Clinical Psychology Review 73 (2019) 101778

Contents lists available at ScienceDirect

Clinical Psychology Review

journal homepage: www.elsevier.com/locate/clinpsychrev

Review

The impact of methodological and measurement factors on transdiagnostic associations with intolerance of uncertainty: A meta-analysis

Peter M. McEvoy[a,b,*], Matthew P. Hyett[a], Sarah Shihata[a], Jordan E. Price[a,c], Laura Strachan[a]

[a] School of Psychology, Curtin University, Perth, Australia
[b] Centre for Clinical Interventions, Perth, Western Australia, Australia
[c] Virginia Commonwealth University, United States

HIGHLIGHTS

- Meta-analysed transdiagnostic associations with intolerance of uncertainty (IU)
- Examined moderators including clinical status, age, sex, IU measure, and symptom measure.
- Extracted 181 studies (N participants = 52,402) with 335 independent effects
- Moderate associations between IU and symptoms were observed across all disorders.
- Some moderating effects were observed, but these were not substantive.

ARTICLE INFO

Keywords:
Intolerance of uncertainty
Anxiety disorders
Obsessive compulsive disorder
Eating disorders
Review
Meta-analysis

ABSTRACT

Intolerance of uncertainty is a dispositional trait associated with a range of psychological disorders, but the influence of methodological factors on theses associations remains unknown. The first aim of this meta-analysis was to quantify the strengths of the association between IU and symptoms of generalised anxiety disorder, social anxiety disorder, panic disorder, agoraphobia, obsessive compulsive disorder, depression, and eating disorders. The second aim was to assess the influence of methodological factors on these relationships, including clinical (vs. non-clinical) status, age group, sex, IU measure, and symptom measure. We extracted 181 studies (*N* participants = 52,402) reporting 335 independent effect sizes (Pearson's *r*). Overall, there was a moderate association between IU and symptoms ($r = 0.51$, 95% CI = 0.50–0.52), although heterogeneity was high ($I^2 = 83.50$, $p < .001$). Some small but significant moderator effects emerged between and within disorders. Effect sizes were not impacted by sample size. The results indicate that IU has robust, moderate associations with a range of disorder symptoms, providing definitive evidence for the transdiagnostic nature of IU.

1. Inroduction

Intolerance of uncertainty (IU) refers to "an individual's dispositional incapacity to endure the aversive response triggered by the perceived absence of salient, key, or sufficient information, and sustained by the associated perception of uncertainty" (Carleton, 2016a, p. 31). IU is a dispositional trait that manifests as negative beliefs and reactions to unpredictable events regardless of their probability, and accounts for cognitive, affective, and behavioural responses to uncertainty (Dugas & Robichaud, 2007; Freeston, Rheaume, Letarte, Dugas, & Ladouceur, 1994; Ladouceur, Gosselin, & Dugas, 2000). Individuals with high IU tend to perceive uncertainty as negative, threatening, and as something to be avoided (Carleton, Norton, & Asmundson, 2007; McEvoy & Mahoney, 2012). Research suggests that IU may impact behaviour and decision-making, and is linked to elevated levels of physiological arousal as well as altered neural responses (Jacoby, Abramowitz, Reuman, & Blakey, 2016; Tanovic, Gee, & Joormann, 2018). Carleton (2016b) posited that IU reflects a fundamental fear of the unknown, which in turn underscores anxiety and neuroticism. The cognitive and behavioural features of anxiety disorders, such as maladaptive or irrational belief systems and associated engagement in safety behaviours, may be driven by a desire to increase control and certainty, thereby alleviating uncertainty and associated distress (Boswell, Thompson-Hollands, Farchione, & Barlow, 2013; Krohne, 1989).

Barlow's (2002) triple vulnerability model suggests that a general

* Corresponding author at: School of Psychology, Curtin University, GPO Box U1987, Perth, Western Australia, 6845, Australia.
E-mail address: peter.mcevoy@curtin.edu.au (P.M. McEvoy).

https://doi.org/10.1016/j.cpr.2019.101778
Received 24 March 2019; Received in revised form 4 August 2019; Accepted 20 August 2019
Available online 31 October 2019
0272-7358/ © 2019 Elsevier Ltd. All rights reserved.

biological vulnerability (e.g., genetics, hormones), general psychological vulnerability (e.g., unpredictable and uncontrollable early environments), and specific psychological vulnerability for particular disorders (e.g., fear of negative evaluation in social anxiety disorder) contribute to the development of emotional disorders. Hong and Cheung (2015) suggested that IU is related to multiple cognitive vulnerability factors (e.g., anxiety sensitivity, fear of negative evaluation, ruminative style), and reflects a common core fear of the unknown. Hong and Cheung (2015) assert that this shared theme of IU underlying most cognitive vulnerabilities is somewhat akin to the generalised psychological vulnerability (i.e., a sense of uncontrollability over internal and external experiences) postulated in the triple vulnerability model (Barlow, 2002). As a general vulnerability factor, IU may account for the high comorbidity observed across disorders and may heighten vulnerability to multiple psychological disorders (Boswell et al., 2013).

The Intolerance of Uncertainty Model (Dugas, Gagnon, Ladouceur, & Freeston, 1998) was initially developed to explain pathological worry, a core feature of generalised anxiety disorder (GAD), and there is now a large body of evidence demonstrating robust associations between IU and GAD symptoms (Buhr & Dugas, 2006; Dugas, Schwartz, & Francis, 2004). IU has also been found to be associated with symptoms of social anxiety disorder (Boelen & Reijntjes, 2009; Carleton, Collimore, & Asmundson, 2007; Whiting et al., 2014), panic disorder and agoraphobia (Carleton, Fetzner, Hackl, & McEvoy, 2013; Mahoney & McEvoy, 2012a, 2012b; McEvoy & Mahoney, 2011), health anxiety (Boelen & Carleton, 2012; Fetzner et al., 2014; Wright, Lebell, & Carleton, 2016), and posttraumatic stress disorder (Banducci, Bujarski, Bonn-Miller, Patel, & Connolly, 2016; Bardeen, Fergus, & Wu, 2012; Boelen, Reijntjes, & Smid, 2016; Oglesby, Boffa, Short, Raines, & Schmidt, 2016). There are established links between IU and obsessive-compulsive and related disorders including obsessive compulsive disorder (OCD; Holaway, Heimberg, & Coles, 2006; Jacoby, Fabricant, Leonard, Riemann, & Abramowitz, 2013; Tolin, Abramowitz, Brigidi, & Foa, 2003), hoarding disorder (Castriotta, Dozier, Taylor, Mayes, & Ayers, 2019; Oglesby et al., 2013), and body dysmorphic disorder (Summers, Matheny, Sarawgi, & Cougle, 2016). IU is also associated with major depressive disorder, which is frequently comorbid with anxiety disorders (Cuijpers et al., 2013; Norton & Price, 2007; Olatunji, Cisler, & Tolin, 2010). However, the association between IU and depression symptoms has not been consistently found and, therefore, some researchers posit that the association may be better accounted for by the frequent co-occurrence between depression and anxiety disorders (Jensen, Cohen, Mennin, Fresco, & Heimberg, 2016; Yook, Kim, Suh, & Lee, 2010). Emerging evidence also supports the relevance of IU to eating disorders (Brown et al., 2017; Kesby, Maguire, Brownlow, & Grisham, 2017; Renjan, McEvoy, Handley, & Fursland, 2016; Sternheim, Startup, & Schmidt, 2011). Moreover, different treatment protocols targeting various anxiety disorders and depression are associated with reductions in IU (Boswell et al., 2013; Dugas & Robichaud, 2007; McEvoy & Erceg-Hurn, 2016; van der Heiden, Muris, & van der Molen, 2012). Taken together, the large body of research is consistent with IU being a transdiagnostic risk and maintaining factor and change mechanism.

Extensive research has been conducted on IU across diagnostic groups, clinical and non-clinical populations, and sexes, as well as child, adolescent, and adult samples. For instance, research suggests that IU is higher in clinical samples relative to non-clinical samples (de Jong-Meyer, Beck, & Riede, 2009; Holaway et al., 2006). However, IU is conceptualised as a dimensional construct whereby the strengths of association between IU and disorder symptoms are comparable across clinical and non-clinical samples (Carleton, 2016b; Carleton et al., 2012). Moreover, although a higher level of IU may translate to a stronger influence of IU on symptoms in clinical samples, it is plausible that range restriction in clinical samples actually attenuates the strength of association compared to unselected samples or non-clinical samples with a larger range of scores (Kennet-Cohen, Kleper, & Turvall, 2018).

The strengths of the association between IU and disorder symptoms may also systematically differ across studies due to the use of different self-report measures to assess IU and disorder-specific symptoms. One of the first IU measures was the 27-item IU Scale (IUS; Freeston et al., 1994), which was originally developed within the context of worry and GAD. The IUS has been found to have good psychometric properties (Freeston et al., 1994), but its underlying factor structure has been found to be complex and unstable, with potential item redundancy (Norton, 2005). In addition, the IUS has been criticized for including items that are more closely linked to symptoms of GAD than other disorders (e.g., *My mind can't be relaxed if I don't know what will happen tomorrow*), thus potentially confounding the measurement of IU with GAD symptoms (Gentes & Ruscio, 2011). Items were subsequently removed from the 27-item version to improve the stability of the measure, resulting in a 12-item short form of the IUS (IUS-12; Carleton, Norton, & Asmundson, 2007). The IUS-12 is highly correlated with the original IUS, has demonstrated high internal consistency and construct validity, and is considered to be less specific to GAD (Carleton, Norton, & Asmundson, 2007; McEvoy & Mahoney, 2011; Norton, 2005).

The IUS-12 comprises two factors; prospective IU (i.e., cognitive appraisals about uncertainty) and inhibitory IU (i.e., behavioural inhibition in the face of uncertainty, McEvoy & Mahoney, 2011). However, recent bifactor analyses have found that in clinical and non-clinical samples the IUS-12 is best represented by a general factor, which explains most of the reliable variance across the items (Hale et al., 2016; Lauriola et al., 2018; Shihata, McEvoy, & Mullan, 2018). These findings indicate that a total score is an appropriate index of IU. The Obsessive Beliefs Questionnaire (OBQ; Obesessive Compulsive Cognitions Working Group (OCCWG), 2003, 2005) incorporates items that assess IU in relation to OCD, whereas the 45-item IU Inventory (IUI; Carleton, Gosselin, & Asmundson, 2010) was designed to assess general IU as well as cognitive and behavioural responses to uncertainty (e.g., avoidance, worry, reassurance seeking). Other adult measures have been used less frequently to measure situation- or disorder-specific IU (i.e., IU in disorder-relevant contexts; IUS situation-specific, Mahoney & McEvoy, 2012a; disorder-specific IU, Thibodeau et al., 2015).

IU measures have also been developed for use in child and adolescent samples. A child version of the IUS has demonstrated good psychometric properties (IUSC; Comer et al., 2009). As for the adult version, research on an abbreviated 12-item IUSC supports a strong general factor (Cornacchio et al., 2018). Thus, although the content of the measures focuses on uncertainty, there has been an evolution in the development of measures with some designed to assess IU in relation to a given disorder (OBQ, IUS), transdiagnostically (IUS-12), within specific situational contexts or diagnostic symptom categories (Disorder Specific IU), and for use in different populations (i.e., IUS-C for children vs. IUS-12 for adults).

The degree to which methodological and measurement issues influence the strengths of the association between IU and symptoms transdiagnostically remains unknown, which leads to uncertainty about the relative role of IU across emotional disorders. Differences in effects across studies may reflect genuine differences between sampled populations, or rather methodological and measurement differences across studies. Understanding the influence of these factors will help to guide theory, research methods, and measure selection in research and clinical contexts. For example, if IU is more strongly associated with symptoms of some disorders than others, and this pattern is found consistently across different measures of IU and symptoms, this would increase confidence that these relationships are not simply an artefact of method variance and it would be important for theorists to refine models to consider a commensurate influence of IU on symptoms. It would be also particularly important for clinicians to consider IU in their case formulations for disorders demonstrating robust strong

associations compared to disorders with weaker associations. On the other hand, if relationships with IU do not significantly differ across syndromes, regardless of the IU or symptom measure used in a study, this would provide compelling evidence that IU is a transdiagnostic factor that can be reliably estimated regardless of the specific measures. IU and outcome measures could then be selected by clinicians and researchers on pragmatic (e.g., brevity to reduce respondent burden) rather than historical grounds (i.e., the disorder-specific literature from which the measure derived). Patterns of association in children and adults would also provide important information about the continuity of these associations over time, with similar associations potentially indicative of IU having a similar relationship with symptoms, whereas different associations over time might reflect a maturation process that would need to be considered in etiological models. Similar associations between non-clinical and clinical associations would also be indicative of IU as a vulnerability factor across the dimensions of IU and symptoms. Robust differential associations across sexes may also help to explain differences in disorder prevalence.

