Published in final edited form as:
J Pers Assess. 2016 ; 98(2): 200–208. doi:10.1080/00223891.2015.1070355.

Resolving Uncertainty About the Intolerance of Uncertainty Scale–12: Application of Modern Psychometric Strategies

Willie Hale, **Miranda Richmond**, **Janet Bennett**, **Tiffany Berzins**, **Alexander Fields**, **David Weber**, **Mark Beck**, and **Augustine Osman**
Department of Psychology, University of Texas at San Antonio

Abstract

In this study, we evaluated the factor structure, reliability estimates, item parameters, and differential correlates of the short form of the Intolerance of Uncertainty Scale (Carleton, Norton, & Asmundson, 2007) in samples of undergraduate women (n = 387) and men (n = 276) ranging in age from 18 to 49 years (M = 20.20, SD = 3.91). This instrument was designed to measure 2 facets of intolerance of uncertainty— prospective anxiety and inhibitory anxiety—although total scores on the measure are often used. A major objective of this study was to determine the degree to which derivation of total versus subscale scores is empirically permissible. Comparison of a bifactor model to a unidimensional model and a 2-factor correlated traits model indicated that the bifactor model exhibited superior fit to the sample data. This model provided evidence of a strong general intolerance of uncertainty factor that was more reliable and accounted for significantly more common variance than either subscale factor. Examination of the item response theory slope parameters revealed negligible bias in the measure's items across genders. Finally, a series of simultaneous regression analyses was conducted to examine differential correlates of the measure's total scale scores for men and women.

Ladouceur, Gosselin, and Dugas (2000) defined intolerance of uncertainty (IU) as "the predisposition to react negatively to an uncertain event or situation, independent of its probability of occurrence and of its associated consequences" (p. 934). The construct reflects "beliefs about the necessity of being certain, about the capacity to cope with unpredictable change, and about adequate functioning in situations that are inherently ambiguous" (Obsessive Compulsive Cognitions Working Group, 1997, p. 678). Freeston, Rhéaume, Letarte, Dugas, and Ladouceur (1994) developed the 27-item Intolerance of Uncertainty Scale (IUS–27) to assess IU. Several studies have shown that IU might exist as a transdiagnostic maintaining factor given its relationships with several psychological problems, including worry (Koemer & Dugas, 2008), generalized anxiety disorder (Behar, DiMarco, Hekler, Mohlman, & Staples, 2009), obsessive–compulsive disorder (Lind & Boschen, 2009), depression (van der Heiden et al., 2010), anxiety disorder (Carleton, Collimore, & Asmundson, 2010), panic disorder, and agoraphobia (Buhr & Dugas, 2009). Carleton, Norton, and Asmundson (2007) created a short version of this measure, the IUS–12, that consists of two factors: prospective and inhibitory anxiety. The purpose of this study

CONTACT: Willie Hale, willie.hale@utsa.edu, Department of Psychology, The University of Texas at San Antonio, 1 UTSA Circle, San Antonio, TX 78249.

was to examine the factorial structure, internal consistency estimates, item parameters, and potential correlates of the 12-item IUS.

Development of the IUS–12

As noted in Jacoby, F.abricant, Leonard, Riemann, and Abramowitz (2013), one of the major critiques of the original 27-item IUS in the extant literature is that factor analytic studies have yielded little consensus about the factorial nature of the construct. Studies using the 27-item version have reported two- (Carleton et al., 2007; Sexton & Dugas, 2009), four- (Buhr & Dugas, 2002), and five-factor solutions (Freeston et al., 1994), many of which contained factors that were difficult to interpret and consisted of items that were cross-loaded on multiple factors. As such, Carleton et al. (2007) sought to produce a shortened version of the scale that would exhibit superior factorial stability while maintaining the original measure's high reliability and construct validity. Using two samples of undergraduate students, they produced the 12-item IUS, which correlated strongly with the 27-item version ($r = .96$, $p < .001$). The resulting measure consisted of two factors. The first factor, prospective anxiety, reflects approach-oriented responses to uncertainty (Birrell, Meares, Wilkinson, & Freeston, 2011); its seven items assess individuals' desire for predictability, propensity for active information seeking to reduce uncertainty, and preference for knowing what future events entail. The second factor, inhibitory anxiety, reflects avoidance-oriented responses to uncertainty; its five items assess the degree to which individuals experience reticence and paralysis when faced with uncertainty. Carleton et al. (2007) reported a high correlation between the two factors ($r = .73$, $p < .001$), and suggested that either total scores or subscale scores for the individual factors could be computed, depending on the needs of the researcher. However, Carleton and colleagues did not provide empirical justification for deriving total scores on the IUS–12. Accordingly, one specific aim of this study was to assess the tenability and empirical foundation for this assertion, given that highly correlated factors provide insufficient evidence for computing total scores for multidimensional instruments (Reise, Moore, & Haviland, 2010).

The IUS–12's correlated two-factor structure has been replicated in several studies using diverse populations, and its construct validity has been established using a variety of related concurrent measures. Table 1 provides a summary of many of these studies. Carleton et al. (2007) used two samples of undergraduate students, whereas Fergus and Wu (2012) used multiple-group confirmatory factor analysis (CFA) to test for measurement invariance of the IUS–12 across African American and White groups. Helsen, Van den Bussche, Vlaeyen, and Goubert (2013) found the same factor structure as prior studies in two samples of Dutch undergraduate students. Carleton et al. (2010) examined the structure of the IUS–12 in a community sample, finding the two-factor structure fit well and that scores on the IUS–12 were related to a myriad of factors including generalized anxiety disorder (GAD), fear of negative evaluation, and social distress. The factor structure and internal consistency estimates of the IUS–12 scale scores have also been replicated in numerous clinical samples with diagnoses including GAD (Khawaja & Yu, 2010), obsessive–compulsive disorder (OCD; Jacoby et al., 2013), as well as anxiety and depression (McEvoy & Mahoney, 2011).

