anxiety consists of four related, but distinct, domain-specific factors (Longley et al., 2005; Taylor & Asmundson, 2004). Some evidence suggests that comparing bifactor models with alternative models may have limited utility because these models often provide better fit to the data due to their inherent flexibility (Bonifay, Lane, & Reise, 2017; Reise, Kim, Mansolf, & Widaman, 2016). As such, in addition to comparing the bifactor model with alternate models, we examined a number of additional statistical indices to further evaluate the proposed bifactor model (Rodriguez, Reise, & Haviland, 2016).

Given that the MIHT is used in medical (e.g., Jones, Hadjistavropoulos, & Sherry, 2012), nonmedical, and samples unselected based on current medical status, it is important to ensure that the factors function consistently across levels of current health status. Given the exploratory nature of this aim, no a priori hypothesis was made regarding measurement invariance. Consistent with previous research (e.g., Alberts et al., 2011; DeMarinis et al., 2009; Kehler & Hadjistavropoulos, 2009; Tang et al., 2009), we expected those with a current self-reported medical problem to report significantly higher levels of health anxiety than those without a current self-reported medical problem.

Method

Participants and Procedure

Participants for all three samples (Sample 1 [n = 824], Sample 2 [n = 348], Sample 3 [n = 354]) were recruited via Amazon Mechanical Turk (MTurk), an online labor market where community adults can complete questionnaires in exchange for financial compensation. MTurk participants have been found to produce high-quality data and be more demographically diverse than undergraduate and other Internet samples (see Chandler & Shapiro, 2016, for a review). The average age and sex makeup of each sample is as follows: Sample 1 = 34.2 years (SD = 12.5) and 60.4% female, Sample 2 = 39.44 years (SD = 13.62) and 62% female, Sample 3 = 30.82 years (SD = 9.41) and 48% female. In Sample 1, 80.9% self-identified as White, 6.9% as Black, 6.4% as Asian, 1.0% as American Indian or Alaska Native, 2.9% endorsed “other,” 1.8% preferred not to respond, and 6.9% of Sample 1 identified their ethnicity as Hispanic. In Sample 2, 83.6% self-identified as White, 6.3% as Black, 4.3% as Asian, 1.1% as Native American, 1.7% endorsed “Biracial or Multiracial,” 0.9% endorsed “other,” and 2.0% of Sample 2 identified their ethnicity as Hispanic. In Sample 3, 67.3% self-identified as White, 8.8% as Black, 5.6% as Asian, 3.1% endorsed “Biracial or Multiracial,” 0.3% endorsed “other,” and 5.9% of Sample 3 identified their ethnicity as Hispanic. The top five medical conditions reported among those in Sample 2 were asthma (24.7%), diabetes (14.1%), heart problems and hypertension (12.7%), arthritis (5.2%), and back and spine problems (4.9%).

To allay concerns regarding data quality, and to ensure participant attentiveness, 139 participants who did not accurately respond to two out of three embedded catch questions were removed from Sample 1 (Oppenheimer, Meyvis, & Davidenko, 2009; Paolacci, Chandler, & Ipeirotis, 2010). Paolacci and Chandler’s (2014) recommendation for improving data quality by restricting MTurk worker approval ratings was used for data collection in Samples 2 and 3. This method has been shown to result in similar levels of data quality in comparison with using catch questions (Peer, Vosgerau, & Acquisti, 2014). Specifically, MTurk workers with approval ratings below 95% were not allowed to participate in data collection for Samples 2 and
3. Additionally, recruitment restrictions were put in place to ensure that (a) Sample 1 participants only completed the study one time and (b) no participant was a member of both Samples 1 and 2.

All study procedures were approved by the local institutional review board. Informed consent and self-report measures could be completed from any computer with Internet access. Data were collected using a secure online survey program. Questionnaires were presented in random order. In Study 2 and Study 3, participants were asked to indicate whether they currently had a physical health problem that had been diagnosed. Participants in Sample 2 indicated that they had a current medical problem, whereas participants in Sample 3 indicated that they did not. Participants were debriefed and paid $0.50 upon study completion, an amount consistent with similar questionnaire studies with MTurk workers (Buhrmester, Kwang, & Gosling, 2011).