Gentes and Ruscio (2011) conducted a meta-analysis comparing associations between symptoms of GAD, major depressive disorder, and OCD with IU measures developed for GAD (i.e., IUS, 56 effect sizes) or OCD (i.e., OBQ, 29 effect sizes). Compared to the OBQ, this study revealed a slightly stronger association between the IUS and symptoms of major depression (0.53 vs. 0.46) and OCD (0.50 vs. 0.42). The IUS was moderately associated with GAD symptoms (0.57), but no studies were identified reporting associations between the OBQ and GAD symptoms. It is unknown how strengths of association between the IUS and OBQ differ across other symptoms of psychopathology or other commonly used IU measures.

The current meta-analysis extends Gentes and Ruscio's (2011) meta-analysis in several ways. First, we provide an updated review that includes the large number of studies conducted beyond August 2010 through to December 2018. Second, included disorders and syndromes have been extended from GAD, major depression, and OCD to include social anxiety disorder, panic disorder and agoraphobia, and eating disorders. Third, we included a broader array of IU measures and versions for both adults and children. Fourth, in addition to syndromes, IU measures, and sample type (i.e., clinical vs. non-clinical), we investigated the impact of age (child vs. adult, average sample age), different symptom measures within syndromes, and sex as potential moderators. Fifth, we compared associations with IU between commonly used narrow (i.e., worry) versus broad symptom measures for GAD, and between a cognitive vulnerability factor (e.g., Anxiety Sensitivity Index) and symptom measures (e.g., Panic Disorder Severity Scale) for panic disorder. These analyses were intended to inform the field about the impact of using these alternative measures on the strengths of association with IU.

The first aim of this study was therefore to use meta-analytic methods to provide a definitive treatise on transdiagnostic associations with IU. Specifically, the first aim was to quantify the strengths of association between IU and symptoms of depression, social anxiety disorder, panic disorder, agoraphobia, GAD, OCD, and eating disorders. The second aim was to provide a quantitative review of the influence of methodological and measurement factors on these associations. The effects were first estimated across all studies, followed by estimates by syndrome (defined by symptom measure), clinical (vs. non-clinical) status, age group (child vs. adult, average sample age), sex, IU measure, and symptom measure. The influence of study sample size was also examined.

2. Methods

2.1. Protocol and registration

The protocol for this study was registered with the International Prospective Register of Systematic Reviews (PROSPERO) on 20th April 2018 (registration number CRD42018092293). Additional self-report measures of IU and of differing syndromes to those specified in the protocol were identified during the literature searches, and thus the final list of main outcome measures for analysis is more extensive than detailed in PROSPERO. The meta-analysis is reported according to the recommendations of the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) Statement (Moher, Liberati, Tetzlaff, Altman, and The PRISMA Group, 2009).

2.2. Literature search

We used PsycINFO, PubMed, and EMBASE databases to search for relevant articles published between 1960 and April 2018 using combinations of the following terms (see Supplementary Material A for full search strategy): *intolerance of uncertainty, IUS, IUI, (in)tolerance for ambiguity, (in)tolerance of ambiguity, need for certainty, need for predictability, need for cognitive closure, mood disorder, depression, major depression, MDD, MDE, anxiety, fear, generalised anxiety, GAD, worry, rumination, obsessive compulsive disorder, OCD, obsession, panic disorder, panic, agoraphobia, social anxiety disorder, social phobia, SAD, fear of negative evaluation, fear of positive evaluation, eating disorder, bulimia, anorexia, EDNOS, binge, purging, dietary restraint, restrictive eating, other specified feeding and eating disorders, OSFED, avoidant restrictive food intake disorder, ARFID*. Consistent with Gentes and Ruscio's (2011) search strategy, we initially included related but distinct constructs that are commonly investigated with IU (e.g., intolerance for ambiguity, need for cognitive closure) to minimise the chances of missing relevant papers. However, the final selection criteria included only IU studies (see Study Selection section). Additional articles were identified through subject heading searches (mapped to each keyword above) across each database. References lists of review papers were examined for additional relevant papers. Searches were limited to English language journal articles or reviews published in peer-reviewed journals, as well as book chapters, and dissertations. A total of 1053 unique manuscripts were initially identified, which were then reviewed by the authors for inclusion or exclusion. An additional 23 manuscripts were identified from the reference lists of review articles, and subsequently reviewed by the authors for inclusion/exclusion.

2.3. Study selection

The abstracts of all publications were screened for mention of IU or intolerance of ambiguity (IA) and GAD, major depressive disorder (MDD), social anxiety disorder (SAD), OCD, panic/agoraphobia, and/or disordered eating (ED). A total of 445 papers were identified as potentially relevant through screening of the abstracts, and the full-texts of these were subsequently reviewed by the authors for inclusion in the final analysis. Papers were considered for inclusion if they measured/reported the following: (a) assessed IU using general (e.g., IUS) or disorder-specific (e.g., OBQ) measures, (b) measured symptoms of GAD, MDD, SAD, OCD, panic/agoraphobia (PD/Ag), and/or disordered eating using validated self-report questionnaires, and (c) reported bivariate correlational or between-groups data (e.g., means and standard deviations of symptom measures for high vs. low intolerance of uncertainty groups) indexing relationships between constructs. Studies examining associations with the IA construct in isolation (i.e., that didn't include a valid measure of IU) were excluded given IA only refers to ambiguity in the present rather than explicitly to uncertain future events, which is the focus of this analysis (Grenier, Barrette, & Ladouceur, 2005). Of the 445 full-text studies identified for review a total of 264 studies were excluded. As detailed in Fig. 1, studies were excluded due to insufficient data to calculate an effect size ($n = 142$), absence of a relevant symptom ($n = 52$) or IU ($n = 50$) measure, sample sizes < 20 ($n = 8$), or because they used an ineligible sample (i.e., a syndrome not included in the current analysis) or separated groups of those high/low on IU through experimental manipulation

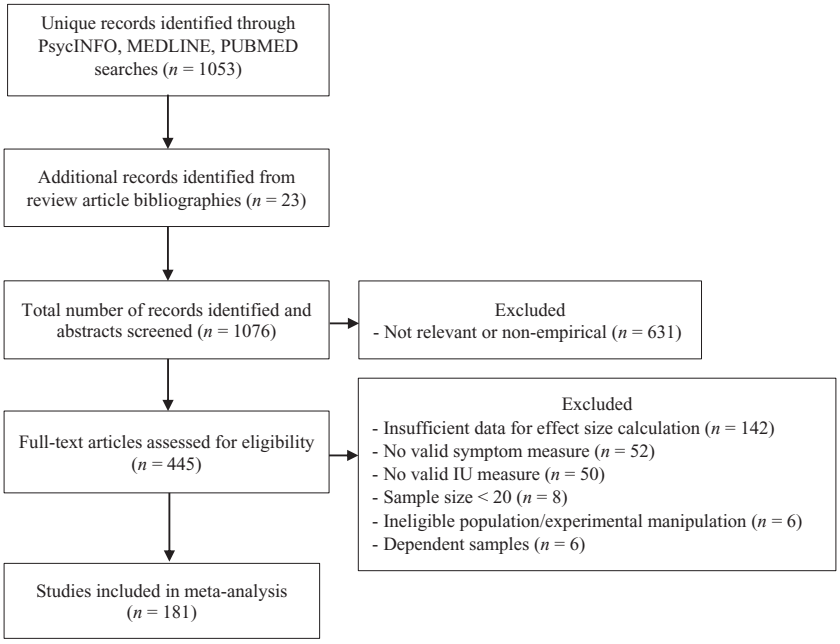

Fig. 1. Flow chart for inclusion and exclusion of studies in meta-analysis.

(n = 6). Authors of 92 studies (of the 445 assessed) that reported IU and symptom measures but not minimum data were contacted via email between September and December 2018 requesting correlational data, which was provided for 33 additional papers. All potentially eligible papers were checked for independence. Six papers reported data from dependent samples and were removed. A further eight effect sizes were duplicated across papers, but these studies reported at least one independent effect so they remained in our study. In total, data from 181 individual papers (N independent effects = 335, Total N participants = 52,402) provided minimum data for inclusion in the current meta-analysis. We conceptualised IU as the predictor variable and symptom measures as dependent variables. Using a pre-planned coding pro forma (see Supplementary Material B), three trained coders (authors MPH, JP, LS) extracted the data.

2.4. Eligible symptom measures

Self-report symptom-based measures of MDD, GAD, OCD, SAD, PD/Ag, or ED were considered eligible for inclusion. Table 1 in Supplementary Material C details all studies used in the meta-analysis. Measures of cognitive vulnerability factors (e.g., meta-cognitive beliefs, fear of negative evaluation) were excluded as these are separable from symptom measures. The one exception to this was the Anxiety Sensitivity Index (ASI). The ASI has been commonly used as a dependent variable in panic disorder studies, so we considered it important to compare associations between the ASI and panic disorder symptom measures. Furthermore, the Penn State Worry Questionnaire (PSWQ) has frequently been used in GAD studies, despite it only narrowly assessing one core feature of the disorder (worry). It is therefore important to determine whether estimates based on the PSWQ are similar to measures that assess broader diagnostic symptoms of GAD (e.g., Generalised Anxiety Disorder-7, GAD-7). These decisions were made to inform researchers about the impact of these commonly measures on the strengths of association. Multiple disorder symptom effects were extracted from studies not reporting on a specific clinical condition (e.g., measures of both MDD and SAD were included if reported from non-clinical or mixed clinical samples). In contrast, only condition specific symptom data were drawn from studies reporting on a single clinical syndrome (e.g., GAD-Q-IV for GAD).

2.5. Effect size coding

Effect sizes were calculated for included studies providing, i) correlations between IU symptom measures and sample size, or ii) between-groups (i.e., mean difference) effect sizes, calculated from the means and standard deviations of symptom measures for high and low IU groups (as determined by cut-off on the IUS). The latter mean difference (d) effect sizes were converted to a correlation (r) using:

$$r = \frac{d}{\sqrt{d^2 + a}}$$

where a is a correction factor for cases where $n_1 \neq n_2$ (when sample sizes are not equal, the correction factor depends on the ratio of n_1 to n_2). Effect sizes for individual studies, and for aggregating across studies, were computed using the common index, r. To assess interrater reliability 20% of publications initially judged as included/excluded were independently coded by another author. A high level of reliability between coders (91.3%) was observed. Disagreements were resolved by discussion and the final coding reflects the consensus between coders.

2.6. Data analysis

2.6.1. Meta-analytic model

Comprehensive Meta-Analysis (CMA) version 3.1 (Borenstein, Hedges, Higgins, & Rothstein, 2013) was used to compute point effect size estimates (95% confidence intervals, CIs) for individual studies, weighted mean effect sizes, heterogeneity statistics, effect size comparisons, and moderator analyses. Random effects models were computed for all weighted comparisons to account for unobserved variance across studies (Hedges, 1994), except where different measures were compared within a syndrome, when it was assumed that measures were estimating the same (fixed) effect.

2.6.2. Heterogeneity

Heterogeneity of effect sizes was measured using the Q and I^2 statistics, which provide an index of the consistency of effects across studies (Higgins, Thompson, Deeks, & Altman, 2003). Cochran's Q statistic (based on the X^2 distribution) tests the null hypothesis that all studies are evaluating the same effect, and is computed by summing the squared deviations of each effect size estimate from the overall meta-

analytic estimate. A shortcoming of the *Q* statistic is that it is overly sensitive to detecting negligible variability in meta-analyses with high numbers of studies (Huedo-Medina, Sanchez-Meca, Marin-Martinez, & Botella, 2006). The I^2 statistic was introduced to overcome the limitations of the *Q* statistic, and measures the *extent* of heterogeneity (rather than just statistically significant heterogeneity in the case of *Q*) measured as a percentage of total variance that is explained by between-study variance. I^2 statistic values of 25, 50, and 75% represent low, medium, and high heterogeneity, respectively (Higgins et al., 2003). The 'between-effect' Q-value reports the statistical significance of the difference between subgroups (i.e., is there significant heterogeneity 'between' groups?). The between-subgroups I^2 value quantifies this heterogeneity between subgroups. CMA also provides I^2 statistics for heterogeneity 'within' each subgroup (e.g., within the clinical samples and separately within the non-clinical samples, when comparing clinical to non-clinical samples). The difference *between* subgroups might be non-significant (Q-statistic), but the heterogeneity *within* subgroups might still be large (I^2 statistic) and significant. Random effects models, allowing estimation of between-study variance and sampling error, were deemed appropriate for comparisons across subgroups (syndromes and measures) as it was predicted that there would be high heterogeneity between studies (differing populations across studies). Between-subgroups Q-values were used to statistically compare point estimates, and significant effects were interpreted as indicating significant differences between two subgroups. In analyses with a significant Q-value and more than two subgroups, *Z*-tests were used to compare each pair of studies (Borenstein, Hedges, Higgins, & Rothstein, 2009).