This study and plan for analysis

The purpose of this study was to address limitations inherent in prior studies involving the IUS–12 by applying several modern psychometric strategies and analyses. First, given that the IUS– 12 was originally conceptualized as a multidimensional instrument, but was determined to be scorable as separate subscales or as a single measure of intolerance of uncertainty (Carleton et al., 2007), we examined coefficient rho (ρ; Raykov, 2009) estimates for the IUS–12 total and subscale scores. To our knowledge, all studies examining the IUS–12 have reported the traditional Cronbach's coefficient alpha (α) as a measure of internal consistency (see Cronbach, 1951). However, numerous well-documented problems exist with coefficient α. For example, coefficient α is a function of the number of items in a scale and also holds impractical assumptions of uncorrelated residual variances, tau-equivalent indicators, or both, which can cause it to either over- or underestimate true reliability (Raykov & Marcoulides, 2015; Zinbarg, Revelle, Yovel, & Li, 2005).

Given the issues with coefficient α, Raykov (2009) proposed a latent variable modeling procedure that computes true score (composite) reliability, coefficient ρ, as the ratio of true score variance to total variance in a unidimensional instrument based on the factor loadings of the items. Being a model-based estimate of internal consistency, coefficient ρ is a more appropriate measure of internal consistency when using CFA, as have all previous studies examining the structure of the IUS–12. It was important to first examine these estimates for both the IUS–12's total and subscale scores given the substantive impact of a measure's reliability on analyses such as CFA. As with coefficient α, values for coefficient ρ greater than .70 were considered acceptable.

Second, based on (a) extremely high intercorrelations among scores on the subscales of the IUS–12 in the current and previous studies (e.g., Carleton et al., 2007), and (b) the IUS–12's authors' explicit recommendation to score the measure at either the subscale or total scale level, we examined the fit of a bifactor model of the IUS–12 using CFA. In bifactor modeling, which can be conducted within exploratory factor analysis (EFA), CFA, and item response theory (IRT) frameworks (Reise et al., 2010), each individual item is specified to load directly on both its specific subscale (or factor) and a general factor that is related to all items in a multidimensional measure. In addition, the specific factors are uncorrelated with each other, and the general factor is uncorrelated with the specific factors. The bifactor model has several advantages over its nested correlated-factor and higher order model brethren. Of particular importance to this investigation is that the bifactor model we considered allows the researcher to simultaneously examine the impact of the general and group-specific (subscale) factors on individual indicators, which can inform judgments about how to best score a multidimensional measure. That is, the degree to which items load on the orthogonal general versus group-specific factors can be used to examine psychometric properties such as coefficient omega (ω; McDonald, 1999) necessary for determining whether most of the variance in a multidimensional instrument is due to subscales that are truly specific (in which case, subscale scorses should be computed) or to a general factor underlying all items (in which case, total scores should be derived). As noted in Reise et al. (2010) neither second-order factors nor highly correlated first-order factors (as is the case for the IUS–12) provide sufficient evidence for the derivation of composite scores in multidimensional instruments.

We conducted confirmatory factor analysis via M*plus* version 7.2 software (Muthén & Muthén, 2012) using the robust weighted least squares (WLSMV) estimator to examine the relative fit of the hypothesized bifactor model against (a) the two-factor correlated model reported by Carleton et al. (2007), and (b) a unidimensional model in which all IUS–12 items were constrained to load onto a single factor. Within each model, we fixed the variance of all factors to 1 to determine the scale of the latent variables, and we did not permit correlated error terms in any of the models. The fit of each respective model was evaluated as acceptable using the following criteria: (a) comparative fit index (CFI) value greater than .95, (b) root mean square error of approximation (RMSEA) value lower than .06 (Hu & Bentler, 1999), and (c) weighted root mean square residual (WRMR) value lower than 1.0 (Yu & Muthén, 2002). Given that the three hypothesized models are nested (Reise et al., 2010), rescaled χ^2 difference tests were used to determine which model fit the data best. Additionally, given that the number of factors underlying responses to the IUS–12 represents a key component of this study, prior to conducting the analyses detailed earlier, we conducted a parallel analysis via Factor 9.3 software (Lorenza-Seva & Ferrando, 2015) to have comparative tests of dimensionality using the current sample.

Third, consistent with the bifactor modeling within CFA (bifactor CFA), we computed the omega hierarchical statistic to examine the amount of variance that is due to variation on a general uncertainty factor. In addition, we bolstered evidence for unidimensionality of the IUS–12 data by computing the expected common variance (ECV). In particular, the ECV is obtained as the sum of squared factor loadings for a general factor (G) divided by the sum of all squared factor loadings (general and specific-group factors). Higher ECV values are indicative of a general factor accounting for higher proportions of the total common variance (see Reise et al., 2010; ten Berge & Sočan, 2004).

Fourth, we examined the degree to which items in the IUS–12 perform similarly for males and females by testing for differential item functioning (DIF) using IRT. Although it is possible to examine measurement invariance between groups within a bifactor CFA model, we chose to assess invariance at the item level using this approach partly because no previous studies in the extant literature have examined the item parameters of the IUS–12. Examining whether differences in the item parameters exist would replicate and extend previous findings regarding the degree to which the IUS–12 items and factors have the same content meaning for men and for women, which is important for making valid comparisons of IUS–12 responses across specific groups (see Vandenberg & Lance, 2000).

Prior to examining the items to detect DIF, we first established (a) unidimensionality of the scale for men and women separately, and (b) configural invariance across both genders, as these represent assumptions of IRT analysis. The IRT–DIF analysis was conducted using the Bock–Aitkin EM algorithm (Bock & Aitkin, 1981) in the Item Response Theory Modeling for Patient Reported Outcomes 2.1 program (IRTPRO 2.1; Cai, du Toit, & Thissen, 2012). Detection of DIF was considered possible for items with DIF statistics (χ^2 c|a) that were significant at the $p < .001$ level, which was followed up by examination of the $S - \chi^2$ item fit statistics by gender for those items that were significant.

Finally, we performed several regression analyses to examine potential correlates of IUS–12 composite scores for men and women. Prior to conducting these analyses, we first conducted independent groups *t* tests on each of subscales and total scores (where applicable) of the IUS–12 and the concurrent measures described later and examined the effect sizes (Cohen's *d*) of any emergent gender differences. Effect sizes of .20, .50, and .80 were considered small, medium, and large, respectively (Cohen, 1977). We then examined the relationships between the IUS–12 and the three concurrent measures for the total sample and each gender using a series of simultaneous regression analysis. To guide interpretation of findings, we considered standardized coefficient (β) values of .10 or less to be weak correlates, whereas those between .10 and .19 were considered moderate and those greater than .20 were considered strong.

