



## Assessing the validity and reliability of the International Anxiety Questionnaire and the International Depression Questionnaire in two bereaved national samples

Hyland, P., Redican, E., Karatzias, T., & Shevlin, M. (2023). Assessing the validity and reliability of the International Anxiety Questionnaire and the International Depression Questionnaire in two bereaved national samples. *Journal of Clinical Psychology and Psychotherapy*, 1-12. Advance online publication. <https://doi.org/10.1002/cpp.2917>

[Link to publication record in Ulster University Research Portal](#)

### Publication Status:

Published online: 29/09/2023

### DOI:

[10.1002/cpp.2917](https://doi.org/10.1002/cpp.2917)

### Document Version

Publisher's PDF, also known as Version of record

### General rights

Copyright for the publications made accessible via Ulster University's Research Portal is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

### Take down policy

The Research Portal is Ulster University's institutional repository that provides access to Ulster's research outputs. Every effort has been made to ensure that content in the Research Portal does not infringe any person's rights, or applicable UK laws. If you discover content in the Research Portal that you believe breaches copyright or violates any law, please contact [pure-support@ulster.ac.uk](mailto:pure-support@ulster.ac.uk).

## RESEARCH ARTICLE

WILEY

# Assessing the validity and reliability of the International Anxiety Questionnaire and the International Depression Questionnaire in two bereaved national samples

Philip Hyland<sup>1</sup> | Enya Redican<sup>2</sup> | Thanos Karatzias<sup>3</sup>  | Mark Shevlin<sup>2</sup>

<sup>1</sup>Department of Psychology, Maynooth University, Maynooth, Ireland

<sup>2</sup>School of Psychology, Ulster University, Ulster, UK

<sup>3</sup>School of Health & Social Care, Napier University, Edinburgh, UK

## Correspondence

Professor Philip Hyland, Department of Psychology, Maynooth University, Maynooth, Ireland.

Email: [philip.hyland@mu.ie](mailto:philip.hyland@mu.ie)

## Abstract

The International Anxiety Questionnaire (IAQ) and International Depression Questionnaire (IDQ) are self-report measures of ICD-11 Generalized Anxiety Disorder (ICD-11 GAD) and ICD-11 Single Episode Depressive Disorder (ICD-11 DD). This study tested the psychometric properties of these scales in two samples of bereaved adults from the United Kingdom and the Republic of Ireland. Confirmatory factor analysis (CFA) was used to test the combined dimensionality and measurement invariance of the IAQ and IDQ across the United Kingdom ( $n = 1012$ ) and Irish ( $n = 1011$ ) samples. Differential item functioning (DIF) was tested using multiple indicator multiple cause (MIMIC) modelling while convergent validity was also assessed. CFA results supported a correlated two-factor model in both samples. The MIMIC model showed that the IDQ item “Had recurrent thoughts of death or suicide” showed DIF and the effect was small. Internal reliability of the scales were high and convergent validity was supported. The prevalence of ICD-11 GAD was 18.6% and 16.1% and ICD-11 DD was 13.8% and 10.5% in the United Kingdom and Irish samples, respectively. Findings of the study provide support for the validity, measurement invariance, and reliability of the IAQ and IDQ among two bereaved national samples.

## KEYWORDS

generalized anxiety disorder, ICD-11, reliability, single episode depressive disorder, validity

## 1 | INTRODUCTION

The eleventh version of the *International Classification of Diseases* (ICD-11: World Health Organization [WHO], 2022) includes updated diagnostic guidelines for anxiety and depressive disorders. Generalized Anxiety Disorder (ICD-11 GAD: code 6B00) is defined by the presence of general apprehension or excessive worry occurring for more days than not, for at least several months, along with other problems (e.g., muscle tension and abdominal

discomfort) that together cause significant distress or impairment. Single Episode Depressive Disorder (ICD-11 DD: code 6A70) is defined by the presence of depressed mood or diminished interest in activities occurring for most of the day, nearly every day, for at least 2 weeks, along with other problems (e.g., feelings of worthlessness and recurrent thoughts of death) that together cause significant distress or impairment. The ICD-11 descriptions of GAD and DD are now the de facto diagnostic models of anxiety and depression.

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial-NoDerivs](https://creativecommons.org/licenses/by-nc-nd/4.0/) License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2023 The Authors. *Clinical Psychology & Psychotherapy* published by John Wiley & Sons Ltd.

There are many empirically supported self-report measures of anxiety—for example, the Beck Anxiety Inventory (Beck et al., 1988) and the GAD-7 (Spitzer et al., 2006)—and depression—for example, the Centre for Epidemiological Studies-Depression scale (Radloff, 1977) and the Patient Health Questionnaire-9 (Kroenke et al., 2001)—that are exceptionally useful for research and clinical practice, but none were designed to capture the symptoms and diagnostic requirements of ICD-11 GAD and DD. Shevlin, Butter, McBride, et al. (2022) recently developed the *International Anxiety Questionnaire* (IAQ) and the *International Depression Questionnaire* (IDQ) to do exactly that. Using a nationally representative sample of adults from the general population of the United Kingdom, Shevlin et al. found that the latent structure of each measure was unidimensional, the items tapped the underlying dimensions at clinically meaningful levels, that each scale produced scores with high internal reliability, each had the capacity to discriminate between individuals with and without a history of treatment for mental health problems, and scores on both measures correlated positively and strongly with established measures of anxiety and depression symptoms. Furthermore, application of the ICD-11 diagnostic guidelines yielded prevalence estimates of 7.1% for GAD and 7.4% for DD. Initial psychometric testing has therefore been promising, but further research is required.