2.6.3. Independence of effect sizes

Meta-analytic studies assume independence of each observation or effect size analogously to the assumption of independence for most statistical modelling. The overarching principle in meta-analyses is that each analysis must only include one effect size of a given construct for each sample (i.e., IU may be measured against different measures of MDD and SAD from any one given sample – giving two effect sizes – but multiple effect sizes cannot be calculated for the association between IU and MDD). The current analysis employed a range of approaches to ensure independence of calculated effect sizes. Where there was more than one independent sample reported in a given publication, data from each sample were coded. For studies that reported data on multiple syndromes (e.g., mixed diagnostic samples reporting correlations between a measure of IU and measures of GAD and MDD for the same sample), "syndrome" was coded as a grouping variable and analyses were run separately for each. Some studies reported association data between a measure of IU and two or more of the same emotional symptom grouping measures (e.g., Yale-Brown Obsessive Compulsive Scale and Obsessive Compulsive Inventory-Revised correlated against the OBQ-PC), in which case correlations for the multiple emotional symptom measures against the IU measure were averaged, thus producing a single mean effect size for those samples. For comparisons across symptom measures, only associations between IU and non-composite measures were included. Where studies reported data on IU and eligible emotional symptom measures at more than one time point (e.g., pre- and-post treatment) only data from the first time point was coded.

2.6.4. Publication/Dissemination bias

The validity of meta-analyses and systematic reviews can be threatened by a range of publication biases, including inclusion or exclusion of publications depending on statistical significance (given publishing standards this most often meant a bias towards studies reporting statistical significance), language bias (i.e., selectively including publications written in English), and outcome bias (selective reporting of some but not all outcomes depending on their statistical significance; Rothstein, Sutton, & Borenstein, 2005). All of these collectively fall under the umbrella term of dissemination bias, and thus have the capacity to systematically undermine the validity of meta-analyses. A range of approaches have been suggested for minimising the impact of publication bias that were implemented in the current meta-analysis, namely prospective registration of the meta-analysis, the inclusion of unpublished dissertations and non-selective reporting of all outcomes (from published studies) regardless of statistical significance. The presence of publication bias was also assessed in two ways in the current meta-analysis. First, funnel plots were produced for each analysis by plotting the precision (1/standard error) of individual studies (on the y axis) with the corresponding effect size (Fisher's Z) for each study. In the absence of bias (and when the individual studies reflect the same overall effect) the plot will resemble an inverted symmetrical funnel shape – studies with larger sample sizes will be distributed at the top around the mean effect size, and studies with smaller sample sizes, and hence greater variability of effect, at the bottom. When bias is present the plot will be asymmetrical at the bottom and calculation of effects across studies will be overestimated in terms of average effect or strength of relationship (Sterne, Becker, & Egger, 2005). Each funnel plot for the current analyses appeared symmetrical (see Supplementary Material D), suggesting an absence of publication bias. Egger's test of the intercept (Egger, Smith, Schneider, & Minder, 1997) was used to examine whether funnel plots for the whole sample of effects, and for the subsample of effects within each symptom group (e.g., MDD, SAD), were symmetrical, whereby a significant effect is indicative of bias. An additional metric, Orwin's fail-safe *N*, was calculated to assess the potential influence of publication bias (Orwin, 1983). The fail-safe *N* partially addresses the so-called 'file drawer problem' (Rosenthal, 1979) – the tendency of researchers to not publish results that are not significant, instead placing them in file drawer with other non-significant findings (Becker, 2005), and thus contributing to publication bias (i.e., selective reporting of significant results). Orwin's fail-safe *N* refers to the number of unpublished studies of negligible effect that would be necessary to nullify the mean effect size in the observed data, and is defined as:

$$\frac{k * (ES - ES_c)}{ES_c}$$

where *k* is the number of studies contributing to the effect size, *ES* is the mean effect size, and ES_c is the criterion effect size that is considered trivial. In the current meta-analysis we set ES_c at $r = 0.10$ following Orwin (1983) and Gentes and Ruscio (2011).

2.6.5. Outlier detection and sensitivity analysis

Sensitivity analysis is an important step for meta-analytic studies, providing insight into whether results are robust (or not) to assumptions that were made when carrying out the analysis (Borenstein et al., 2009); for example, whether removal of studies identified as outliers influence outcomes. A "remove-one" analysis was used for the current meta-analysis in CMA to assess the impact of each study on the combined effect. This approach first analyses all included studies and then reruns for all studies except the second, and so on, until all individual studies have been excluded from an analysis (keeping all other studies).

2.6.6. Moderation analysis

Meta-regression, as applied to meta-analyses, works similarly to traditional regression methods in assessing the relationship between one or more moderators and a dependent variable, but instead of moderators being at the individual subject level they are at the study level, and effect sizes act as dependent variables rather than individual subject scores (Borenstein et al., 2009). The meta-regression model applies a weight for each study. Meta-regression was used to investigate the impact of the age, sex, and sample size on the strength of the bivariate associations between IU and symptom measures.

3. Results

3.1. Effects across all studies

The random effects were first estimated across all studies, followed by estimates by syndrome (defined by symptom measure), clinical (vs. non-clinical) status, age group (child vs. adult, average sample age), sex distribution, IU measure, and symptom measure. The influence of sample size was also examined. The overall random-effects point estimate (95% CI) was 0.51 (0.50–0.52, $p < .001$, $N = 335$ independent effects). The Q-value was significant 2024.01 ($p < .001$) and the I^2 was indicative of high heterogeneity (83.50), which was consistent with the a priori decision to use a random-effects model. Egger's test was non-significant, $t(333) = 1.22$, p (2-tailed) = 0.22, suggesting that the random-effects estimate was not biased. Orwin's fail-safe N was 1559, indicating that 1559 unpublished studies with no effect would be necessary to render the effect in this study trivial (< 0.10), suggesting that the file drawer problem was unlikely to have biased results. One study removed point estimate statistics (i.e., removing each effect size and calculating the overall random-effects model) varied between 0.507 and 0.509 ($ps < 0.001$), indicating that removal of individual studies had minimal impact on the overall effect size. A bibliography of all included studies is provided in Supplementary Material E.

3.1.1. Syndrome comparison

A comparison by syndrome yielded an overall random-effects point estimate of 0.52 (0.51–0.53), a significant between-effect Q-value of 118.68 ($p < .001$, $N = 335$ independent effects), and an overall I^2 indicative of high heterogeneity (83.50). Random-effects estimates (CIs) for each syndrome were: 0.50 (depression, 0.47–0.52), 0.50 (social anxiety, 0.46–0.55), 0.42 (OCD, 0.40–0.45), 0.47 (panic/agoraphobia, 0.43–0.52), 0.57 (GAD, CI = 0.56–0.59), and 0.40 (eating disorders, 0.32–0.47). Z-tests revealed that the estimate was highest for GAD (all $ps < 0.01$), lowest for eating disorders (all $ps < 0.01$, except $p = .09$ in comparison to OCD), and depression was higher than OCD ($p < .001$). No other differences were statistically significant ($ps > 0.05$). Within each syndrome, I^2 was indicative of high heterogeneity and varied between 71.25 ($p < .001$, GAD) and 89.40 ($p < .001$, SAD) for all disorders except eating disorders, which was low to medium (37.66, $p = .12$). Forest plots of all studies for each syndrome are provided in Supplementary Material F.

3.1.2. Clinical status comparison

A comparison by clinical status (clinical vs. non-clinical) yielded an overall random-effects point estimate of 0.51 (0.49–0.52), a significant between-effect Q-value of 31.37 ($p < .001$, $n = 318$ independent effects), and an I^2 indicative of high heterogeneity (83.74). The random-effects point estimates for clinical samples (0.45, CI = 0.42–0.47, $n = 100$) was lower than for non-clinical samples (0.53, CI = 0.51–0.54, $n = 218$). I^2 values within clinical (81.88, $p < .001$) and non-clinical (82.63, $p < .001$) studies were indicative of high heterogeneity.

3.1.3. Sex comparison

A bivariate meta-regression indicated that the proportion of women in each sample moderated the relationship between IU and symptoms, with a significant slope coefficient of 0.003 (0017–0.0043, Fisher's Z = 4.58, $p < .001$), indicating that for each percentage increase in the proportion of women the strength of association between IU and symptoms increased by 0.003. The proportion of total between-study variance explained by sex was 10% ($R^2 = 0.10$).

3.1.4. Age group comparison

A comparison of child versus adult studies yielded an overall point estimate of 0.51 (0.49–0.52), a non-significant between-effect Q-value of 0.01 ($p = .92$, $n = 238$ independent effects), and an I^2 indicative of high heterogeneity (85.40). Random-effects point estimates for children and adults were 0.50 (0.42–0.58) and 0.51 (0.49–0.52), respectively. I^2 values were indicative of high heterogeneity within child (90.14, $p < .001$, $n = 15$) and adult (84.64, $p < .001$, $n = 223$) studies. A bivariate meta-regression with mean age of each sample yielded a non-significant slope coefficient of −0.0015 (−0.0035–0.0005, Fisher's Z = −1.46, $p = .14$). The proportion of total between-study variance explained by age was 1% ($R^2 = 0.01$).

3.1.5. IU measure comparison

A comparison of IU measures with at least 10 studies each (IUS-27, IUS-12, OBQ-44 item version) yielded an overall random-effects point estimate of 0.51 (0.49–0.52), a significant between-effects Q-value of 42.13 ($p < .001$, $n = 305$ independent effects), and an I^2 value indicative of high heterogeneity (83.63). Random-effects point estimates for the IUS-27, IUS-12, and OBQ-44 were 0.53 (0.52–0.55), 0.49 (0.47–0.51), and 0.40 (0.35–0.44), respectively. Z-tests revealed that all estimates significantly differed from each other ($ps < 0.01$). I^2 values for studies using the IUS-27 (80.23, $n = 178$), IUS-12 (85.27, $n = 101$), and OBQ-44 (70.01, $n = 26$) were indicative of high heterogeneity (all $ps < 0.001$).

3.1.6. Symptom measure comparison

Symptom measures with at least 10 studies each were compared, which included the Beck Depression Inventory (BDI), Depression Anxiety Stress Scales-Depression (DASS-D), Yale-Brown Obsessive Compulsive Scale (Y-BOCS), Obsessive Compulsive Inventory-Revised (OCI-R), Anxiety Sensitivity Index (ASI), and Penn State Worry Questionnaire (PSWQ). The overall point estimate was 0.53 (0.51–0.54), the between-effect Q-value was significant (76.31, $p < .001$, $n = 206$ independent effects), and the I^2 value was indicative of high heterogeneity (80.15). Random-effects point estimates were 0.50 (0.47–0.53, BDI), 0.50 (0.45–0.54, DASS-D), 0.33 (0.26–0.41, Y-BOCS), 0.45 (0.41–0.49, OCI-R), 0.49 (0.44–0.54, ASI), 0.57 (0.56–0.59, PSWQ). Z-tests revealed that the BDI ($p < .01$) and ASI ($p = .03$) estimates were significantly higher than the Y-BOCS, and the BDI was also higher than the OCI-R ($p = .02$). No other differences were statistically significant ($ps > 0.06$). The I^2 values were indicative of medium to high heterogeneity in studies using each measure: BDI (77.20, $p < .001$), DASS-D (55.20, $p = .004$), Y-BOCS (56.87, $p = .01$, OCI-R (81.28, $p < .001$), ASI (78.86, $p < .001$), PSWQ (70.78, $p < .001$).

3.1.7. Impact of sample size

Sample size varied from 21 to 2450 ($M = 270.95$, $SD = 269.02$), with skewness of 2.77 and kurtosis of 13.86. Reigning in the sample sizes larger than 1000 to equal 1000 ($n = 6$) resulted in acceptable skewness of 1.45 and kurtosis of 1.58. A bivariate meta-regression between study sample size (reigned) and the relationship between IU and symptoms yielded a slope efficient of < 0.001 (< −0.001 – < 0.001, Fisher's Z = 0.53, $p = .60$). The proportion of total between-study variance explained by sample size was < 1% ($R^2 < 0.01$).