**Measure**

As described, the 31-item MIHT (Longley et al., 2005) was used in the present study because it assesses for the four factors of health anxiety associated with contemporary cognitive–behavioral models (i.e., Affective, Cognitive, Perceptual, Behavioral). Each item is rated on a 5-point scale, based on how much participants believe each item pertains to them ranging from 1 (strongly disagree) to 5 (strongly agree). Higher scores on the MIHT scales indicate greater health anxiety, with a possible total score range from 31 to 155. Previous research has provided psychometric support for the MIHT, including evidence of internal consistency, retest reliability, convergent and discriminant validity (Longley et al., 2005), and structural support for both the correlated four-factor (Longley et al., 2005, Witthöft et al., 2015) and higher-order models (MacSwain et al., 2009; Stewart et al., 2008).

**Data Analytic Strategy**

The following four models were examined in all three of our samples. The first model was a correlated four-factor model that consisted of seven items with primary loadings on Factor I (cognitive: dysfunctional beliefs about health), eight items with primary loadings on Factor II (behavioral: avoidance, particularly reassurance seeking), nine items with primary loadings on Factor III (perceptual: hypervigilance to physical sensations), and seven items with primary loadings on Factor IV (affective: health-related fear). No secondary loadings were modeled, but the factors were allowed to intercorrelate. Model 2 was a one-factor model; each of the MIHT items loaded onto one factor. Model 3 was a second-order model in which the correlations among the first-order factors in Model 2 were removed and direct pathways were added from a second-order factor to the first-order factors. Model 4 was a bifactor model in which all 31 items were simultaneously loaded onto a general factor and each of their respective domain-specific factors. In the bifactor model, all factor covariances were fixed to zero (Brown, 2015).

A multiple-groups CFA framework was used to examine measurement invariance (Brown, 2015). Specifically, restrictive models were used to test for (a) configural invariance (equal form) and (b) metric/scalar invariance (equal factor loadings, indicator thresholds) between those with (n = 348) and without (n = 354) a self-reported medical problem. To test configural invariance, the adequacy of the MIHT factor structure was examined simultaneously in the two groups. To test metric/scalar invariance, factor loadings and indicator thresholds were constrained to equality. Because MIHT item responses result in ordered – categorical data, metric/scalar invariance were simultaneously modeled (e.g., Brown, 2015; Ebesutani et al., 2014; Olatunji et al., 2017).

Analysis of covariance was conducted using SPSS Statistics (Version 22; IBM Corp., 2013) to test the hypothesis that those with a self-reported medical problem would report significantly higher levels of health anxiety in comparison to those who did not report a current medical problem. Group served as the independent variable and the MIHT total score and subscale scores served as dependent variables in separate models. Gender and age served as covariates in all models to account for between-groups differences in these demographic variables.

All models were tested using Mplus 7.4 (Muthen & Muthen, 2015). Mean- and variance-adjusted weighted least squares (WLSMV) estimation was used to test all models because MIHT item responses are ordered categorical (Asparouhov, 2005). In comparison with other robust maximum likelihood estimation, WLSMV has been shown to be less biased and more accurate in estimating factor loadings, especially when sample sizes exceed 200 (Li, 2016). Three of the most recommended (Brown, 2015; Hu & Bentler, 1999) fit indices were used to evaluate the models: the Tucker–Lewis fit index (TLI), comparative fit index (CFI), and root mean square error of approximation (RMSEA). The following guidelines were used to evaluate fit: CFI and TLI should be close to .95, RMSEA should be close to .06, and the upper limit of the 90% RMSEA confidence interval (CI) should not exceed .10 (Hu & Bentler, 1999; Kline, 2011).

To evaluate model comparisons, the chi-square difference test was used (using the DIFFTEST function in Mplus; Muthen & Muthen, 2015). However, the chi square difference test is highly influenced by sample size, and thus, chi-square difference tests may indicate a significant difference when the magnitude of differences are actually trivial (Cheung & Rensvold, 2002). To address this issue, we also used alternative tests for comparing models (Brown, 2015; Kline, 2011; i.e., examining RMSEA 90% CIs and change.
in CFI ($\Delta$CFI). If models have overlapping 90% RMSEA CIs, differences in model fit are considered nonsignificant (Wang & Russell, 2005). $\Delta$CFI values of less than or equal to .002 (Meade, Johnson, & Braddy, 2008) or less than or equal to .01 (Cheung & Rensvold, 2002) may indicate functionally trivial differences in parameter estimates.