In this study, we further tested the reliability and validity of the IAQ and IDQ scores in two samples of bereaved adults from the United Kingdom and the Republic of Ireland. Similar to the initial validation study, the aims are to (1) describe the distribution of item and scale level scores and (2) assess construct validity using factor analytic methods. This study recruited bereaved participants because research has shown that bereavement increases the probability of the onset of multiple mental health problems including depression (e.g., Pham et al., 2018; Zisook & Kendler, 2007) and anxiety (e.g., Shear & Skritskaya, 2012; Thimm et al., 2020). This study also addressed the important psychometric issue of measurement invariance; this refers to the extent that a scale measures the same construct, in the same way, for different people. The ability to make valid comparison of anxiety and depression, for example, between different age groups or countries, is based on the assumption that the items contained within these scales operate equivalently for these different groups of interest, or that the item performance is “invariant” (Shevlin, Butter, McBride, et al., 2022). There has been extensive invariance testing of other measures of anxiety and depression such as the DASS-21 (Scholten et al., 2017) and the PHQ-9/GAD-7 (Shevlin, Butter, McBride, et al., 2022).

In this study, we hypothesized that (1) a correlated 2-factor model of the IAQ and the IDQ would provide acceptable fit to the sample data, (2) that scores on the IAQ and IDQ would have high internal reliability, (3) that there would be no differential item functioning associated with age, gender, or nationality (UK/Ireland), (4) that scores on the IAQ and IDQ would be strongly and positively correlated with an independent measure of anxiety and depression symptoms as well as a measure of prolonged grief disorder symptoms, and (5) the IAQ and IDQ scores would distinguish between

### Key Practitioner Message

- This study demonstrates that the IAQ and IDQ are psychometrically sound measures of ICD-11 GAD and DD and can assess anxiety and depression, respectively, in the same manner regardless of country, sex, and age.
- This study demonstrates that a substantial minority of bereaved individuals in the United Kingdom and Irish populations meet criteria for diagnosis of probable ICD-11 GAD and DD. These findings emphasize the mental health burden of bereavement and the need for routine screening of anxiety and depression following bereavement.
- This study highlights that the co-occurrence of anxiety and depression is common and that the IAQ and IDQ can be used to assess ICD-11 mixed depressive and anxiety disorder. Hence, clinicians may wish to deploy the IDQ and IAQ to identify the presence of depression and anxiety, respectively or to assess for a combination of both.
- This study highlights the need for the development of a clinician-administered measure of ICD-11 GAD and DD. However, the sound psychometric properties of the IAQ and IDQ support their use to assess for symptoms of anxiety and depression, respectively.

those individuals with and without a history of treatment for mental health problems.

## 2 | METHODS

### 2.1 | Participants and procedures

Data were collected from a sample of bereaved adults from the United Kingdom (UK:  $N = 1012$ ) and the Republic of Ireland (Ireland:  $N = 1011$ ) using identical procedures. The survey company Qualtrics was employed to recruit participants in each nation from existing, actively managed, double-opt-in research panels via email, SMS, or in-app notifications. Inclusion criteria were that respondents were aged 18 years or older, were residing in the United Kingdom or Ireland, respectively, could complete the survey in English, and had answered “Yes” to the following question that screened for lifetime bereavement: “During your life have you known anyone who has died (e.g., a partner, parent, child, close friend)?” An a priori power analysis was conducted to determine the optimal sample size needed to detect prolonged grief disorder with an assumed prevalence rate of 2.4% among bereaved adults (see Shevlin et al., 2023, for justification), with a precision of 1%, and a confidence level of 95% (optimal sample size was not determined by the IAQ and IDQ due to the data used in the current study belonging to an overall project largely focused on PGD). This resulted in a required sample size of  $N = 900$ . Given the size of

the respective populations of the United Kingdom and Ireland and the available panel members in each nation, we set our target sample sizes for 1000 participants in both the United Kingdom and Ireland. The UK data were collected between 19 April and 13 August 2022, and the Irish data were collected from 21 April 2022 and 12 September 2022. Ethical approval for the collection of all data was provided by the research ethics committee at Ulster University (Reference number: FCPSY-22-026-A). Quota sampling was used to select participants in a manner that resulted in the composition of the UK and Irish samples being closely aligned to the sex, age, and regional distributions of the respective populations. Key demographic information for each sample is presented in Table 1.