3.2. Depression studies

The overall random-effect point estimate was 0.50 (0.47–0.52) with a significant Q-value of 319.91 ($p < .001$, $n = 76$ independent effects). I^2 was indicative of high heterogeneity (76.56, $p < .001$). Egger's test was non-significant, $t(74) = 1.48$, p (2-tailed) = 0.14. Orwin's fail-safe N was 347.

3.2.1. Clinical status comparison

A comparison by clinical status (clinical vs. non-clinical) yielded an overall random-effects point estimate of 0.50 (0.47–0.52), a non-significant between-effects Q-value of 1.47 ($p = .23$, $n = 73$ independent effects), and an I^2 indicative of high heterogeneity (77.39). The

random-effects point estimates were 0.46 (0.40–0.52) and 0.50 (0.48–0.53) for clinical (n = 17) and non-clinical (n = 56) samples, respectively. I^2 values of were indicative of high heterogeneity within clinical (83.61) and non-clinical (74.80) samples (ps < 0.001).

3.2.2. Sex comparison

A bivariate meta-regression with proportion of women in each sample yielded a non-significant slope coefficient of 0.0016 (−0.0009–0.0041, Fisher's Z = 1.27, p < .21). The proportion of total between-study variance explained by sex was 2% (R^2 = 0.02).

3.2.3. Age group comparison

Only one child study assessed depression symptom, so a comparison between adult and child samples could not be completed. A bivariate meta-regression with mean age of each sample yielded a non-significant slope coefficient of −0.0027 (−0.0055–0.0001, Fisher's Z = −1.91, p = .056). The proportion of total between-study variance explained by age was 5% (R^2 < 0.05).

3.2.4. IU measure comparison

A comparison of IU measures with at least 10 studies each (IUS-27, IUS-12) yielded an overall random-effects point estimate of 0.49 (0.47–0.51) and a significant between-effects Q-value of 14.63 (p < .001, n = 67 independent effects). Random-effects point estimates for the IUS-27 and IUS-12 were 0.53 (0.50–0.56) and 0.44 (0.41–0.47), respectively. I^2 values for each measure indicated moderate (IUS-12 = 55.15) to high (IUS-27 = 73.11) heterogeneity (ps ≤ 0.001).

3.3. Social anxiety studies

The overall random-effect point estimate was 0.50 (0.46–0.55) and the Q-value of 292.31 was significant (p < .001, n = 32 independent effects). I^2 was indicative of high heterogeneity (89.40, p < .001). Egger's test was non-significant, t(31) = 0.37, p (2-tailed) = 0.72. Orwin's fail-safe N was 143.

3.3.1. Clinical status comparison

A comparison by clinical status (clinical vs. non-clinical) yielded an overall random-effects point estimate of 0.49 (0.44–0.54) and a significant between-effects Q-value of 4.30 (p = .04, n = 31 independent effects). The random-effects point estimates were 0.44 (0.37–0.51) and 0.54 (0.48–0.60) for clinical (n = 12) and non-clinical (n = 19) samples, respectively. I^2 values were indicative of high heterogeneity in clinical (85.68) and non-clinical (90.64) samples (ps < 0.001).

3.3.2. Sex comparison

A bivariate meta-regression with proportion of women in each sample and IU yielded a significant slope coefficient of 0.006 (0.0011–0.0098, Fisher's Z = 2.47, p < .01). The proportion of total between-study variance explained by sex was 19% (R^2 = 0.19).

3.3.3. Age group comparison

Only three child studies assessed social anxiety symptoms, so a comparison between adult and child samples could not be completed. A bivariate meta-regression with mean age of each sample yielded a non-significant slope coefficient of 0.0035 (−0.0036–0.0105, Fisher's Z = 0.96, p = .34). The proportion of total between-study variance explained by age was < 1% (R^2 < 0.01).

3.3.4. IU measure comparison

A comparison of IU measures with at least 10 studies each (IUS-27, IUS-12) yielded an overall random-effects point estimate of 0.51 (0.46–0.55) and a non-significant between-effects Q-value of 0.12 (p = .73, n = 28 independent effects). Random-effects point estimates for the IUS-27 and IUS-12 were 0.50 (0.44–0.56) and 0.52 (0.44–0.59), respectively. I^2 values were indicative of high heterogeneity within studies using the IUS-27 (83.69, n = 15) and IUS-12 (92.68, n = 17, ps < 0.001).

3.4. Panic/agoraphobia studies

The overall random-effect point estimate was 0.47 (0.43–0.52) with a significant Q-value of 172.16 (p < .001, n = 29 independent effects). I^2 was indicative of high heterogeneity (83.74, p < .001). Egger's test was non-significant, t(27) = 1.12, p (2-tailed) = 0.27. Orwin's fail-safe N was 114.

3.4.1. Clinical status comparison

A comparison by clinical status (clinical vs. non-clinical) yielded an overall random-effects point estimate of 0.45 (0.41–0.49) and a significant between-effects Q-value of 13.30 (p < .001, n = 28 independent effects). The random-effects point estimates were 0.37 (0.31–0.43) and 0.52 (0.47–0.57) for clinical (n = 10) and non-clinical (n = 18) samples, respectively. I^2 values of were indicative of moderate to large heterogeneity within studies using clinical (62.94) and non-clinical (82.31) samples (ps < 0.001).

3.4.2. Sex comparison

A bivariate meta-regression with proportion of women in each sample yielded a non-significant slope coefficient of 0.0017 (−0.0023–0.0057, Fisher's Z = 0.84, p = .40). The proportion of total between-study variance explained by sex was < 1% (R^2 < 0.01).

3.4.3. Age group comparison

Only one child study assessed panic symptoms, so a comparison between adult and child samples could not be completed. A bivariate meta-regression with mean age of each sample yielded a non-significant slope coefficient of −0.0008 (−0.0069–0.0053 Fisher's Z = −0.26, p = .79). The proportion of total between-study variance explained by age was < 1% (R^2 < 0.01).

3.4.4. IU measure comparison

A comparison of IU measures with at least 10 studies each (IUS-27, IUS-12) yielded an overall random-effects point estimate of 0.46 (0.41–0.51) and a non-significant between-effects Q-value of 0.045 (p = .83, n = 26 independent effects). Random-effects point estimates for the IUS-27 and IUS-12 were 0.46 (0.37–0.54) and 0.47 (0.41–0.52), respectively. I^2 values were indicative of high heterogeneity for studies using the IUS-27 (78.60, n = 10) and IUS-12 (86.37, n = 16, ps < 0.001).

3.4.5. Symptom measure comparison

The ASI is commonly used as a dependent variable in panic studies, despite it being a cognitive vulnerability factor rather than a symptom measure. To inform the field about the impact of using the ASI versus a symptom measure, we directly compared associations with the PDSS. The overall fixed-effect point estimate (assuming both measures are estimating the same effect) was 0.46 (0.44–0.47) and the between-effects Q-value was significant (24.68, p < .001, n = 27 independent effects). Fixed-effect point estimates were 0.48 (0.46–0.50, ASI) and 0.38 (0.35–0.42, PDSS). I^2 values were indicative of high heterogeneity for the ASI (84.30, n = 21 studies) and PDSS (73.96, n = 6 studies, ps < 0.01).

3.5. Generalised anxiety disorder studies

The overall random-effect point estimate was 0.57 (0.56–0.59) and the Q-value was significant (413.91, p < .001, n = 120 independent effects). Egger's test was non-significant, t(118) = 0.38, p (2-tailed) = 0.71. Orwin's fail-safe N was 660. I^2 of 71.25 (p < .001). I^2 was indicative of high heterogeneity (71.25, p < .001).

3.5.1. Clinical status comparison

A comparison by clinical status (clinical vs. non-clinical) yielded an overall random-effects point estimate of 0.57 (0.56–0.58) and a significant between-effects Q-value of 5.14 ($p = .02$, $n = 112$ independent effects). The random-effects point estimates were 0.54 (0.51–0.57) and 0.58 (0.56–0.60) for clinical ($n = 29$) and non-clinical ($n = 83$) samples, respectively. I^2 values of were indicative of moderate to high heterogeneity for studies using clinical (63.94) and non-clinical (71.20) samples ($ps < 0.001$).

3.5.2. Sex comparison

A bivariate meta-regression between proportion of women in each sample and IU yielded a significant slope coefficient of 0.0028 (0.010–0.0045, Fisher's Z = 3.12, $p = .002$). The proportion of total between-study variance explained by sex was 15% ($R^2 = 0.15$).

3.5.3. Age group comparison

Seven studies assessed GAD symptoms in children. Although this fell below the a priori criterion of 10, we ran an exploratory analysis. A comparison by age (child vs. adult) yielded an overall random-effects point estimate of 0.58 (0.56–0.60) and a non-significant between-effects Q-value of 0.71 ($p = .40$, $n = 91$ independent effects). The random-effects point estimates were 0.62 (0.52–0.71, $n = 7$) and 0.58 (0.56–0.60, $n = 84$) for child and adult studies, respectively. I^2 values were indicative of high heterogeneity within child (84.81) and adult (73.69) samples ($ps < 0.001$). A bivariate meta-regression with mean age of each sample yielded a non-significant slope coefficient of −0.001 (−0.004–0.002, Fisher's Z = −0.72, $p = .47$). The proportion of total between-study variance explained by age was < 1% ($R^2 < 0.01$).

3.5.4. IU measure comparison

A comparison of IU measures with at least 10 studies each (IUS-27, IUS-12) and the IUS-C ($n = 6$ studies and was included as an exploratory analysis) yielded an overall random-effects point estimate of 0.58 (0.56–0.59) and a non-significant between-effects Q-value of 5.45 ($p = .07$, $n = 116$ independent effects). Random-effects point estimates for the IUS-27, IUS-12, and IUS-C were 0.58 (0.57–0.60), 0.55 (0.52–0.58), and 0.63 (0.49–0.74), respectively. I^2 values were indicative of high heterogeneity for the IUS-27 (64.89, $n = 75$), IUS-12 (76.58, $n = 35$), and IUS-C (87.07, $n = 6$, all $ps < 0.001$).

3.5.5. Symptom measure comparison

The PSWQ is commonly used as a dependent variable in GAD studies, despite it being a narrow symptom measure (worry) compared to broader measures of GAD symptoms. To inform the field about the impact of using the PSWQ versus a GAD symptom measure, we directly compared associations. The overall fixed effect point estimate (assuming the PSWQ and symptom measures are estimating the same effect) was 0.57 (0.57–0.58) and the between-effects Q-value was not significant (0.08, $p = .77$, $n = 118$ independent effects). Fixed-effect point estimates for each group of studies were 0.57 (0.57–0.58, PSWQ) and 0.58 (0.56–0.60, GAD symptom measures). I^2 values were indicative of high heterogeneity for studies using the PSWQ (70.87, $n = 104$ studies) and symptom measures (78.03, $n = 14$ studies, $ps < 0.001$).

3.6. Obsessive compulsive disorder studies

The overall random-effect point estimate was 0.42 (0.40–0.45) with a significant Q-value of 299.87 ($p < .001$, $n = 69$ independent effects). I^2 was indicative of high heterogeneity (77.32, $p < .001$). Egger's test was non-significant, $t(67) = 0.43$, p (2-tailed) = 0.67. Orwin's fail-safe N was 246.

3.6.1. Clinical status comparison

A comparison by clinical status (clinical vs. non-clinical) yielded an overall random-effects point estimate of 0.41 (0.38–0.44) and a significant between-effects Q-value of 9.29 ($p = .002$, $n = 65$ independent effects). The random-effects point estimates were 0.36 (0.32–0.41) and 0.45 (0.42–0.49) for clinical ($n = 29$) and non-clinical ($n = 36$) samples, respectively. I^2 values were indicative of high heterogeneity for clinical (68.99) and non-clinical (78.56) samples ($ps < 0.001$).

3.6.2. Sex comparison

A bivariate meta-regression with proportion of women in each sample yielded a significant slope coefficient of 0.0071 (0.043–0.0099, Fisher's Z = 5.01, $p < .001$). The proportion of total between-study variance explained by sex was 42% ($R^2 = 0.42$).

3.6.3. Age group comparison

Only three studies investigated OCD symptoms in exclusive child samples, so a comparison between adult and child samples could not be completed. A bivariate meta-regression with mean age of each sample yielded a non-significant slope coefficient of −0.0025 (−0.0073–0.0023, Fisher's Z = −1.02, $p = .31$). The proportion of total between-study variance explained by age was < 1% ($R^2 < 0.01$).

3.6.4. IU measure comparison

A comparison of IU measures with at least 10 studies each (IUS-27, IUS-12, OBQ-44) yielded an overall random-effects point estimate of 0.43 (0.40–0.45) and a non-significant between-effects Q-value of 5.12 ($p = .077$, $n = 60$ independent effects). Random-effects point estimates for the IUS-27, IUS-12, and OBQ-44 were 0.46 (0.42–0.50), 0.43 (0.36–0.49), and 0.38 (0.33–0.43), respectively. I^2 values were indicative of high heterogeneity for studies using the IUS-27 (74.01, $n = 26$), IUS-12 (83.36, $n = 13$), and OBQ-44 (73.10, $n = 21$, all $ps < 0.001$).