Method

Participants and procedure

To increase the heterogeneity of the sample, participants ($N = 663$) were recruited from various introductory psychology classes at a medium-sized university in the Midwest and a large public university in the Southwest. The combined sample (there were no statistically significant differences obtained for the demographic variables) included 276 men (*M* age = 20.37, *SD* = 3.86 years) and 387 women (*M* age = 20.07, *SD* = 3.94 years), ranging in age from 18 to 49 years (*M* age = 20.20, *SD* = 3.91). Regarding ethnicity, the sample was 48.70% Hispanic American, 28.20% White, 8.10% African American, 7.80% Asian American, and 7.10% other ethnicities.

Institutional review board (IRB) approval was obtained from each university prior to conducting the study. Participants were recruited through the psychology research pools, and each participant received course research credit. Data were collected in small structured groups of approximately 50 students during each session, with each session lasting for approximately 1 hr. All participants provided informed consent before completing the questionnaire packet.

Measures

In addition to a basic demographic information questionnaire, all participants completed the IUS–12, along with concurrent validation self-report measures of anxiety sensitivity, generalized anxiety, and various psychological symptoms.

IUS–12

The IUS–12 (Carleton et al., 2007) is comprised of 12 items that assess reactions to impending uncertainty, ambiguous situations, and the future. It consists of two factors: prospective anxiety (7 items; e.g., "I can't stand being taken by surprise") and inhibitory anxiety (5 items; e.g., "I must get away from all uncertain situations"). Each item is assessed using a 5-point Likert scale ranging from 1 (*not at all characteristic of me*) to 5 (*entirely* *characteristic of me*). Significant support was found for composite reliability of the total IUS–12 scores (ρ = .92, 95% CI [.91, .93]) and the Inhibitory (ρ = .87, 95% CI [.85, .89]) and Prospective (ρ = .87, 95% CI = [.86, .89]) subscale scores.

The Anxiety Sensitivity Index–3

The Anxiety Sensitivity Index–3 (Taylor et al., 2007) is an 18-item self-report questionnaire that assesses the tendency to fear anxiety-related symptoms resulting from the belief that such sensations could have harmful social, psychological, and physiological consequences. It is composed of three 6-item subscales: physical (e.g., "When my stomach is upset, I worry that I might be seriously ill"), cognitive (e.g., "When I cannot keep my mind on a task, I worry that I might be going crazy"), and social (e.g., "I worry that other people will notice my anxiety"). Participants are asked to endorse the extent of their agreement with each item on a 5-point scale ranging from 0 (*very little*) to 4 (*very much*). Subscale and total scores are calculated by summing relevant items. The internal consistency in the current sample was high for the total scale score (ρ = .91, 95% CI [.90, .92]) as well as the physical (ρ = .87, 95% CI [.85, .89]), cognitive (ρ = .89, 95% CI [.87, .91]), and social (ρ = .79, 95% CI [.76, .84]) sub-scale scores.

The Generalized Anxiety Disorder Questionnaire Version 4

The Generalized Anxiety Disorder Questionnaire Version 4 (GAD–Q–IV; Newman et al., 2002) is a 9-item measure designed to assess the presence of distressing, abnormally high, uncontrollable worry and related symptoms. It consists of five dichotomous (*yes*–*no*) items measuring incidences of worry, a checklist of six GAD symptoms that is scored dichotomously, one item assessing frequently worried about topics, and two questions assessing an individual's GAD-related distress and impairment scored on 9-point scales ranging from 0 (*none*) to 8 (*very severe*). This measure can be used as a screening tool for GAD or can be used to obtain continuous scores of GAD severity, provided all questions are answered, which was the case in this investigation.

The Symptom Assessment–45

The Symptom Assessment–45 (SA-45; Strategic Advantage, 1998) is a 45-item self-report instrument that consists of nine 5-item subscales assessing several types of psychopathology, including interpersonal sensitivity, phobic anxiety, somatization, and paranoid ideation. Each item is assessed using a 5-point scale ranging from 1 (*not at all*) to 5 (*extremely*). The internal consistency in this sample ranged from moderate for the Psychoticism subscale (ρ = .67, 95% CI [.60, .72]) to very high for the Interpersonal Sensitivity subscale (ρ = .88, 95% CI [.86, .89]).

Results

Confirmatory factor analysis

Table 2 provides the standardized factor loadings (upper portion) and fit statistics (lower portion) for each of the three models that were compared. Both the unidimensional model—robust $\chi^2(54, N = 663)$ *D* 464.52, $p < .001$, CFI = .96, RMSEA = .11, WRMR = 1.51—and the two-factor correlated model— robust $\chi^2(53, N = 663) = 308.78$, $p < .001$, CFI = .97, RMSEA = .09, and WRMR = 1.20—demonstrated moderate to somewhat poor fit to the sample data. Only the bifactor model— robust $\chi^2(43, N = 663) = 158.28$, $p < .001$, CFI = .99, RMSEA = .06, and WRMR = .80—met all preestablished criteria. Rescaled χ^2 difference tests indicated that the bifactor model fit the data significantly better than did the competing unidimensional model, robust $\Delta\chi^2(12) = 244.57$, $p < .001$, and the two-factor correlated traits model, robust $\Delta\chi^2(11) = 133.64$, $p < .001$. These results are consistent with the findings of the parallel analysis, which suggested extracting a single factor from the data.

Examination of the bifactor CFA model and omega hierarchical

Given that the bifactor model exhibited superior fit relative to the unidimensional and two-factor correlated models, we next evaluated the bifactor CFA model to determine the nature of the resulting general and specific factors. Table 2 (see columns under "Bifactor model") details the factor loadings, communalities, and uniqueness for each item in the bifactor solution (top portion), the percentage of the total and common variance accounted for by the general and specific factors, as well as the resulting coefficient ω values associated with the resulting factors (bottom portion). As can be seen in Table 2, all 12 items loaded higher on the general IU factor (with loadings ranging from .68–.80) than on their respective group-specific factors (with loadings ranging from .17–.56 for inhibitory IU, and from −.17–.66 for prospective IU), providing initial evidence of a strong general IU factor (Reise et al., 2010). Moreover, only one and two items loaded higher than .50 on the specific prospective IU and inhibitory IU factors, respectively, which is far fewer than the four needed to provide sufficient empirical support for computing group-specific factor scores (Reise et al., 2010). This result does not provide strong empirical support for computing scores for the IUS–12 group-specific factors.