In addition to a basic examination of fit statistics, the following statistical indices were used to further evaluate the bifactor models (Dueber, 2016; Rodriguez et al., 2016). OmegaH ($\omega_H$) reflects the proportion of variance in MIHT scores attributable to a single general factor, whereas OmegaHS ($\omega_{HS}$) reflects the proportion of variance in scores attributable to each domain-specific factor after removing the variance due to the general factor. Explained common variance (ECV) is calculated by dividing variance explained by the general factor by variance explained by both general and specific factors, and thus, serves as a better index of the unidimensionality of a measure in comparison with $\omega_H$, which is best understood as a measure of general factor reliability. Item explained common variance (I-ECV) is an indicator of the amount of common variance for each MIHT item attributable to the general factor. Percentage of uncontaminated correlations (PUC) characterizes the percentage of MIHT item correlations contaminated by variance attributed to the general and domain-specific factors. Average relative parameter bias (ARPB) serves as an indicator of the bias across parameters if items are forced into a unidimensional, versus multidimensional, structure. Factor determinacy (FD), the correlation between factors and factors scores, serves as an indicator of the degree to which factor scores are of practical value and should be used in measurement models. Values should be at least .80 or higher, and values >.90 are preferred (Gorsuch, 1983).

**Results**

**Confirmatory Factor Analysis**

Fit statistics for all four models, in each sample, are presented in Table 1. The correlated four-factor model provided adequate fit in Samples 1 and 2, with RMSEA, CFI, and TLI close to specified guidelines. In sample 3, however, CFI and TLI were below specified guidelines (i.e., CFI = .894, TLI = .885). The correlated four-factor model provided a significantly better fit to the data than the one-factor model which evidenced poor fit across all three samples. The second-order model provided an almost identical fit in comparison with the correlated four-factor model. Although the chi square difference test indicated a significant difference between the correlated four-factor model and the higher-order model across all three groups, overlapping RMSEA 90% CIs and $\Delta$CFI values indicate

<p>| Table 1. Goodness-of-Fit Statistics for Tested Models. |</p>
<table>
<thead>
<tr>
<th>Model</th>
<th>$\chi^2$</th>
<th>df</th>
<th>$\Delta\chi^2$</th>
<th>RMSEA</th>
<th>LL</th>
<th>UL</th>
<th>CFI</th>
<th>TLI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample 1 (n = 824)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>One-factor</td>
<td>15334.412</td>
<td>434</td>
<td>2013.449$^a$</td>
<td>.204</td>
<td>.201</td>
<td>.207</td>
<td>.582</td>
<td>.552</td>
</tr>
<tr>
<td>Correlated four-factor</td>
<td>2163.885</td>
<td>428</td>
<td>32.352$^b$</td>
<td>.070</td>
<td>.067</td>
<td>.073</td>
<td>.951</td>
<td>.947</td>
</tr>
<tr>
<td>Second-order</td>
<td>2049.326</td>
<td>430</td>
<td>412.765$^c$</td>
<td>.068</td>
<td>.065</td>
<td>.071</td>
<td>.955</td>
<td>.951</td>
</tr>
<tr>
<td>Bifactor</td>
<td>1412.574</td>
<td>403</td>
<td>—</td>
<td>.055</td>
<td>.052</td>
<td>.058</td>
<td>.972</td>
<td>.967</td>
</tr>
<tr>
<td>Sample 2: Medical (n = 348)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>One-factor</td>
<td>3356.471</td>
<td>434</td>
<td>560.673$^a$</td>
<td>.139</td>
<td>.135</td>
<td>.144</td>
<td>.662</td>
<td>.638</td>
</tr>
<tr>
<td>Correlated four-factor</td>
<td>1162.559</td>
<td>428</td>
<td>6.039$^b$</td>
<td>.070</td>
<td>.065</td>
<td>.075</td>
<td>.915</td>
<td>.908</td>
</tr>
<tr>
<td>Second-order</td>
<td>1117.280</td>
<td>430</td>
<td>246.325$^c$</td>
<td>.068</td>
<td>.063</td>
<td>.073</td>
<td>.921</td>
<td>.914</td>
</tr>
<tr>
<td>Bifactor</td>
<td>799.682</td>
<td>403</td>
<td>—</td>
<td>.053</td>
<td>.048</td>
<td>.059</td>
<td>.954</td>
<td>.947</td>
</tr>
<tr>
<td>Sample 3: Nonmedical (n = 354)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>One-factor</td>
<td>4057.089</td>
<td>434</td>
<td>606.316$^a$</td>
<td>.154</td>
<td>.149</td>
<td>.158</td>
<td>.596</td>
<td>.567</td>
</tr>
<tr>
<td>Correlated four-factor</td>
<td>1376.221</td>
<td>428</td>
<td>19.886$^b$</td>
<td>.079</td>
<td>.074</td>
<td>.084</td>
<td>.894</td>
<td>.885</td>
</tr>
<tr>
<td>Second-order</td>
<td>1374.558</td>
<td>430</td>
<td>287.179$^c$</td>
<td>.079</td>
<td>.074</td>
<td>.083</td>
<td>.895</td>
<td>.886</td>
</tr>
<tr>
<td>Bifactor</td>
<td>959.340</td>
<td>403</td>
<td>—</td>
<td>.062</td>
<td>.057</td>
<td>.068</td>
<td>.938</td>
<td>.928</td>
</tr>
<tr>
<td>Measurement invariance</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Configural</td>
<td>1756.926</td>
<td>806</td>
<td>—</td>
<td>.058</td>
<td>.054</td>
<td>.062</td>
<td>.946</td>
<td>.938</td>
</tr>
<tr>
<td>Metric/scalar</td>
<td>1945.748</td>
<td>951</td>
<td>313.109$^d$</td>
<td>.055</td>
<td>.051</td>
<td>.058</td>
<td>.944</td>
<td>.945</td>
</tr>
</tbody>
</table>