3.6.5. Symptom measure comparison

Associations between IU and the self-report and clinician-rated versions of the Y-BOCS were compared to determine whether effects differed by reporter. A fixed-effect model was used because it was assumed that both versions were attempting to estimate the same effect. The overall fixed-effect point estimate was 0.34 (0.30–0.39) and the between-effects Q-value was significant (3.88, $p = .049$, $n = 11$ independent effects). Fixed-effect point estimates were 0.30 (0.23–0.36, clinician-rated) and 0.39 (0.32–0.45, self-reported). The I^2 values were indicative of low heterogeneity within clinician-rated studies (31.56, $p = .20$, $n = 6$ studies) and medium to high heterogeneity within self-report studies (66.67, $p = .02$, $n = 5$ studies).

3.7. Eating disorder studies

The overall random-effect point estimate was 0.40 (0.32–0.47) with a non-significant Q-value (12.83, $p = .118$, $n = 9$ independent effects). Egger's test was non-significant, $t(7) = 0.82$, p (2-tailed) = 0.44. Orwin's fail-safe N was 31. I^2 was indicative of low to medium heterogeneity (37.66, $p = .118$). Due to the low number of studies in each comparison group, exploratory analyses were performed for some variables as detailed below.

3.7.1. Clinical status comparison

A comparison by clinical status (clinical vs. non-clinical) yielded an overall random-effects point estimate of 0.39 (0.31–0.46) and a non-significant between-effects Q-value of 0.075 ($p = .784$, $n = 9$ independent effects). The random-effects point estimates were 0.37 (0.25–0.48) and 0.40 (CI = 0.29–0.49) for clinical ($n = 3$) and non-clinical ($n = 6$) samples, respectively. I^2 values were indicative of very low (< 0.01, $p = .537$) or moderate 54.30 ($p = .053$) and non-significant heterogeneity for clinical and non-clinical samples,

respectively.

3.7.2. Sex comparison

Seven of the nine studies only included women, and one had 98% women, so there was insufficient variance to calculate this model.

3.7.3. Age group comparison

Most studies included children and adults ($n = 6$), with the remaining studies including only adults, so this analysis could not be calculated. A bivariate meta-regression with mean age of each sample yielded a non-significant slope coefficient of 0.0050 (−0.0136–0.0235, Fisher's Z = 0.53, $p = .599$). The proportion of total between-study variance explained by age was 2% ($R^2 = 0.02$).

3.7.4. IU measure comparison

Eight of the nine studies used the IUS-27 and one used the IUS-12, so meaningful comparisons could not be made.

4. Discussion

4.1. The jury is in: IU is certainly transdiagnostic

The first aim of this meta-analysis was to quantify the strengths of association between IU and symptoms of depression, social anxiety disorder, panic disorder, agoraphobia, GAD, OCD, and eating disorders. The extensive body of evidence across most of these symptom domains facilitated the calculation of precise point estimates with narrow confidence intervals. There was no evidence that publication bias or study sample size influenced results. The strengths of association between IU and symptoms ranged from 0.40 (eating disorders) to 0.57 (GAD), and was significantly higher for GAD and significantly lower for eating disorders, compared to all other syndromes. The association with depression symptoms was also stronger than for OCD. Although reliable, it is important to note that these differences were small and the practical or clinical implications are likely to be minimal. Consistent with our findings, Gentes and Ruscio's (2011) meta-analysis using two definitions of IU within the GAD and OCD literatures, operationalised by the IUS-27 and OBQ-44, respectively, found no evidence that IU was specific to GAD, depression, or OCD. Our study extends these findings to panic disorder/agoraphobia, social anxiety disorder, and eating disorders, and to a range of different measures of IU and symptoms. Our findings allow us to definitively conclude that IU has robust moderate transdiagnostic associations with symptoms of anxiety and related disorders, depression, and eating disorders.

4.2. Moderators have minimal influence on associations between IU and symptoms

The second aim was to investigate how syndrome, clinical status, sex, age, IU measure, and symptom measure influence these associations. Heterogeneity in associations within all syndromes was high, except for eating disorders, suggesting that study methods, measures, and samples other than those assessed in the current study may indeed moderate these associations. The strength of association was weaker for clinical than non-clinical samples when all studies were included (0.45 vs. 0.53), and also within studies assessing panic/agoraphobia (0.37 vs. 0.52), OCD (0.36 vs. 0.45), and GAD (0.54 vs. 0.58) symptoms. One explanation of these differences is range attenuation within clinical samples, where participants are selected on the basis of elevated symptoms (direct range restriction) or diagnosis (indirect range restriction). IU and symptoms are considered to be dimensional (e.g., Asmundson, Weeks, Carleton, Thibodeau, & Fetzner, 2011; Carleton et al., 2012; Olatunji, Broman-Fulks, Bergman, Green, & Zlomke, 2010; Olatunji, Williams, Haslam, Abramowitz, & Tolin, 2008; Ruscio, Borkovec, & Ruscio, 2001; Ruscio, Ruscio, & Keane, 2002), so artificially restricting the range of scores will attenuate the strengths of association (Kennet-Cohen et al., 2018). It is unclear whether effect differences of 0.15 or less are meaningful in practice, but it is important to be aware that restricting samples to the clinical range may reduce the strength of association somewhat, at least for panic/agoraphobia, OCD, and GAD.

Across all studies the proportion of women was positively and significantly associated with higher associations between IU and symptoms, explaining 10% of the variance in the relationship, although the impact of sex varied across syndrome. Sex explained a small and non-significant proportion of variance in the strengths of association between IU and panic/agoraphobia (< 1%) and depression (2%) symptoms, a larger and significant proportion of variance for GAD (15%) and social anxiety (19%) symptoms, and a substantial proportion of variance in OCD (42%) symptoms. There was insufficient variance across eating disorder studies to investigate the impact of sex.

The varied influence of sex on associations was unexpected. There is evidence of measurement invariance across men and women for the IUS-12 (Bottesi, Noventa, Freeston, & Ghisi, 2019; Carleton, Mulvogue, et al., 2012; Dekkers, Jansen, Salemink, & Huizenga, 2017; Roma & Hope, 2017), suggesting that men and women demonstrate similar response profiles on this measure. There is also evidence of measurement invariance across sex for at least some of the most common symptom measures identified in our study, including the BDI (Whisman, Judd, Whiteford, & Gelhorn, 2013), DASS (Gomez, Summers, Summers, Wolf, & Summers, 2014), Social Phobia and Anxiety Inventory (Bunnell, Joseph, & Beidel, 2013), ASI (Jardin et al., 2018), EDE-Q (Penelo, Negrete, Portell, & Raich, 2013), and the PSWQ in children (Păsărelu et al., 2017) and adults (Suárez-Falcón, Monroy-Cifuentes, & Ruiz, 2018). We were unable to find any studies of measurement invariance across men and women for the Y-BOCS or OBQ-44, and while there is some evidence of measurement invariance for the OCI-child version (Rodríguez-Jiménez et al., 2017) no studies have been conducted in adults using the OCI-R. If there are systematic differences in response styles between men and women that are inherent in these measures, the moderating effect of sex may simply be a measurement artefact rather than a genuine difference between sexes. Therefore, the finding that sex was a significant moderator of the relationship between IU and OCD symptoms needs to be interpreted with caution until invariance has been demonstrated, which is an important future research direction. Notwithstanding this caveat, researchers investigating relationships between IU and symptoms may need to account for sex in statistical models, particularly for SAD, GAD, and OCD. Future research into sex-linked genetic, biological, hormonal, and cultural factors that could explain differences in IU levels, and associated impacts on symptoms for social anxiety, depression, and OCD, would be informative.

Comparisons between child and adult samples across all studies and across GAD studies in particular, for which there were sufficient studies for a meaningful comparison, revealed non-significant differences. Average sample age as a continuous variable across all studies was also a non-significant moderator (explaining 1% of the variance). Age was also not a significant moderator for any of the analyses within specific syndromes and the proportion of variance explained by age was small (≤5%). This finding indicates continuity of associations between IU and symptoms across the lifespan.

Only three IU measures had been used in at least 10 studies and were therefore compared. Overall, the IUS-27 was most strongly associated with symptoms, followed by the IUS-12 and OBQ-44, respectively. The difference between the IUS-27 and IUS-12 was small (0.04), as was the difference between the IUS-12 and OBQ-44 (0.09), and thus the impact of IU measure on associations with symptoms does not appear to be substantial. Within symptom syndromes, compared to the IUS-12, the IUS-27 enjoyed a significant but only slightly stronger association with depression (0.44 vs. 0.53) symptoms. Consistent with Gentes and Ruscio (2011), the IUS-27 was also more strongly associated with OCD symptoms than the OBQ-44 (0.46 vs. 0.38), but this difference was not statistically significant. The strengths of associations were

not significantly different between IU measures for social anxiety, GAD, or panic disorder and agoraphobia symptoms. Too few studies measured eating disorder symptoms to compare associations across IU measures. Although not all measures could be compared within all syndromes due to an insufficient number of studies (e.g., OBQ-44 in GAD studies), for the comparisons that could be made the impact on associations of using different IU measures was minimal. Although there might be important theoretical reasons for using a particular IU or symptom measure our findings suggest that, for total scores at least, all measures of IU yield similar strengths of association across anxiety and related disorders. Indeed, the finding that the magnitude of association between the IUS-27 and OCD symptoms was similar to the IU measure developed within the OCD literature (i.e., OBQ-44), suggests that it may not be necessary for OCD studies to rely on the OBQ-44. The IUS-12 has the advantages of possibly being more transdiagnostic in content (Gentes & Ruscio, 2011) and brevity, with the latter reducing respondent burden across clinical and research contexts.

The comparisons across different symptom measures revealed that associations did not significantly differ across the BDI, DASS-D, or PSWQ. The association for the Y-BOCS was significantly weaker than for the BDI ($r_{diff} = 0.17$) and ASI ($r_{diff} = 0.16$), and the BDI was also higher than the OCI-R ($r_{diff} = 0.05$), although the differences were small in magnitude. The weaker association for the two OCD measures, albeit more pronounced for the Y-BOCS, may reflect a slightly weaker association for OCD symptoms than depression symptoms. Alternatively, the difference for the Y-BOCS might mainly reflect method variance, given that this comparison combined self- and clinician-rated versions of the Y-BOCS, and the clinician-rated version yielded weaker effects than the self-report version. Regardless, it is unlikely that differences of this small magnitude would have any practical impact. The relatively comparable strengths of association across syndromes and most measures provide further support for the transdiagnostic nature of IU.

The use of the ASI in panic disorder studies appears to strengthen the relationship compared to the panic disorder severity scale, although the difference was small ($r_{diff} = 0.10$). Similarly, a narrow measure of GAD symptoms (PSWQ) yielded a similar estimate as broader symptom measures, indicating that studies using these different outcomes can be combined. Overall, with few exceptions, our findings indicate that the strengths of the relationship between IU and symptoms are only weakly, if at all, influenced by syndrome, clinical status, sex, age, IU measure, symptom measure, or sample size.

4.3. Implications

The most important implication of these findings is that for the purposes of case formulation, treatment evaluation, and research the relationships between IU and symptoms can be accurately assessed regardless of the range of clinical, demographic, and measurement factors investigated in this study. Clinicians and researchers would be best served to select measures for theoretical (e.g., specific interest in measure subscales, item content that is most closely aligned with the study aims) or pragmatic (e.g., brevity) reasons, rather than on the basis of client presentation. Nonetheless, it is important to be aware that restricting samples along the dimension of IU or symptoms (e.g., clinical samples), sex, IU measure, and symptom measure, or indeed their combination, may have a small but reliable influence over strengths of association with symptoms. In most instances, however, the impact of these factors is likely to be negligible. Clinicians and researchers can proceed with confidence that if they use well-validated IU and symptom measures they are likely to reliably estimate associations.