Partitioning of the variance due to the factors in the bifactor CFA model provided further support for a strong general IU factor and against the specific inhibitory and prospective IU factors. Results indicated that the general IU factor accounted for substantial variance among the items, accounting for 47.40% and 80.30% of the total and shared variance, respectively, whereas the inhibitory IU and prospective IU factors accounted for 6.60% and 5.00% of the total variance and 11.20% and 8.50% of the shared variance, respectively.

Aside from the substantial difference in attributable variance, considerable differences in the reliability estimates associated with the factors also emerged. Omega hierarchical (ω_h) and omega subscale (ω_s) estimates (McDonald, 1999), which represent the appropriate measure of the reliability of the latent constructs in a bifactor model with the effects of other constructs removed (Reise et al., 2010) are also presented in Table 2 (lower portion). The general IU factor exhibited high reliability ($\omega_h = .88$), whereas the group-specific inhibitory IU ($\omega_s = .20$) and prospective IU ($\omega_s = .03$) factors did not. These results indicate that the group-specific factors, despite each accounting for a small proportion of the variance in IUS–12 scores, do not possess sufficient reliable variance for interpretation (Gignac & Watkins, 2013), whereas the general IU factor does. Accordingly, the preponderance of the evidence supports deriving only total scores on the IUS–12.

Differential item functioning

Based on the findings that the IUS–12 should best be scored as a unidimensional instrument, we next sought to determine the extent to which the items performed similarly across genders, using IRT to detect DIF in the items. Prior to examining the items for DIF, we first established (a) unidimensionality of the scale for men and women separately, and (b) configural invariance across both genders. We found that, after correlating the residual terms for Items 6 and 7 for both genders, the unidimensional baseline models exhibited adequate fit for both men (CFI = .93, RMSEA = .07, and SRMR = .06) and women (CFI = .96, RMSEA = .06, and SRMR = .04). Additionally we found evidence of configural invariance across genders (CFI = .95, RMSEA = .07, and SRMR = .04). Together, these results provided the permissibility to examine the IUS–12 for DIF.

The DIF statistics (χ^2 c|a) and their associated *p* values are presented in Table 3, alongside the item-level diagnostic statistics ($S - \chi^2$) and the slope (*a*) parameters for each gender. Using the guidelines proposed by Baker and Kim (2004), we found that for males, the *a* parameters ranged from moderate, $a = 0.86$ (Item 4), to high, $a = 1.47$ (Item 12). Similarly, for females, the *a* parameters ranged from moderate, $a = .86$ (Items 8 and 11), to high, $a = 1.36$ (Item 6). As can be seen in Table 3, the IRT–DIF analysis could only detect evidence of DIF for a single item (Item 10). Accordingly, we next examined the item-level diagnostic statistics for this item, but found that the item-level diagnostic statistics were nonsignificant for both gender groups, indicating satisfactory fit of this item for both males and females. Given these mixed findings, we examined the correlation between total scores of the full scale using all 12 items and scores using all items except Item 10. These scores were highly correlated for the total sample ($r = .99$), as well as for both males ($r = .99$) and females ($r = .99$). Accordingly, we concluded that any DIF detected in Item 10 was negligible and proceeded to test for gender differences using the total IUS–12 scores based on all 12 items.

Analyses of correlates of the IUS–12 for men and women

After demonstrating that a single common dimension (i.e., general IU) is most representative of the items in the IUS–12 and showing that the items are ostensibly invariant across gender, we decided to examine potential gender differences in correlates for the IUS–12 total scale scores. First, we conducted independent groups *t* tests on each of the subscales and total scores (where applicable) of the IUS–12 as well as the ASI–3, GAD– Q–IV, and the SA–45. Results of this analysis revealed that men and women differed significantly on the GAD–Q–IV total scale score ($M_{men} = 4.43$, $SD_{men} = 3.55$; $M_{women} = 5.41$, $SD_{women} = 3.44$), $t(661) = -3.58$, $p < .001$, Cohen's $d = .28$, consistent with prevalence rates of GAD in the general population. Gender differences were also found on the SA–45's Interpersonal Sensitivity ($M_{men} = 9.59$, $SD_{men} = 4.75$; $M_{women} = 10.76$, $SD_{women} = 5.40$), $t(661) = -2.88$, $p < .01$, Cohen's $d = .22$, and Obsessive–Compulsiveness ($M_{men} = 10.97$, $SD_{men} = 4.94$; $M_{women} = 11.75$, $SD_{women} = 4.91$), $t(661) = -1.99$, $p < .05$, Cohen's $d = .15$, subscale scores. However, the magnitude of the effect sizes associated with these differences can be considered trivial to small.

We next examined the relationship between scores on the IUS–12 and scores on the three concurrent measures for the total sample and each gender using a series of simultaneous regression analyses. Results are presented in Table 4. To guide interpretation of findings, we considered standardized coefficient (β) values of .10 or less to be weak correlates, whereas those between .10 and .19 were considered moderate, and those greater than .20 were considered strong.

For men, we found that the ASI–3's Cognitive Concerns, $\beta = .26$, $t(272) = 2.74$, $p < .01$, and Social Concerns, $\beta = .48$, $t\ (272) = 6.89$, $p < .001$, subscale scores exhibited strong significant relationships with scores on the IUS–12, whereas the Physical Concerns subscale score did not significantly predict IUS–12 scores when controlling for the other two subscale scores. This model accounted for approximately 41% (R^2) of the variance in IUS–12 scores. A second regression analysis using the GAD–Q–IV showed that scores on this measure significantly predicted IUS–12 scores, $\beta = .65$, $t(272) = 13.96$, $p < .001$, accounting for 41% (R^2) of the variance in IUS–12 scores. A third regression analysis using scores on all the sub-scales of the SA–45 to predict IUS–12 scores showed that only the Depression, $\beta = .18$, $t(272) = 2.45$, $p < .05$, Obsessive– Compulsiveness, $\beta = .33$, $t(272) = 4.64$, $p < .001$, and Paranoid Ideation, $\beta = .15$, $t(272) = 2.00$, $p < .05$, subscale scores predicted IUS–12 scores when all subscale scores were entered simultaneously. This model accounted for 46% (R^2) of the variance in IUS–12 scores.