Note. df = degrees of freedom; RMSEA = root mean square error of approximation; CI = confidence interval; LL = lower limit; UL = upper limit; CFI = comparative fit index; TLI = Tucker–Lewis fit index. Models computed using mean- and variance-adjusted weighted least squares estimation. $\Delta\chi^2$ computed using Mplus 7.4 DIFFTEST function.

$^a$$\Delta\chi^2$ comparing one-factor and correlated four-factor models. $^b$$\Delta\chi^2$ comparing second-order and correlated four-factor models. $^c$$\Delta\chi^2$ comparing second-order and bifactor models. $^d$$\Delta\chi^2$ comparing metric/scalar and configural models.
Table 2. Standardized Factor Loadings From Bifactor Model in All Three Samples.

<table>
<thead>
<tr>
<th>Item #</th>
<th>General factor</th>
<th>Behavioral</th>
<th>Cognitive</th>
<th>Perceptual</th>
<th>Affective</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>S1</td>
<td>S2</td>
<td>S3</td>
<td>S1</td>
<td>S2</td>
</tr>
<tr>
<td>1</td>
<td>.36</td>
<td>.17</td>
<td>.38</td>
<td>.33</td>
<td>.26</td>
</tr>
<tr>
<td>5</td>
<td>.21</td>
<td>.28</td>
<td>.23</td>
<td>.76</td>
<td>.68</td>
</tr>
<tr>
<td>14</td>
<td>.43</td>
<td>.56</td>
<td>.44</td>
<td>.75</td>
<td>.51</td>
</tr>
<tr>
<td>16</td>
<td>.41</td>
<td>.52</td>
<td>.29</td>
<td>.69</td>
<td>.68</td>
</tr>
<tr>
<td>18</td>
<td>.40</td>
<td>.49</td>
<td>.37</td>
<td>.54</td>
<td>.45</td>
</tr>
<tr>
<td>20</td>
<td>.61</td>
<td>.63</td>
<td>.56</td>
<td>.52</td>
<td>.33</td>
</tr>
<tr>
<td>25</td>
<td>.46</td>
<td>.47</td>
<td>.56</td>
<td>.46</td>
<td>.57</td>
</tr>
<tr>
<td>26</td>
<td>.44</td>
<td>.55</td>
<td>.46</td>
<td>.72</td>
<td>.52</td>
</tr>
<tr>
<td>2</td>
<td>.48</td>
<td>.44</td>
<td>.52</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>.52</td>
<td>.36</td>
<td>.58</td>
<td>.68</td>
<td>.57</td>
</tr>
<tr>
<td>8</td>
<td>.52</td>
<td>.64</td>
<td>.70</td>
<td>.74</td>
<td>.47</td>
</tr>
<tr>
<td>10</td>
<td>.49</td>
<td>.36</td>
<td>.60</td>
<td>.71</td>
<td>.62</td>
</tr>
<tr>
<td>15</td>
<td>.47</td>
<td>.42</td>
<td>.63</td>
<td>.75</td>
<td>.63</td>
</tr>
<tr>
<td>27</td>
<td>.50</td>
<td>.49</td>
<td>.52</td>
<td>.77</td>
<td>.70</td>
</tr>
<tr>
<td>31</td>
<td>.41</td>
<td>.43</td>
<td>.51</td>
<td>.77</td>
<td>.75</td>
</tr>
</tbody>
</table>