4.4. Strengths, limitations and future research directions

The main strength of our study was the large sample of effects, which facilitated precise point estimates and the capacity to investigate the influence of a range of potential moderators. However, our study was limited in several important ways, which invites opportunities for future research. First, there were insufficient studies to conduct all planned comparisons within all syndromes, so some of our conclusions may not generalise across all disorders. There are also additional sociodemographic and cultural factors (e.g., ethnicity) that might be important moderators that were not investigated in this study. Second, unlike a previous meta-analysis (Gentes & Ruscio, 2011), we included symptom measures that assessed cardinal features of disorders but that did not necessarily assess the full breadth of symptoms within a disorder (e.g., PSWQ for GAD). This decision was made because such measures are commonly used to assess important dimensions of particular disorders (e.g., pathological worry in GAD), and thus it is important for researchers and clinicians to be informed about how these decisions may or may not influence the relationships between the measures they select and IU. It is also important to note that very few symptom measures were used in enough studies to be included in the comparisons, so restricting measures to only those that comprehensively assess symptoms of a particular syndrome would have limited the number of variables we could investigate. It is, however, important to acknowledge that the relationships investigated here for such measures may not necessarily generalise to measures of the whole syndrome, and also that some constructs (including worry) are transdiagnostic (Harvey, Watkins, Mansell, & Shafran, 2004; McEvoy, Mahoney, & Moulds, 2010; Segerstrom, Tsao, Alden, & Craske, 2000), and as such may not be unique to the syndrome to which it was allocated in our study. For instance, associations between IU and worry may be similar across syndromes and are unlikely to be unique to GAD.

A third limitation of our study is that we only investigated bivariate associations, and therefore did not take into account comorbid symptoms and other relevant factors (e.g., negative affect, neuroticism, disorder-specific and transdiagnostic cognitive vulnerability factors). There is an extensive body of research demonstrating that IU is uniquely associated with symptoms after accounting for these factors (Hong, 2013; Koerner, Mejia, & Kusec, 2017; Mahoney & McEvoy, 2012b; Norton & Mehta, 2007; Norton, Sexton, Walker, & Norton, 2005; Sexton, Norton, Walker, & Norton, 2003; Shihata, McEvoy, & Mullan, 2017), but the strengths of associations would certainly be attenuated to a greater or lesser degree depending upon which constellation of factors was included in the model. This issue is complex as studies tend to control for a unique array of factors, and thus questions about the unique contributions of IU and other factors to the prediction of symptoms may best be resolved within individual studies or with more narrowly focused meta-analyses.

A fourth limitation is that the cross-sectional associations investigated in this meta-analysis cannot provide direct evidence of causation. Longitudinal and experimental studies, which can demonstrate temporal primacy of IU with respect to emotional disorder symptoms, are required to build the case that changes in IU are causally related to changes in symptoms. Intervention studies have shown that changes in IU are associated with changes in symptoms across different treatment formats (McEvoy & Erceg-Hurn, 2016), but evidence of temporal relationships is required during the change process. It is well-known that cross-sectional associations can lead to biased estimates of prospective associations (Cole & Maxwell, 2003; Maxwell & Cole, 2007), so our findings do not speak to sensitivity to change over time or IU as a mechanism of change, which are important areas for future research. A recent systematic review of experimental and temporal precedence studies of IU with respect to anxiety, mood, and OCD symptoms ($N = 15$ studies in 12 articles) found that the strongest support for IU as a causal mechanism for anxiety-related symptoms, with some evidence for negative affect (Rosser, 2019). These findings were qualified by the recommendation that more experimental and clinical research is required before concluding that IU is a causal, transdiagnostic mechanism. Very few studies had investigated depression-linked relationships, and the evidence for a causal relationship

with OCD-related symptoms was weak. No studies were identified that experimentally examined IU in eating disorders.

A fifth limitation is that we restricted our study to total scores on IU measures. There is some psychometric evidence of multidimensionality for measures of IU (Boelen & Lenferink, 2018; Freeston et al., 1994; McEvoy & Mahoney, 2011), and there is some evidence that different components of IU are more or less strongly associated with different emotional disorders. For instance, McEvoy and Mahoney (2012) found that prospective IU (cognitive component), was uniquely associated with GAD (i.e., worry) and OCD symptoms, whereas inhibitory IU (behavioural component) was uniquely associated with panic/agoraphobia, social anxiety, and depression symptoms. Nelson, Liu, Sarapas, and Shankman (2016) also found evidence of divergent validity in a study using an experimental paradigm to investigate the startle reflex, with inhibitory IU being associated with attenuated startle potentiation and prospective IU with heightened startle potentiation (see Tanovic et al., 2018, for a review). On the other hand, the factor structure of the IUS-27 has been found to be somewhat unstable (McEvoy & Mahoney, 2011; Norton, 2005; Roma & Hope, 2017), and more recent psychometric analyses using bifactor modelling have failed to find evidence of strong subfactors in the IUS-12, and rather support a conclusion that most of the reliable variance in IU measures may in fact be shared across items and subfactors (Hale et al., 2016; Lauriola et al., 2018; Shihata et al., 2018). Investigating differential associations across subscales for the multitude of IU and symptom measure subscales was beyond the scope of this study, particularly given the uncertainty that recent research casts on the validity of IU subfactors. Future work should consider using bifactor models of IU measures to determine whether subfactors explain unique, reliable, and meaningful variance in dependent variables.

There is overwhelming support for the transdiagnostic nature of IU, and therefore research for the purposes of examining cross-sectional relationships between self-reported IU and the syndromes included in this study is no longer required. Future research should move on to answering important questions about causal relationships between IU and symptoms measures (Rosser, 2019; Shihata et al., 2016), and about how IU manifests across various levels of analysis, including cognition, behaviour, neurobiology, genetics, lab-based fear learning and extinction, and so on (see McEvoy, Carleton, Correa, Shankman, & Shihata, 2019). Investigating how self-reported IU reflects the outcome of interactions amongst these factors, and how these interactions can facilitate up-regulation or down-regulation of psychopathological symptoms, will move the literature forward in important ways in terms of our understanding of these mechanisms and opportunities for intervention. For instance, fear conditioning studies have demonstrated an association between post-extinction fear recovery (i.e., relapse) and self-reported IU following standard extinction procedures, but this association disappears when novel stimuli are paired with conditioning stimuli during extinction (Dunsmoor, Campese, Ceceli, LeDoux, & Phelps, 2015; Lucas, Luck, & Lipp, 2018). Understanding the mechanisms through which novel stimuli during extinction reduces the impact of IU on relapse will have direct translational implications for clinical practice.

4.5. Conclusions

The extensive body of IU research conducted to date enabled us to calculate precise estimates of the magnitudes of association between IU and symptoms of anxiety disorders, depression, and eating disorders, along with the influence of a range of clinical, demographic, and measurement factors on these associations. This meta-analysis provides strong evidence that IU is robustly and moderately associated with symptoms of anxiety, depression, and eating disorders. Moreover, the strengths of these associations were relatively unaffected by disorder syndrome, clinical status, sex, age, IU measure, or symptom measure. Clinicians and researchers can select measures on pragmatic or theoretical grounds with the knowledge that the relationships between IU and self-reported symptom measures can be reliably estimated. We encourage future research that improves our understanding of temporal and causal relationships between IU and symptoms, and how self-reported IU reflects and interacts with the operation of units of analysis within the cognitive, behavioural, neural, physiological, social, and affective domains.

Role of funding sources

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Contributors

Peter McEvoy contributed to development of the research questions, design, method, analysis, and manuscript writing. Matthew Hyett contributed to development of the research questions, design, method, data extraction, analysis, and manuscript writing. Sarah Shihata contributed to the manuscript writing. Jordan Price contributed to the development of the research questions, design, method, and data extraction. Laura Strachan contributed to the design and data extraction. All authors reviewed and approved the final manuscript.

Declaration of Competing Interest

All authors declare they have no conflicts of interest.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.cpr.2019.101778.

References

Asmundson, G. J. G., Weeks, J. W., Carleton, R. N., Thibodeau, M. A., & Fetzner, M. G. (2011). Revisiting the latent structure of the anxiety sensitivity construct: More evidence of dimensionality. *Journal of Anxiety Disorders, 25*, 138–147. https://doi.org/10.1016/j.janxdis.2010.08.013.

Banducci, A. N., Bujarski, S. J., Bonn-Miller, M. O., Patel, A., & Connolly, K. M. (2016). The impact of intolerance of emotional distress and uncertainty on veterans with co-occurring PTSD and substance use disorders. *Journal of Anxiety Disorders, 41*, 73–81. https://doi.org/10.1016/j.janxdis.2016.03.003.

Bardeen, J. R., Fergus, T. A., & Wu, K. D. (2012). The interactive effect of worry and intolerance of uncertainty on posttraumatic stress symptoms. *Cognitive Therapy and Research, 37*, 742–751. https://doi.org/10.1007/s10608-012-9512-1.

Barlow, D. (2002). *Anxiety and its disorders* (2nd ed.). New York: The Guilford Press, Inc.

Becker, B. J. (2005). Failsafe N or file-drawer number. In H. R. Rothstein, A. J. Sutton, & M. Borenstein (Eds.). *Publication Bias in meta-analysis: Prevention, assessment and adjustments*. Chichester, UK: John Wiley & Sons, Ltd.

Boelen, P. A., & Carleton, R. N. (2012). Intolerance of uncertainty, hypochondriacal concerns, obsessive-compulsive symptoms, and worry. *Journal of Nervous and Mental Disease, 200*, 208–213. https://doi.org/10.1097/NMD.0b013e318247cb17.

Boelen, P. A., & Lenferink, L. I. M. (2018). Latent class analysis of indicators of intolerance of uncertainty. *Scandinavian Journal of Psychology, 59*, 243–251. https://doi.org/10.1111/sjop.12440.

Boelen, P. A., & Reijntjes, A. (2009). Intolerance of uncertainty and social anxiety. *Journal of Anxiety Disorders, 23*, 130–135. https://doi.org/10.1016/j.janxdis.2008.04.007.

Boelen, P. A., Reijntjes, A., & Smid, G. E. (2016). Concurrent and prospective associations of intolerance of uncertainty with symptoms of prolonged grief, posttraumatic stress, and depression after bereavement. *Journal of Anxiety Disorders, 41*, 65–72. https://doi.org/10.1016/j.janxdis.2016.03.004.

Borenstein, M., Hedges, L., Higgins, J., & Rothstein, H. (2009). *Introduction to meta-analysis*. Chichester, UK: John Wiley & Sons, Ltd.

Borenstein, M., Hedges, L., Higgins, J., & Rothstein, H. (2013). *Comprehensive meta-analysis version 3*. Englewood, NJ: Biostat.

Boswell, J. F., Thompson-Hollands, J., Farchione, T. J., & Barlow, D. H. (2013). Intolerance of uncertainty: A common factor in the treatment of emotional disorders. *Journal of Clinical Psychology, 69*, 630–645. https://doi.org/10.1002/jclp.21965.

Bottesi, G., Noventa, S., Freeston, M. H., & Ghisi, M. (2019). Seeking certainty about intolerance of uncertainty: Addressing old and new issues through the intolerance of uncertainty scale-revised. *PLoS ONE, 14*, e0211929. https://doi.org/10.1371/journal.pone.0211929.

Brown, M., Robinson, L., Campione, G. C., Wuensch, K., Hildebrandt, T., & Micali, N. (2017). Intolerance of uncertainty in eating disorders: A systematic review and meta-