For women, we found similar results for both the ASI–3 and the GAD–Q–IV scale scores. With respect to the ASI–3, the Physical Concerns subscale score did not predict IUS–12 scores, whereas the Cognitive Concerns, $\beta = .34$, $t(383) = 4.65$, $p < .001$, and Social Concerns, $\beta = .27$, $t(383) = 4.74$, $p < .001$, subscale scores predicted IUS–12 scores, accounting for 36% (R^2) of the variance in IUS–12 scores. A second regression analysis using the GAD–Q–IV total scale score showed that scores on this measure significantly predicted IUS–12 scores, $\beta = .65$, $t(383) = 17.03$, $p < .001$, accounting for 43% (R^2) of the variance in IUS–12 scores. A third regression analysis, using scores on all the subscales of the SA–45 to predict IUS–12 scores showed that only the Obsessive–Compulsiveness sub-scale, $\beta = .20$, $t(383) = 3.51$, $p < .001$, the Paranoid Ideation subscale, $\beta = .23$, $t(383) = 3.29$, $p < .001$, and the Anxiety (as opposed to the Depression subscale for men) subscale, $\beta = .20$, $t(383) = 3.20$, $p < .01$, scores predicted IUS–12 scores when all subscale scores were entered simultaneously. This model accounted for 51% (R^2) of the variance in IUS–12 scores. Despite the differential correlates observed for the third set of regressions, follow-up tests revealed no significant differences between males and females in the amount of variance accounted for in any of the three models we tested (i.e., the R^2 statistics were not significantly different across genders).

Discussion

IU has been established as a very important psychological construct in the extant literature. Significant and theoretically expected relationships have been found between IU and constructs spanning several types of psychopathology, indicating that IU might serve as a transdiagnostic factor. The IUS–12 (Carleton et al., 2007) has become increasingly used as a prominent measure of IU given its brevity (relative to the original measure) and demonstrable construct validity. That said, several psychometric issues have yet to be adequately addressed with respect to this measure. This study sought to address many of these issues by (a) comparing the fit of competing models for the IUS–12, (b) evaluating the specificity of the IUS–12's previously defined factors, (c) exploring the item parameters of the IUS–12 by gender, and (d) assessing the relationships between the IUS–12 and several concurrent measures.

With respect to the first two aims listed earlier, we compared the fit of a bifactor CFA model to a unidimensional model and to the correlated two-factor model identified by Carleton et al. (2007) and other researchers. Results indicated that the bifactor model fit the data better than both competing models, providing evidence of a strong general factor underlying all items. The general IU factor exhibited high reliability and accounted for nearly 50% of the total variance and 80% of the shared variance in IUS–12 scores. Given this evidence, combined with the fact (a) that few items loaded highly on the orthogonal group specific (subscale) factors, (b) that these factors accounted for relatively small proportions of both total and shared variance, and (c) that these factors exhibited poor reliability when removing the effects of the other factors, we are confident in making a few suggestions as to how the IUS–12 should be used in practice. First, we would suggest that future researchers always model the construct as a bifactor model in future studies examining the relationships between intolerance of uncertainty and theoretically related constructs. Second (and more practically), with respect to clinical usage, we believe that the bifactor model provides clear empirical support for deriving total scores on the measure. Given that value of over 80% of the common variance in IUS–12 scores, clinicians can rest assured that the total scores truly reflect the general factor. Finally, we would recommend avoiding scoring the subscales altogether. The strong support for the unidimensionality of the IUS–12 reported herein bolsters support for IU's standing as a transdiagnostic maintaining factor, given that unidimensional constructs are more likely than multidimensional constructs to exhibit invariance in form and function across individuals with different types of psychopathology.

Measurement invariance of the IUS–12 across gender and ethnicity has been established using CFA. However, consistent with the third aim listed earlier, we also examined the a parameters associated with the IUS–12 to determine whether items in the IUS–12 perform differently for males than for females. This was done because no studies in the existing literature have investigated the item parameters associated with this measure; examining whether the items in the IUS–12 exhibit DIF would replicate and extend previous CFA findings regarding the invariance of the measure. Mixed evidence of DIF was found for only one item (Item 10); however, this DIF was deemed negligible given near-perfect correlations between IUS–12 total scores and total scores with Item 10 removed. As such, researchers who attempt to obtain mean differences between genders using the measure need not concern themselves with bias in the items. That said, future investigations should examine whether item bias exists in the IUS–12 scores across other demographic variables such as age and ethnicity.

Finally, we assessed the relationships between total IUS–12 scores and scores on several different concurrent measures in a series of regression analyses, finding similar relationships as have other studies using the IUS–12 (e.g., Khawaja & Yu, 2010; McEvoy & Mahoney, 2011). The IUS–12 exhibited significant zero-order relationships with all three of the ASI–3's subscale scores, although only the social concerns and cognitive concerns subscale scores remained significant (for males, females, and the total sample) when entering them all simultaneously. The zero-order correlations observed herein were similar to prior studies (e.g., Carleton et al., 2010) examining the relationships between these two measures. Scores on the GAD–Q–IV and the IUS–12 exhibited extremely similar relationships with respect to the total samples in this (r = .64) and previous studies (r = .61; Carleton et al., 2007). Finally, we found that the IUS–12 total score was significantly related to all of the subscale scores of the SA–45 at the zero-order level, but that only two subscale scores (Obsessive–Compulsiveness and Paranoid Ideation) were significantly related across genders when all subscale scores were entered simultaneously into the equations. Furthermore, the regression analysis indicated that the SA–45 Anxiety and Depression subscale scores were differentially related to IUS–12 scores as a function of gender. The Anxiety subscale score was significantly related to IUS–12 scores for women but not men, and the opposite was true with respect to the Depression subscale scores. Future studies should investigate the cause of these differential relationships.