Note. S1 = Sample 1 (n = 824); S2 = Sample 2: Medical (n = 348); S3 = Sample 3: Nonmedical (n = 354). All factor loadings, except for those identified by *, were significant at p < .05.

*p > .05.

that the differences in these models, across all three samples, are trivial in magnitude. The bifactor model provided adequate fit to the data in all three samples. The significant Δχ², nonoverlapping RMSEA 90% CIs, and ΔCFI > .01 all supported the bifactor model as providing better fit to the data than the second-order model. See Table 2 for factor loadings from the bifactor model in each sample. With the exception of the items of the Affective domain of the MIHT, all items exhibited significant factor loadings (p < .05) on their domain-specific factors in all three samples. Item loadings on the domain-specific factors tended to be higher than on the general factor, especially the Perceptual factor, which was the only factor with items that did not exhibit significant loadings across all three samples on the general factor.

Bifactor Model Evaluation

Strong reliability was exhibited for both the general (ω ranging from .95 to .96) and four domain-specific factors (ωS ranging from .88 to .96) in all three samples. As can be seen in Table 3, the general factor accounted for a substantial proportion of variance in MIHT item scores (ωH values ranging from .62 to .71). After removing the variance due to the general factor, both Behavioral and Cognitive factors exhibited ωHS values greater than .50 in two of three samples (ωHS values ranging from .40 to .67). The Perceptual factor exhibited large ωHS values (ranging from .66 to .85) and the Affective factor exhibited small ωHS values (ranging from .10 to .21). Unidimensional item pools are suggested by ECV values >.85 (Stucky & Edelen, 2015). The ECV
values across all three samples in the present study (ranging from .36 to .46) indicate that the general factor accounts for the minority of variance, thus supporting a multidimensional, rather than unidimensional, conceptualization of the MIHT. A multidimensional model is further supported by I-ECV averages from .37 to .46. In fact, only five items in Samples 1 and 3, and only four items in Sample 2, reflected content that largely represents that of the general factor, rather than being domain specific (i.e., I-ECV > .85; Stucky & Edelen, 2015). Importantly, all of these items were from the Affective factor in all three samples. The redundancy of this factor with the general factor is further suggested by an examination of residual variances of the factors. The Affective factor was the only domain-specific factor to exhibit nonsignificant residual variance. This was observed in all three samples (ps ranging from .09 to .44). PUC values of .77, in combination with the described ECV values, indicate that the multidimensional nature of the MIHT does not necessarily disqualify the use of a total score (Reise, Bonifay, & Haviland, 2013). ARPB values ranged from .17 to .41, thus, further supporting multidimensionality and indicating that the MIHT domain-specific factors require modeling (Rodriguez et al., 2016). Finally, FD values for the general domain and three of four domain-specific factors (i.e., Behavioral, Cognitive, Perceptual) were between .87 and .96 across samples, thus suggesting adequate FD. In contrast, FD values of .77, .74, and .85 for the Affective factor increases concerns surrounding the use of this factor as a summed subscale score and as a latent variable in a SEM framework.