analysis. *European Eating Disorders Review, 25*, 329–343. https://doi.org/10.1002/erv.2523.
Buhr, K., & Dugas, M. J. (2006). Investigating the construct validity of intolerance of uncertainty and its unique relationship with worry. *Journal of Anxiety Disorders, 20*, 222–236. https://doi.org/10.1016/j.janxdis.2004.12.004.
Bunnell, B. E., Joseph, D. L., & Beidel, D. C. (2013). Measurement invariance of the social phobia and anxiety inventory. *Journal of Anxiety Disorders, 27*, 84–91. https://doi.org/10.1016/j.janxdis.2012.09.001.
Carleton, R. N. (2016a). Into the unknown: A review and synthesis of contemporary models involving uncertainty. *Journal of Anxiety Disorders, 39*, 30–43. https://doi.org/10.1016/j.janxdis.2016.02.007.
Carleton, R. N. (2016b). Fear of the unknown: One fear to rule them all? *Journal of Anxiety Disorders, 41*, 5–21. https://doi.org/10.1016/j.janxdis.2016.03.011.
Carleton, R. N., Collimore, K. C., & Asmundson, G. J. (2007). Social anxiety and fear of negative evaluation: Construct validity of the BFNE-II. *Journal of Anxiety Disorders, 21*, 131–141. https://doi.org/10.1016/j.janxdis.2006.03.010.
Carleton, R. N., Fetzner, M. G., Hackl, J. L., & McEvoy, P. (2013). Intolerance of uncertainty as a contributor to fear and avoidance symptoms of panic attacks. *Cognitive Behaviour Therapy, 42*, 328–341. https://doi.org/10.1080/16506073.2013.792100.
Carleton, R. N., Gosselin, P., & Asmundson, G. J. (2010). The intolerance of uncertainty index: Replication and extension with an English sample. *Psychological Assessment, 22*, 396–406. https://doi.org/10.1037/a0019230.
Carleton, R. N., Mulvogue, M. K., Thibodeau, M. A., McCabe, R. E., Antony, M. M., & Asmundson, G. J. (2012). Increasingly certain about uncertainty: Intolerance of uncertainty across anxiety and depression. *Journal of Anxiety Disorders, 26*, 468–479. https://doi.org/10.1016/j.janxdis.2012.01.011.
Carleton, R. N., Norton, M. A., & Asmundson, G. J. (2007). Fearing the unknown: A short version of the intolerance of uncertainty scale. *Journal of Anxiety Disorders, 21*, 105–117. https://doi.org/10.1016/j.janxdis.2006.03.014.
Carleton, R. N., Weeks, J. W., Howell, A. N., Asmundson, G. J., Antony, M. M., & McCabe, R. E. (2012). Assessing the latent structure of the intolerance of uncertainty construct: An initial taxometric analysis. *Journal of Anxiety Disorders, 26*, 150–157. https://doi.org/10.1016/j.janxdis.2011.10.006.
Castriotta, N., Dozier, M. E., Taylor, C. T., Mayes, T., & Ayers, C. R. (2019). Intolerance of uncertainty in hoarding disorder. *Journal of Obsessive-Compulsive and Related Disorders, 21*, 97–101. https://doi.org/10.1016/j.jocrd.2018.11.005.
Cole, D. A., & Maxwell, S. E. (2003). Testing mediational models with longitudinal data: Questions and tips in the use of structural equation modeling. *Journal of Abnormal Psychology, 112*, 558–577. https://doi.org/10.1037/0021-843X.112.4.558.
Comer, J. S., Roy, A. K., Furr, J. M., Gotimer, K., Beidas, R. S., Dugas, M. J., & Kendall, P. C. (2009). The intolerance of uncertainty scale for children: A psychometric evaluation. *Psychological Assessment, 21*, 402–411. https://doi.org/10.1037/a0016719.
Cornacchio, D., Sanchez, A. L., Coxe, S., Roy, A., Pincus, D. B., Read, K. L., ... Comer, J. S. (2018). Factor structure of the intolerance of uncertainty scale for children. *Journal of Anxiety Disorders, 53*, 100–107. https://doi.org/10.1016/j.janxdis.2017.07.003.
Cuijpers, P., Sijbrandij, M., Koole, S. L., Andersson, G., Beekman, A. T., & Reynolds, C. F., 3rd (2013). The efficacy of psychotherapy and pharmacotherapy in treating depressive and anxiety disorders: A meta-analysis of direct comparisons. *World Psychiatry, 12*, 137–148. https://doi.org/10.1002/wps.20038.
Dekkers, L. M. S., Jansen, B. R. J., Salemink, E., & Huizenga, H. M. (2017). Intolerance of uncertainty scale: Measurement invariance among adolescent boys and girls and relationships with anxiety and risk taking. *Journal of Behavior Therapy and Experimental Psychiatry, 55*, 57–65. https://doi.org/10.1016/j.jbtep.2016.11.009.
Dugas, M. J., Gagnon, F., Ladouceur, R., & Freeston, M. H. (1998). Generalized anxiety disorder: A preliminary test of a conceptual model. *Behaviour Research and Therapy, 36*, 215–226. https://doi.org/10.1016/S0005-7967(97)00070-3.
Dugas, M. J., & Robichaud, M. (2007). *Cognitive-behavioral treatment for generalized anxiety disorder: From science to practice.* New York, NY, US: Routledge/Taylor & Francis Group.
Dugas, M. J., Schwartz, A., & Francis, K. (2004). Brief report: Intolerance of uncertainty, worry, and depression. *Cognitive Therapy and Research, 28*, 835–842. https://doi.org/10.1007/s10608-004-0669-0.
Dunsmoor, J. E., Campese, V. D., Ceceli, A. O., LeDoux, J. E., & Phelps, E. A. (2015). Novelty-facilitated extinction: Providing a novel outcome in place of an expected threat diminishes recovery of defensive responses. *Biological Psychiatry, 78*, 203–209. https://doi.org/10.1016/j.biopsych.2014.12.008.
Egger, M., Smith, G. D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *British Medical Journal, 315*, 629–634. https://doi.org/10.1136/bmj.315.7109.629.
Fetzner, M. G., Asmundson, G. J., Carey, C., Thibodeau, M. A., Brandt, C., Zvolensky, M. J., & Carleton, R. N. (2014). How do elements of a reduced capacity to withstand uncertainty relate to the severity of health anxiety? *Cognitive Behaviour Therapy, 43*, 262–274. https://doi.org/10.1080/16506073.2014.929170.
Freeston, M., Rheaume, J., Letarte, H., Dugas, M. J., & Ladouceur, R. (1994). Why do people worry? *Personality and Individual Differences, 17*, 791–802. https://doi.org/10.1016/0191-8869(94)90048-5.
Gentes, E. L., & Ruscio, A. M. (2011). A meta-analysis of the relation of intolerance of uncertainty to symptoms of generalized anxiety disorder, major depressive disorder, and obsessive-compulsive disorder. *Clinical Psychology Review, 31*, 923–933. https://doi.org/10.1016/j.cpr.2011.05.001.
Gomez, R., Summers, M., Summers, A., Wolf, A., & Summers, J. (2014). Depression anxiety stress Scales-21: Measurement and structural invariance across ratings of men and women. *Assessment, 21*, 418–426. https://doi.org/10.1177/1073191113514106.
Grenier, S., Barrette, A.-M., & Ladouceur, R. (2005). Intolerance of uncertainty and intolerance of ambiguity: Similarities and differences. *Personality and Individual Differences, 39*, 593–600. https://doi.org/10.1016/j.paid.2005.02.014.
Hale, W., Richmond, M., Bennett, J., Berzins, T., Fields, A., Weber, D., ... Osman, A. (2016). Resolving uncertainty about the intolerance of uncertainty scale-12: Application of modern psychometric strategies. *Journal of Personality Assessment, 98*, 200–208. https://doi.org/10.1080/00223891.2015.1070355.
Harvey, A., Watkins, E., Mansell, W., & Shafran, R. (2004). *Cognitive behavioral processes across psychological disorders: A transdiagnostic approach to research and treatment.* New York: Oxford University Press Inc.
Hedges, L. V. (1994). Fixed effect models. In H. Cooper, & L. V. Hedges (Eds.). *The handbook of research synthesis* (pp. 285–299). New York: Russell Sage Foundation.
van der Heiden, C., Muris, P., & van der Molen, H. T. (2012). Randomized controlled trial on the effectiveness of metacognitive therapy and intolerance-of-uncertainty therapy for generalized anxiety disorder. *Behaviour Research and Therapy, 50*, 100–109. https://doi.org/10.1016/j.brat.2011.12.005.
Higgins, J. P., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring inconsistency in meta-analyses. *British Medical Journal, 327*, 557–560. https://doi.org/10.1136/bmj.327.7414.557.
Holaway, R. M., Heimberg, R. G., & Coles, M. E. (2006). A comparison of intolerance of uncertainty in analogue obsessive-compulsive disorder and generalized anxiety disorder. *Journal of Anxiety Disorders, 20*, 158–174. https://doi.org/10.1016/j.janxdis.2005.01.002.
Hong, R. Y. (2013). From dispositional traits to psychopathological symptoms: Social-cognitive vulnerabilities as intervening mechanisms. *Journal of Psychopathology and Behavioral Assessment, 35*, 407–420. https://doi.org/10.1007/s10862-013-9350-9.
Hong, R. Y., & Cheung, M. W. L. (2015). The structure of cognitive vulnerabilities to depression and anxiety: Evidence for a common core etiologic process based on a meta-analytic review. *Clinical Psychological Science, 3*, 892–912. https://doi.org/10.1177/2167702614553789.
Huedo-Medina, T. B., Sanchez-Meca, J., Marin-Martinez, F., & Botella, J. (2006). Assessing heterogeneity in meta-analysis: Q statistic or I2 index? *Psychological Methods, 11*, 193–206. https://doi.org/10.1037/1082-989x.11.2.193.
Jacoby, R. J., Abramowitz, J. S., Reuman, L., & Blakey, S. M. (2016). Enhancing the ecological validity of the beads task as a behavioral measure of intolerance of uncertainty. *Journal of Anxiety Disorders, 41*, 43–49. https://doi.org/10.1016/j.janxdis.2016.02.003.
Jacoby, R. J., Fabricant, L. E., Leonard, R. C., Riemann, B. C., & Abramowitz, J. S. (2013). Just to be certain: Confirming the factor structure of the intolerance of uncertainty scale in patients with obsessive-compulsive disorder. *Journal of Anxiety Disorders, 27*, 535–542. https://doi.org/10.1016/j.janxdis.2013.07.008.
Jardin, C., Paulus, D. J., Garey, L., Kauffman, B., Bakhshaie, J., Manning, K., ... Zvolensky, M. J. (2018). Towards a greater understanding of anxiety sensitivity across groups: The construct validity of the anxiety sensitivity Index-3. *Psychiatry Research, 268*, 72–81. https://doi.org/10.1016/j.psychres.2018.07.007.
Jensen, D., Cohen, J. N., Mennin, D. S., Fresco, D. M., & Heimberg, R. G. (2016). Clarifying the unique associations among intolerance of uncertainty, anxiety, and depression. *Cognitive Behaviour Therapy, 45*, 431–444. https://doi.org/10.1080/16506073.2016.1197308.
de Jong-Meyer, R., Beck, B., & Riede, K. (2009). Relationships between rumination, worry, intolerance of uncertainty and metacognitive beliefs. *Personality and Individual Differences, 46*, 547–551. https://doi.org/10.1016/j.paid.2008.12.010.
Kennet-Cohen, T., Kleper, D., & Turvall, E. (2018). Standard errors and confidence intervals for correlations corrected for indirect range restriction: A simulation study comparing analytic and bootstrap methods. *British Journal of Mathematical and Statistical Psychology, 71*, 39–59. https://doi.org/10.1111/bmsp.
Kesby, A., Maguire, S., Brownlow, R., & Grisham, J. R. (2017). Intolerance of uncertainty in eating disorders: An update on the field. *Clinical Psychology Review, 56*, 94–105. https://doi.org/10.1016/j.cpr.2017.07.002.
Koerner, N., Mejia, T., & Kusec, A. (2017). What's in a name? Intolerance of uncertainty, other uncertainty-relevant constructs, and their differential relations to worry and generalized anxiety disorder. *Cognitive Behaviour Therapy, 46*, 141–161. https://doi.org/10.1080/16506073.2016.1211172.
Krohne, H. W. (1989). The concept of coping modes: Relating cognitive person variables to actual coping behavior. *Advances in Behaviour Research and Therapy, 11*, 235–248. https://doi.org/10.1016/0146-6402(89)90027-1.
Ladouceur, R., Gosselin, P., & Dugas, M. J. (2000). Experimental manipulation of intolerance of uncertainty: A study of a theoretical model of worry. *Behaviour Research and Therapy, 38*, 933–941. https://doi.org/10.1016/S0005-7967(99)00133-3.
Lauriola, M., Mosca, O., Trentini, C., Foschi, R., Tambelli, R., & Carleton, R. N. (2018). The intolerance of uncertainty inventory: Validity and comparison of scoring methods to assess individuals screening positive for anxiety and depression. *Frontiers in Psychology, 9*, 388. https://doi.org/10.3389/fpsyg.2018.00388.
Lucas, K., Luck, C. C., & Lipp, O. V. (2018). Novelty-facilitated extinction and the reinstatement of conditional human fear. *Behaviour Research and Therapy, 109*, 68–74. https://doi.org/10.1016/j.brat.2018.08.002.
Mahoney, A. E., & McEvoy, P. M. (2012a). Trait versus situation-specific intolerance of uncertainty in a clinical sample with anxiety and depressive disorders. *Cognitive Behaviour Therapy, 41*, 26–39. https://doi.org/10.1080/16506073.2011.622131.
Mahoney, A. E., & McEvoy, P. M. (2012b). A transdiagnostic examination of intolerance of uncertainty across anxiety and depressive disorders. *Cognitive Behaviour Therapy, 41*, 212–222. https://doi.org/10.1080/16506073.2011.622130.
Maxwell, S. E., & Cole, D. A. (2007). Bias in cross-sectional analyses of longitudinal mediation. *Psychological Methods, 12*, 23–44. https://doi.org/10.1037/1082-989X.12.1.23.
McEvoy, P. M., Carleton, R. N., Correa, K., Shankman, S. A., & Shihata, S. (2019). Intolerance of uncertainty. In B. O. Olatunji (Ed.). *The Cambridge handbook of anxiety and related disorders*. Cambridge, UK: Cambridge University Press.
McEvoy, P. M., & Erceg-Hurn, D. M. (2016). The search for universal transdiagnostic and