This study is not without its limitations, which must be considered to properly contextualize the findings reported here. First, this study was conducted solely with undergraduate students, and the nature of IU, and its relationships to the concurrent measures we examined, might differ in various ways for older adults, who theoretically are not dealing with issues of emerging adulthood. Additionally, it would be beneficial to attempt to replicate our findings with independent clinical samples. Second, this study did not attempt to examine other important properties of the IUS–12, including the stability of the bifactor structure over time. Finally, only self-report measures were used in this study and the cross-sectional and correlational design employed cannot be used to make causal claims. Further research using diverse samples and methodologies are needed to address these issues. These limitations notwithstanding, this study provided invaluable insight into the nature of IU as measured by the IUS–12. Many of the findings have important clinical and research implications that will prove fruitful for future investigations of the construct.

References

Baker, FB.; Kim, SH., editors. Item response theory: Parameter estimation techniques. 2nd. New York, NY: Marcel Dekker; 2004.

Behar E, DiMarco ID, Hekler EB, Mohlman J, Staples AM. Current theoretical models of generalized anxiety disorder (GAD): Conceptual review and treatment implications. Journal of Anxiety Disorders. 2009; 23:1011–1023. [PubMed: 19700258]

Birrell J, Meares K, Wilkinson A, Freeston M. Toward a definition of intolerance of uncertainty: A review of factor analytical studies of the Intolerance of Uncertainty Scale. Clinical Psychology Review. 2011; 31:1198–1208. [PubMed: 21871853]

Bock R, Aitkin M. Marginal maximum likelihood estimation of item parameters: Application of an EM algorithm. Psychometrika. 1981; 46:443–459.

Buhr K, Dugas MJ. The Intolerance of Uncertainty Scale: Psychometric properties of the English version. Behaviour Research and Therapy. 2002; 40:931–946. [PubMed: 12186356]

Buhr K, Dugas MJ. The role of fear of anxiety and intolerance of uncertainty in worry: An experimental manipulation. Behaviour Research and Therapy. 2009; 47:215–223. [PubMed: 19159867]

Cai, L.; du Toit, S.; Thissen, D. IRTPRO 2.1: Flexible, multidimensional, multiple categorical IRT modeling [Computer software]. Chicago, IL: Scientific Software International; 2012.

Carleton R, Collimore K, Asmundson G. "It's not just the judgments—It's that I don't know": Intolerance of uncertainty as a predictor of social anxiety. Journal of Anxiety Disorders. 2010; 24:189–195. [PubMed: 19931391]

Carleton R, Norton P, Asmundson G. Fearing the unknown. A short version of the Intolerance of Uncertainty Scale. Journal of Anxiety Disorders. 2007; 21:105–117. [PubMed: 16647833]

Cohen, J. Statistical power analysis for the behavioral sciences. New York, NY: Routledge; 1977.

Cronbach LJ. Coefficient alpha and the internal structure of a test. Psychometrika. 1951; 16:297–334.

Fergus T, Wu K. The Intolerance of Uncertainty Scale: Measurement invariance, population heterogeneity, and its relation with worry among self-identifying white and black respondents. Assessment. 2012; 20:555–564. [PubMed: 22984142]

Freeston M, Rhéaume J, Letarte H, Dugas M, Ladouceur R. Why do people worry? Personality and Individual Differences. 1994; 17:791–802.

Gignac G, Watkins M. Bifactor modeling and the estimation of model-based reliability in the WAIS-IV. Multivariate Behavioral Research. 2013; 48:639–662. [PubMed: 26741057]

Helsen K, Van den Bussche E, Vlaeyen JWS, Goubert L. Confirmatory factor analysis of the Dutch Intolerance of Uncertainty Scale: Comparison of the full and short version. Journal of Behavior Therapy and Experimental Psychiatry. 2013; 44:21–29. [PubMed: 22842528]

Hu L, Bentler P. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. Structural Equation Modeling. 1999; 6:1–55.

Jacoby R, Fabricant L, Leonard R, Riemann B, Abramowitz J. Just to be certain: Confirming the factor structure of the Intolerance of Uncertainty Scale in patients with obsessive-compulsive disorder. Journal of Anxiety Disorders. 2013; 27:535–543. [PubMed: 23973743]

Khawaja N, Yu L. A comparison of the 27-item and 12-item Intolerance of Uncertainty Scales. Clinical Psychologist. 2010; 14:97–106.

Koemer N, Dugas MJ. An investigation of appraisals in individuals vulnerable to excessive worry: The role of intolerance of uncertainty. Cognitive Therapy and Research. 2008; 32:619–638.

Ladouceur R, Gosselin F, Dugas M. Experimental manipulation of intolerance of uncertainty: A study of a theoretical model of worry. Behaviour Research and Therapy. 2000; 38:933–941. [PubMed: 10957827]

Lind C, Boschen M. Intolerance of uncertainty mediates the relationship between responsibility beliefs and compulsive checking. Journal of Anxiety Disorders. 2009; 23:1047–1052. [PubMed: 19656653]

Lorenza-Seva, U.; Ferrando, P. FACTOR (Version 9.3) [Software]. 2015. Retrieved from http://psico.fcep.urv.es/utilitats/factor/Download.html

McDonald, RP. Test theory: A unified treatment. Mahwah, NJ: Erlbaum; 1999.

McEvoy P, Mahoney A. Achieving certainty about the structure of intolerance of uncertainty in a treatment-seeking sample with anxiety and depression. Journal of Anxiety Disorders. 2011; 25:112–122. [PubMed: 20828984]

Muthen, L.; Muthen, B. Mplus user's guide. Los Angeles, CA: Muthen & Muthen; 2012.

Newman M, Zuellig A, Kachin K, Constantino M, Przeworski A, Erickson T, Cashman-McGrath L. Preliminary reliability and validity of the Generalized Anxiety Disorder Questionnaire-IV: A revised self-report diagnostic measure of generalized anxiety disorder. Behavior Therapy. 2002; 33:215–233.

Obsessive Compulsive Cognitions Working Group. Cognitive assessment of obsessive-compulsive disorder. Behaviour Research and Therapy. 1997; 35:667–681. [PubMed: 9193129]

Raykov T. Evaluation of scale reliability for unidimensional measures using latent variable modeling. Measurement and Evaluation in Counseling and Development. 2009; 42:223–232.