Comparing Those With and Without a Self-Reported Medical Problem

As described, the bifactor model exhibited adequate model fit among those with (Sample 2) and without (Sample 3) a self-reported medical problem. Although there was a significant \( \Delta \chi^2 \). RMSEA 90% CIs were overlapping and there was no decrement in CFI between the configural and metric/scalar invariance models, thus indicating that the more restrictive model (metric/scalar) did not evidence a significant degradation in model fit compared with the less restrictive model (configural). These results support the presence of measurement invariance between those with and without a self-reported medical problem.

As predicted, participants with a self-reported medical problem reported higher levels of health anxiety (MIHT total score: \( M = 99.36; SD = 17.49 \)) than those who did not report a current medical problem (\( M = 94.16, SD = 17.68 \)), \( F(1, 701) = 18.17, p < .001 \). The mean difference was small in magnitude (Cohen’s \( d = 0.30 \)). Additionally, participants with a self-reported medical problem scored significantly higher than those who did not report a current medical problem on the following three MIHT factors: Cognitive, \( M = 18.71, SD = 6.18 \) versus \( M = 17.19, SD = 6.18 \), \( F(1, 701) = 14.87, p < .001 \); Perceptual, \( M = 34.37, SD = 5.67 \) versus \( M = 33.15, SD = 6.01 \), \( F(1, 701) = 4.51, p = .03 \); and Affective, \( M = 21.49, SD = 5.88 \) versus \( M = 19.63, SD = 6.22 \), \( F(1, 701) = 14.49, p < .001 \). These group differences were small in magnitude (ds ranging from 0.21 to 0.31). The between-groups difference in scores on the Behavioral factor did not reach statistical significance, self-reported medical problem: \( M = 24.79, SD = 6.42 \) versus absence of a medical problem: \( M = 24.20, SD = 6.40 \), \( F(1, 701) = 3.35, p = .07 \).

Discussion

In the present study, we examined the factorial structure of the MIHT and measurement invariance based on current health status. Consistent with previous research (Longley et al., 2005; MacSwain et al., 2009; Stewart et al., 2008; Witthöft et al., 2015), we found that both the correlated four-factor and hierarchical models of the MIHT provided adequate fit to the data and neither model evidenced superior fit. The bifactor model, consisting of a general health anxiety factor and four domain-specific factors, provided the best fit across all three samples. Importantly, statistical indices derived from all three bifactor models presented a consistent picture of a self-report measure that is strongly multidimensional.

Although the general health anxiety factor accounted for significant variance, an overwhelming amount of variance was contributed by the domain-specific factors, thus...
supporting a multidimensional structure. ECV values were below .50 across samples, thus failing to support a robust univariate structure. In addition, the ARPB value, which indicates bias across parameters if items are forced into a unidimensional structure, was elevated above an acceptable range in each sample. The majority of reliable variance in three of the domain-specific factors (i.e., Cognitive, Behavioral, Perceptual) was independent of the general factor. Furthermore, nearly all of the MIHT items from these three subscales had I-ECV values indicating greater contribution of the items to the respective domain-specific factors than to the general health anxiety factor. The Affective domain is the notable exception. The same four items (i.e., 6, 7, 11, and 12), all from the Affective factor, reflected content that largely represented the general factor, rather than being domain specific (Stucky & Edelen, 2015). This finding is consistent with previous research examining a second-order model of the MIHT, which suggested the possibility that the Affective factor was redundant with the general factor (MacSwain et al., 2009; Stewart et al., 2008).

In the present study, examination of residual variances of the factors provided additional evidence of the redundancy of this factor; the Affective factor was the only domain-specific factor to exhibit nonsignificant residual variance, which is indicative of a weak factor in which variance is largely explained by the general factor. This effect was observed in all three samples. An examination of the items with nonsignificant I-ECV values from the Affective subscale suggests a common theme, broad health-related worry with potentially serious implications (e.g., “death,” “disease,” “something serious”). Whereas the other domains consist of content that is important, but perhaps more peripheral (e.g., reassurance seeking, hypervigilance for physical sensations, alienation due to disease conviction), the Affective domains appears to consist of content that is at the core of the health anxiety construct. This content-based explanation may aid in better understanding present and past results related to this specific domain.