trans-therapy change processes: Evidence for intolerance of uncertainty. *Journal of Anxiety Disorders, 41*, 96–107. https://doi.org/10.1016/j.janxdis.2016.02.002.
McEvoy, P. M., & Mahoney, A. E. (2011). Achieving certainty about the structure of intolerance of uncertainty in a treatment-seeking sample with anxiety and depression. *Journal of Anxiety Disorders, 25*, 112–122. https://doi.org/10.1016/j.janxdis.2010.08.010.
McEvoy, P. M., Mahoney, A. E., & Moulds, M. L. (2010). Are worry, rumination, and post-event processing one and the same? Development of the repetitive thinking questionnaire. *Journal of Anxiety Disorders, 24*, 509–519. https://doi.org/10.1016/j.janxdis.2010.03.008.
McEvoy, P. M., & Mahoney, A. E. J. (2012). To be sure, to be sure: Intolerance of uncertainty mediates symptoms of various anxieyt disorders and depression. *Behavior Therapy, 43*, 533–545. https://doi.org/10.1016/j.beth.2011.02.007.
Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & The PRISMA Group (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS Medicine, 6*, e1000097. https://doi.org/10.1371/journal.pmed1000097.
Nelson, B. D., Liu, H., Sarapas, C., & Shankman, S. A. (2016). Intolerance of uncertainty mediates the relationship between panic and the startle reflex in anticipation of unpredictable threat. *Journal of Experimental Psychopathology, 7*, 172–189. https://doi.org/10.5127/jep.048115.
Norton, P. J. (2005). A psychometric analysis of the intolerance of uncertainty scale among four racial groups. *Journal of Anxiety Disorders, 19*, 699–707. https://doi.org/10.1016/j.janxdis.2004.08.002.
Norton, P. J., & Mehta, P. D. (2007). Hierarchical model of vulnerabilities for emotional disorders. *Cognitive Behaviour Therapy, 36*, 240–254. https://doi.org/10.1080/16506070701628065.
Norton, P. J., & Price, E. C. (2007). A meta-analytic review of adult cognitive-behavioral treatment outcome across the anxiety disorders. *Journal of Nervous and Mental Disease, 195*, 521–531. https://doi.org/10.1097/01.nmd.0000253843.70149.9a.
Norton, P. J., Sexton, K. A., Walker, J. R., & Norton, G. R. (2005). Hierarchical model of vulnerabilities for anxiety: Replication and extension with a clinical sample. *Cognitive Behaviour Therapy, 34*, 50–63. https://doi.org/10.1080/16506070410005401.
Obesessive Compulsive Cognitions Working Group (OCCWG) (2003). Psychometric validation of the obsessive beliefs questionnaire and the interpretation of intrusions inventory: Part I. *Behaviour Research and Therapy, 41*, 863–878. https://doi.org/10.1016/S0005-7967(02)00099-2.
Obesessive Compulsive Cognitions Working Group (OCCWG) (2005). Psychometric validation of the obsessive belief questionnaire and interpretation of intrusions inventory—Part 2: Factor analyses and testing of a brief version. *Behaviour Research and Therapy, 43*, 1527–1542. https://doi.org/10.1016/j.brat.2004.07.010.
Oglesby, M. E., Boffa, J. W., Short, N. A., Raines, A. M., & Schmidt, N. B. (2016). Intolerance of uncertainty as a predictor of post-traumatic stress symptoms following a traumatic event. *Journal of Anxiety Disorders, 41*, 82–87. https://doi.org/10.1016/j.janxdis.2016.01.005.
Oglesby, M. E., Medley, A. N., Norr, A. M., Capron, D. W., Korte, K. J., & Schmidt, N. B. (2013). Intolerance of uncertainty as a vulnerability factor for hoarding behaviors. *Journal of Affective Disorders, 145*, 227–231. https://doi.org/10.1016/j.jad.2012.08.003.
Olatunji, B. O., Broman-Fulks, J. J., Bergman, S. M., Green, B. A., & Zlomke, K. R. (2010). A taxometric investigation of the latent structure of worry: Dimensionality and associations with depression, anxiety, and stress. *Behavior Therapy, 41*, 212–228. https://doi.org/10.1016/j.beth.2009.03.001.
Olatunji, B. O., Cisler, J. M., & Tolin, D. F. (2010). A meta-analysis of the influence of comorbidity on treatment outcome in the anxiety disorders. *Clinical Psychology Review, 30*, 642–654. https://doi.org/10.1016/j.cpr.2010.04.008.
Olatunji, B. O., Williams, B. J., Haslam, N., Abramowitz, J. S., & Tolin, D. F. (2008). The latent structure of obsessive-compulsive symptoms: A taxometric study. *Depression and Anxiety, 25*, 956–968. https://doi.org/10.1002/da.20387.
Orwin, R. G. (1983). A fail-safe N for effect size in meta-analysis. *Journal of Educational Statistics, 8*, 157–159. https://doi.org/10.2307/1164923.
Păsărelu, C. R., Dobrean, A., Balazsi, R., Predescu, E., Şipos, R., & Lupu, V. (2017). The Penn State worry questionnaire for children: Age, gender and clinical invariance. *Child Psychiatry and Human Development, 48*, 359–369. https://doi.org/10.1007/s10578-016-0663-2.
Penelo, E., Negrete, A., Portell, M., & Raich, R. M. (2013). Psychometric properties of the eating disorder examination questionnaire (EDE-Q) and norms for rural and urban adolescent males and females in Mexico. *PLoS ONE, 8*, e83245. https://doi.org/10.1371/journal.pone.0083245.
Renjan, V., McEvoy, P. M., Handley, A. K., & Fursland, A. (2016). Stomaching uncertainty: Relationships among intolerance of uncertainty, eating disorder pathology, and comorbid emotional symptoms. *Journal of Anxiety Disorders, 41*, 88–95. https://doi.org/10.1016/j.janxdis.2016.03.008.
Rodríguez-Jiménez, T., Godoy, A., Piqueras, J. A., Gavino, A., Martínez-González, A. E., & Foa, E. B. (2017). Factor structure and measurement invariance of the obsessive-compulsive inventory – Child version (OCI-CV) in general population. *European Journal of Psychological Assessment, 33*, 97–103. https://doi.org/10.1027/1015-5759/a000276.
Roma, V. G., & Hope, D. A. (2017). Are we certain about which measure of intolerance of uncertainty to use yet? *Cognitive Behaviour Therapy, 46*, 327–338. https://doi.org/10.1080/16506073.2016.1252792.
Rosenthal, R. (1979). The 'file-drawer problem' and tolerance for null results. *Psychological Bulletin, 86*, 638–641. https://doi.org/10.1037/0033-2909.86.3.638.
Rosser, B. A. (2019). Intolerance of uncertainty as a transdiagnostic mechanism of psychological difficulties: A systematic review of evidence pertaining to causality and temporal precedence. *Cognitive Therapy and Research, 43*, 438–463. https://doi.org/10.1007/s10608-018-9964-z.
Rothstein, H. R., Sutton, A. J., & Borenstein, M. (2005). *Publication bias in meta-analysis: Prevention, assesment and adjustments.* Chichester, UK: John Wiley & Sons, Ltd.
Ruscio, A. M., Borkovec, T. D., & Ruscio, J. (2001). A taxometric investigation of the latent structure of worry. *Journal of Abnormal Psychology, 110*, 413–422. https://doi.org/10.1037/0021-843X.110.3.413.
Ruscio, A. M., Ruscio, J., & Keane, T. M. (2002). The latent structure of posttraumatic stress disorder: A taxometric investigation of reactions to extreme stress. *Journal of Abnormal Psychology, 111*, 290–301. https://doi.org/10.1037/0021-843X.111.2.290.
Segerstrom, S. C., Tsao, J. C. I., Alden, L. E., & Craske, M. G. (2000). Worry and rumination: Repetitive thought as a concomitant and predictor of negative mood. *Cognitive Therapy and Research, 24*, 671–688. https://doi.org/10.1023/A:1005587311498.
Sexton, K. A., Norton, P. J., Walker, J. R., & Norton, G. R. (2003). Hierarchical model of generalized and specific vulnerabilities in anxiety. *Cognitive Behaviour Therapy, 32*, 82–94. https://doi.org/10.1080/16506070302321.
Shihata, S., McEvoy, P. M., Mullan, B., & Carleton, R. N. (2016). Intolerance of uncertainty in emotional disorders: what uncertainties remain? *Journal of Anxiety Disorders, 41*, 115–124. https://doi.org/10.1016/j.janxdis.2016.05.001.
Shihata, S., McEvoy, P. M., & Mullan, B. A. (2017). Pathways from uncertainty to anxiety: An evaluation of a hierarchical model of trait and disorder-specific intolerance of uncertainty on anxiety disorder symptoms. *Journal of Anxiety Disorders, 45*, 72–79. https://doi.org/10.1016/j.janxdis.2016.12.001.
Shihata, S., McEvoy, P. M., & Mullan, B. A. (2018). A bifactor model of intolerance of uncertainty in undergraduate and clinical samples: Do we need to reconsider the two-factor model? *Psychological Assessment, 30*, 893–903. https://doi.org/10.1037/pas0000540.
Sterne, J. A. C., Becker, B. J., & Egger, M. (2005). The funnel plot. In H. R. Rothstein, A. J. Sutton, & M. Borenstein (Eds.). *Publication bias in meta-analysis: Prevention, assessment and adjustments* (pp. 75–98). Chichester, England: John Wiley & Sons.
Sternheim, L., Startup, H., & Schmidt, U. (2011). An experimental exploration of behavioral and cognitive-emotional aspects of intolerance of uncertainty in eating disorder patients. *Journal of Anxiety Disorders, 25*, 806–812. https://doi.org/10.1016/j.janxdis.2011.03.020.
Suárez-Falcón, J. C., Monroy-Cifuentes, A., & Ruiz, F. J. (2018). Penn State worry Questionnaire-11 validity in Colombia and factorial equivalence across gender and nonclinical and clinical samples. *Anales de Psicología, 34*, 451–457. https://doi.org/10.6018/analesps.34.3.300281.
Summers, B. J., Matheny, N. L., Sarawgi, S., & Cougle, J. R. (2016). Intolerance of uncertainty in body dysmorphic disorder. *Body Image, 16*, 45–53. https://doi.org/10.1016/j.bodyim.2015.11.002.
Tanovic, E., Gee, D. G., & Joormann, J. (2018). Intolerance of uncertainty: Neural and psychophysiological correlates of the perception of uncertainty as threatening. *Clinical Psychology Review, 60*, 87–99. https://doi.org/10.1016/j.cpr.2018.01.001.
Thibodeau, M. A., Carleton, R. N., McEvoy, P. M., Zvolensky, M. J., Brandt, C. P., Boelen, P. A., ... Asmundson, G. J. (2015). Developing scales measuring disorder-specific intolerance of uncertainty (DSIU): A new perspective on transdiagnostic. *Journal of Anxiety Disorders, 31*, 49–57. https://doi.org/10.1016/j.janxdis.2015.01.006.
Tolin, D. F., Abramowitz, J. S., Brigidi, B. D., & Foa, E. B. (2003). Intolerance of uncertainty in obsessive-compulsive disorder. *Journal of Anxiety Disorders, 17*, 233–242. https://doi.org/10.1016/S0887-6185(02)00182-2.
Whisman, M. A., Judd, C. M., Whiteford, N. T., & Gelhorn, H. L. (2013). Measurement invariance of the Beck depression inventory–second edition (BDI-II) across gender, race, and ethnicity in college students. *Assessment, 20*, 419–428.
Whiting, S. E., Jenkins, W. S., May, A. C., Rudy, B. M., Davis, T. E., 3rd, & Reuther, E. T. (2014). The role of intolerance of uncertainty in social anxiety subtypes. *Journal of Clinical Psychology, 70*, 260–272. https://doi.org/10.1002/jclp.22024.
Wright, K. D., Lebell, M. A., & Carleton, R. N. (2016). Intolerance of uncertainty, anxiety sensitivity, health anxiety, and anxiety disorder symptoms in youth. *Journal of Anxiety Disorders, 41*, 35–42. https://doi.org/10.1016/j.janxdis.2016.04.011.
Yook, K., Kim, K. H., Suh, S. Y., & Lee, K. S. (2010). Intolerance of uncertainty, worry, and rumination in major depressive disorder and generalized anxiety disorder. *Journal of Anxiety Disorders, 24*, 623–628. https://doi.org/10.1016/j.janxdis.2010.04.00.