Raykov T, Marcoulides GA. A direct latent variable modeling based method for point and interval estimation of coefficient alpha. Educational and Psychological Measurement. 2015; 75(1):146–156.

Reise S, Moore T, Haviland M. Bifactor models and rotations: Exploring the extent to which multidimensional data yield univocal scale scores. Journal of Personality Assessment. 2010; 92:544–559. [PubMed: 20954056]

Sexton K, Dugas M. Defining distinct negative beliefs about uncertainty: Validating the factor structure of the Intolerance of Uncertainty Scale. Psychological Assessment. 2009; 21:176–186. [PubMed: 19485672]

Strategic Advantage. Symptom Assessment–45 questionnaire (SA– 45) technical manual. Toronto, ON, Canada: Multi-Health Systems; 1998.

Taylor S, Zvolensky M, Cox B, Deacon B, Heimberg R, Ledley D, Cardenas SJ. Robust dimensions of anxiety sensitivity: Development and initial validation of the Anxiety Sensitivity Index-3. Psychological Assessment. 2007; 19:176–188. [PubMed: 17563199]

ten Berge JM, Socan G. The greatest lower bound to the reliability of a test and the hypothesis of unidimensionality. Psychometrika. 2004; 69:613–625.

van der Heiden C, Melchoir K, Muris P, Bouwmeester S, Bos A, van der Molen H. A hierarchical model for the relationships between general and specific vulnerability factors and symptom levels of generalized anxiety disorder. Journal of Anxiety Disorders. 2010; 24:284–280. [PubMed: 20074909]

Vandenberg R, Lance C. A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. Organizational Research Methods. 2000; 3:4–70.

Yu, C.; Muthén, B. Evaluation of model fit indices for latent variable models with categorical and continuous outcomes. Paper presented at the annual meeting of the American Educational Research Association; New Orleans, LA. 2002 Apr.

Zinbarg RE, Revelle W, Yovel I, Li W. Cronbach's α, Revelle's *b* and McDonald's ωh: Their relations with each other and two alternative conceptualizations of reliability. Psychometrika. 2005; 70:1–11.

Table 1

Key internal consistency, structural, and concurrent validation analyses of the Intolerance of Uncertainty Scale–12.

Study/ Author	Sample	Concurrent validation measures	Internal consistency approach	Modeling approach
Carleton, Norton, and Asmundson (2007)	Undergrad students: Regina sample (*n* = 254); Houston sample (*n* = 818)	Beck Anxiety Inventory, Beck Depression Inventory, Generalized Anxiety Disorder Questionnaire–4, Penn State Worry Questionnaire	Regina sample: Total IU α = .91, PIU α = .85, IIU α = .85; Houston sample: Not reported	Separate CFAs for both samples
Carleton, Collimore, & Asmundson (2010)	Community sample (*N* = 286)	Anxiety Sensitivity Index–3, Brief Fear of Negative Evaluation Scale, Generalized Anxiety Disorder Assessment, Social Avoidance and Social Distress Scale, Social Interaction Phobia Scale, Positive and Negative Affect Schedule	Total IU α = .92, PIU α = .87, IIU α = .90	Regression analyses
Khawaja & Yu (2010)	Clinical (GAD) sample (*n* = 50); Nonclinical sample (*n* = 57)	Anxious Thoughts Inventory, Meta-Cognitions Questionnaire, Penn State Worry Questionnaire, State–Trait Anxiety Inventory	Clinical sample: Total IU α = .87, PIU α = .86, IIU α = .72; Nonclinical sample: Total IU α = .92, PIU α = .86, IIU α = .89	Comparison of reliability estimates across samples
McEvoy & Mahoney (2011)	Treatment seeking sample with anxiety and depression (*N* = 463)	Eyesenck Personality Questionnaire, Penn State Worry Questionnaire, Body Sensations and Agoraphobic Cognitions Questionnaires, Social Phobia Scale, Social Interaction Anxiety Scale, Padua Inventory, Beck Depression Inventory, Kessler Psychological Distress Scale	Total IU α = .93, PIU α = .88, IIU α = .88	CFA
Fergus & Wu (2012)	Undergraduate students: White (*n* = 1,185); African American (*n* = 301)	Penn State Worry Questionnaire	Not reported	Multiple groups CFA
Helsen et al. (2013)	Dutch undergraduate students: Calibration sample (*n* = 483); Validation sample (*n* = 483)	Penn State Worry Questionnaire, Beck Depression Inventory, State–Trait Anxiety Inventory	Total IU α = .83, PIU α = .78, IIU α = .72 (all based on total sample)	Separate CFAs for both samples
Jacoby et al. (2013)	*N* = 205 patients with OCD	Yale–Brown OCD Scale, Dimensional OCD Scale, Obsessive Beliefs Questionnaire–44, Beck Depression Inventory	Total IU α = .93, PIU α = .90, IIU α = .90	CFA

Note. CFA = confirmatory factor analysis; IU = intolerance of uncertainty; PIU = prospective intolerance of uncertainty; IIU = inhibitory intolerance of uncertainty; GAD = generalized anxiety disorder; OCD = obsessive–compulsive disorder.

Table 2

Standardized factor loadings, confirmatory factor analysis fit statistics of competing models for Intolerance of Uncertainty Scale–12, and variance partitioning based on bifactor model.