Taken together, study findings indicate that the MIHT subscales, with the exception of the Affective subscale, are meaningfully independent from the general factor. As such, continued use of the MIHT subscales is warranted. When examining structural models that include the MIHT, researchers should consider using multidimensional latent variable model specifications that account for the general and domain-specific factors (e.g., bifactor models). However, it is advisable to consider removing the latent Affective domain when using the 31 items of the MIHT to model the factor structure of health anxiety, as the items of this domain appear to better represent the general health anxiety construct (Meyer & Brown, 2013; Reise, 2012).\(^1\) In addition, although the MIHT has a robust multidimensional factor structure, an examination of ECV values in combination with PUC values supports the use of a total score (Reise et al., 2013).

Results from an examination of measurement invariance suggest that a bifactor model of the MIHT is an accurate operationalization for those with or without a current self-reported medical condition. These results suggest that the MIHT total and subscale scores warrant use irrespective of the current health status of respondents. Differences between those with and without a self-reported medical condition on the MIHT total and subscale scores should be considered an accurate reflection of true group differences rather than indicating that observed differences are a function of item content that promotes differential responding. Consistent with previous research (e.g., Alberts et al., 2011; DeMarinis et al., 2009; Kehler & Hadjiistavropoulos, 2009; Tang et al., 2009), participants with a current self-reported medical problem reported significantly higher levels of health anxiety in comparison with those who reported the absence of a current medical problem. Differences in MIHT scores were only small in magnitude between the two groups. The small differences between groups suggests that the mean differences may not be conceptually important, other than to suggest that a current medical problem may intensify, to a small degree, one’s average level of health anxiety.

Study limitations must be acknowledged. The use of a large community sample in this study (a) is consistent with a dimensional conceptualization of health anxiety and (b) ensured that our analytic approach was significantly powered to detect effects. Additionally, research supports the use of MTurk for data collection (Chandler & Shapiro, 2016). Nevertheless, MTurk samples should not be considered representative of the general population (Paolacci & Chandler, 2014). For example, evidence suggests that MTurk samples tend to be younger and more highly educated than general population samples (Paolacci & Chandler, 2014). As such, replication of study results in both general population and clinical samples (e.g., somatic symptom disorder and illness anxiety; American Psychiatric Association, 2013), with more individuals scoring at the high end of the continuum of health anxiety, will help ensure that results generalize. Extant research has examined potential ethnoracial differences on alternative measures of health anxiety, with results suggesting ethnoracial invariance of those measures (Fergus, Kelley, & Griggs, 2016). Future samples with greater racial/ethnic diversity are needed to examine if a bifactor model of the MIHT evidences ethnoracial invariance. Finally, given the heterogeneous nature of our sample of participants with a current self-reported medical problem, it will be important to determine whether the factor structure of the MIHT is invariant across specific diagnoses (e.g., Alberts et al., 2011).

Despite these limitations, the present results provide evidence that a bifactor model most accurately represents the structure of the MIHT. The MIHT appears to be strongly multidimensional, with three of the four subscales providing substantive value. In addition, continued use of the total
score is recommended and medical status invariance is consistent with evidence of health anxiety as a dimensional construct, thus supporting use of the MIHT in both clinical and nonclinical samples.

Declaration of Conflicting Interests
The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding
The author(s) received no financial support for the research, authorship, and/or publication of this article.

Note
1. CFA was used to test the post hoc hypothesis that a modified bifactor model of the MIHT (i.e., with the Affective domain removed: one general factor and three domain-specific factors [cognitive, perceptual, and behavioral]) would provide adequate fit to the data. The modified model provided adequate fit to the data in all three samples: Sample 1: χ2 = 1459.539 (df = 410, p < .001), CFI = .971, TLI = .967, RMSEA = .056, 90% CI [.053, .059], Δχ2 = 65.732, p < .001; Sample 2: χ2 = 835.506 (df = 410, p < .001), CFI = .951, TLI = .944, RMSEA = .055, 90% CI [.049, .060], Δχ2 = 46.380, p < .001; Sample 3: χ2 = 960.919 (df = 410, p < .001), CFI = .939, TLI = .930, RMSEA = .062, 90% CI [.057, .062], Δχ2 = 20.561, p = .005. Although the Δχ2 was significant in all three samples, overlapping RMSEA 90% CIs, and ΔCFI values suggest that the modified bifactor model did not exhibit a significant decrement in model fit compared with the original MIHT bifactor model.

References