		Correlated factors model		Bifactor model				
Item	**Unidimensional model**	**Inhibitory**	**Prospective**	**General**	**Inhibitory**	**Prospective**	h^2	u^2
3. Uncertainty keeps me from living a full life.	.75	.78		.72	.24		.58	.42
6. When it is time to act, uncertainty paralyzes me.	.84	.86		.71	.55		.81	.19
7. When I am uncertain I can't function very well.	.82	.85		.70	.56		.80	.20
10. The smallest doubt can stop me from acting.	.75	.77		.68	.38		.61	.39
12. I must get away from all uncertain situations.	.81	.85		.80	.17		.67	.33
1. Unforeseen events upset me greatly.	.74		.76	.77		−.17	.59	.41
2. It frustrates me not having all the information I need.	.73		.75	.75		.02	.56	.44
4. One should always look ahead so as to avoid surprises.	.71		.73	.73		.05	.53	.47
5. A small unforeseen event can spoil everything, even with the best of planning.	.75		.77	.78		−.12	.61	.39
8. I always want to know what the future has in store for me.	.67		.68	.68		.07	.46	.54
9. I can't stand being taken by surprise.	.75		.77	.76		.06	.58	.42
11. I should be able to organize everything in advance.	.73		.75	.75		.66	.56	.44
WLSMV χ^2	464.52	308.78*			158.28*			
df	54	53			42			
CFI	0.96	0.97			0.99			
RMSEA [90% CI]	.11 [.10, .12]	.09 [.07, .10]			.06 [.05, .07]			
WRMR	1.51	1.20			0.80			
% Total variance (bifactor model)				47.40	6.60	5.00	Error = 41.00	
% Common variance (bifactor model)				80.30	11.20	8.50		

Item	Unidimensional model	Correlated factors model		Bifactor model				
		Inhibitory	**Prospective**	**General**	**Inhibitory**	**Prospective**	h^2	u^2
Coefficient omega				ω_h = .88	ω_s = .20	ω_s = .03		

Note. N = 663. h^2 = communality; u^2 = uniqueness; WLSMV = weighted least squares means and variance adjusted χ^2; *df* = degrees of freedom; CFI = robust comparative fit index; RMSEA = root mean square error of approximation; CI = confidence interval; WRMR = weighted root mean square residual; ω_h = omega hierarchical; ω_s = omega subscale.

*Statistically different (p < .001) from previous model. In the correlated factors model, the correlation between the factors was r = 88.

Table 3

Item response theory modeling: Differential item functioning analysis.

	Fit statistics									
	Males[a]				Females[b]				DIF statistics	
Item	***a* (*SE*)**	***S*–χ^2**	***df***	***p***	***a* (*SE*)**	***S*–χ^2**	***df***	***p***	**χ^2 c|a**	***p***
1	1.12 (.22)	75.24	61	.10	0.98 (.15)	97.03	78	.07	0.40	0.99
2	1.08 (.20)	92.59	71	.04	0.98 (.15)	98.28	80	.08	1.20	0.88
3	1.10 (.22)	73.15	57	.07	1.15 (.18)	108.4	80	.02	3.90	0.42
4	0.86 (.18)	85.44	81	.35	1.09 (.16)	90.15	84	.30	7.80	0.10
5	1.06 (.21)	92.41	73	.06	1.20 (.19)	99.78	76	.04	12.60	0.01
6	1.13 (.24)	57.15	56	.43	1.36 (.22)	88.79	63	.02	3.10	0.54
7	1.13 (.23)	62.81	59	.34	1.31 (.20)	90.85	76	.12	6.10	0.19
8	0.92 (.19)	100.6	77	.04	0.86 (.14)	104.98	94	.21	1.90	0.76
9	1.15 (.25)	63.77	62	.42	1.01 (.16)	89.96	79	.19	3.20	0.52
10	1.02 (.23)	77.8	63	.10	1.05 (.16)	95.67	83	.16	17.60	0.00
11	1.16 (.24)	76.93	69	.24	0.86 (.14)	113.44	91	.06	5.50	0.24
12	1.47 (.32)	45.72	47	.53	1.27 (.20)	71.03	65	.28	4.10	0.40

Note. DIF = differential item functioning; *a* = discrimination parameter; *SE* = standard error; $S - \chi^2$ = item fit statistics; *df* = degrees of freedom. Values shown in bold represent items where significant DIF was potentially detected.

[a] *n* = 276.

[b] *n* = 387.

Table 4

Simultaneous regression analysis: Correlates of the IUS–12 scale scores.

	IUS–12					
	Men[a]		Women[b]		Total sample	
Measure	***r***	**β**	***r***	**β**	***r***	**β**
ASI–3	$R^2 = .41^{**}$		$R^2 = .36^{**}$		$R^2 = .37$	
Physical concerns	$.48^{**}$	−.06	$.50^{**}$	.06	$.49^{**}$	.01
Cognitive concerns	$.56^{**}$	26^{**}	$.57^{**}$	$.34^{**}$	$.56^{**}$	$.31^{**}$
Social concerns	$.63^{**}$	$.48^{**}$	$.54^{**}$	$.27^{**}$	$.57^{**}$	$.35^{**}$
GAD–Q–IV	$R^2 = .41^{**}$		$R^2 = .43^{**}$		$R^2 = .41^{**}$	
Generalized anxiety	$.65^{**}$	$.65^{**}$	$.66^{**}$	$.65^{**}$	$.64^{**}$	$.64^{**}$
SA–45	$R^2 = .46^{**}$		$R^2 = .51^{**}$		$R^2 = .48^{**}$	
Anxiety	$.56^{**}$	.04	$.62^{**}$	$.20^{**}$	$.59^{**}$	$.15^{**}$
Depression	$.58^{**}$	$.18^{*}$	$.56^{**}$	.09	$.57^{**}$	$.12^{**}$
Hostility	$.32^{**}$	.00	$.41^{**}$	−.02	$.37^{**}$	−.01
Interpersonal sensitivity	$.60^{**}$	.10	$.63^{**}$	.09	$.61^{**}$	.09
Obsessive–compulsiveness	$.63^{**}$	$.33^{**}$	$.62^{**}$	$.20^{**}$	$.62^{**}$	$.25^{**}$
Paranoid ideation	$.54^{**}$	$.15^{*}$	$.64^{**}$	$.23^{**}$	$.60^{**}$	$.19^{**}$
Phobic anxiety	$.45^{**}$	.08	$.49^{**}$	.07	$.47^{**}$	$.07^{*}$
Psychoticism	$.31^{**}$	−.11	$.49^{**}$	.02	$.42^{**}$	−.03
Somatization	$.35^{**}$	.01	$.46^{**}$	−.03	$.42^{**}$	−.02

Note. IUS–12 = Intolerance of Uncertainty Scale-12; ASI–3 = Anxiety Sensitivity Index–3; GAD–Q–IV= Generalized Anxiety Disorder Questionnaire Version 4; SA–45 = Symptom Assessment–45.

[a] $n = 276$.

[b] $n = 387$.

* $p < .05$.

** $p < .01$.