## 2.2 | Materials

ICD-11 GAD: The IAQ (Shevlin, Butter, McBride, et al., 2022) is a self-report measure designed to capture all diagnostic requirements for ICD-11 GAD. Participants answer eight questions that assess the two “essential” (questions 1 and 2) and six ‘accompanying’ (questions 3 to 8) symptoms of GAD, based on the following instruction: “**Over the last several months**, how frequently have you had the following feelings, thoughts, and behaviours?” These questions are answered using a 5-point Likert scale where 0 = *Never*, 1 = *Only a few days*, 2 = *Half the days*, 3 = *Most days*, and 4 = *Every day*. Symptom endorsement is based on responses of 3 or 4 on the Likert scale. There is an additional question measuring functional impairment (“Have these experiences caused problems in personal, family, social, educational, occupational, or other important areas of your life?”) that is answered on a “Yes” or “No” basis. The IAQ can be used to measure symptom severity or to identify probable diagnostic status. The severity scoring method involves summing responses to the eight questions, producing possible scores ranging from 0 to 32. The diagnostic criteria for ICD-11 GAD requires that four or more symptoms be endorsed with at least one from questions 1 or 2 (i.e., the essential symptoms), and functional impairment is also endorsed.

ICD-11 DD: The IDQ (Shevlin, Butter, McBride, et al., 2022) is a self-report measure designed to capture all diagnostic requirements for ICD-11 DD. Participants answer nine questions assessing the two “essential” (questions 1 and 2) and seven “accompanying” (questions 3 to 8) symptoms of DD based on the following instruction: “**Over the last two weeks**, how frequently have you had the following feelings, thoughts, and behaviours?” These questions are answered using a 5-point Likert scale where 0 = *Never*, 1 = *Only a few days*, 2 = *Half the days*, 3 = *Most days*, and 4 = *Every day*. Symptom endorsement is based on responses of 3 or 4 on the Likert scale. There is an additional question measuring functional impairment (“Have these experiences caused problems in personal, family, social, educational, occupational, or other important areas of your life?”) that is answered on a “Yes” or “No” basis. The IDQ can be used to measure symptom severity or to identify probable diagnostic status. The severity scoring method involves summing responses to the nine questions, producing possible

scores ranging from 0 to 36. The diagnostic criteria for ICD-11 DD requires that five or more symptoms be endorsed with at least one being from question 1 or 2 (i.e., the essential symptoms), and functional impairment is also endorsed.

Anxiety and depression symptoms: The Patient Health Questionnaire-4 (PHQ-4; Kroenke et al., 2009) was used as a concurrent measure of anxiety and depression symptoms. The PHQ-4 includes the two core items measuring depression symptoms from the PHQ-9 and the two core items measuring generalized anxiety symptoms from the GAD-7. As with the parent scales, the PHQ-4 includes the stem question: “Over the last 2 weeks, how often have you been bothered by the following problems?” A 4-point Likert response scale is used where 0 = *not at all*, 1 = *several days*, 2 = *more than half the days*, and 3 = *nearly every day*, and possible scores range from 0 to 12. Previous research has shown that the PHQ-4 produces reliable and valid scores in general population samples (e.g., Löwe et al., 2010), and the internal reliability of the scale scores in the UK ( $\alpha = .93$ ) and Irish ( $\alpha = .93$ ) samples were excellent.

Prolonged grief symptoms: The International Prolonged Grief Disorder Scale (IPGDS; Killikelly et al., 2021) is a self-report measure aligned to the ICD-11 description of Prolonged Grief Disorder. The IPGDS includes two items measuring the core symptoms of longing for the deceased and pre-occupation with the deceased, and 10 items measuring different forms of emotional pain associated with bereavement (e.g., intense feelings of sorrow, anger, guilt, and being unable to accept the loss). Participants indicate the frequency of these symptoms over the past week on a 5-point Likert scale ranging from 0 (*Not at all*) to 4 (*Extremely*). Scores can range from 0 to 48 with higher scores reflecting higher symptoms of prolonged grief. The psychometric properties of the IPGDS scale scores have been supported in several national samples (Killikelly et al., 2020, 2021), and the internal reliability of the scale scores in the UK ( $\alpha = .94$ ) and Irish ( $\alpha = .92$ ) samples were excellent.

Treatment for mental health problems: Participants were asked to provide information about their current or past treatment for a mental health problem. They received the following statement: “Mental health difficulties are very common. It will help us understand our survey results if you would tell us whether you currently or have in the past received treatment (medication or talking therapies) for these kinds of difficulties.” Three response options were provided including “I have never received treatment for mental health problems,” “I have received treatment for mental health problems in the past,” and “I am currently receiving treatment for mental health problems.”

## 2.3 | Data analysis

The analyses were conducted in five phases. First, descriptive statistics for the individual items and the summed scores on the IAQ and the IDQ were calculated, and cross-country differences were tested.

**TABLE 1** Demographic characteristics of the UK and Irish samples.

Ireland (N = 1011)	%	UK (N = 1012)	%
<b>Sex</b>		<b>Sex</b>	
Female	52.5	Female	51.3
Male	47.3	Male	47.9
<b>Age</b>		<b>Age</b>	
18–24	9.2	18–24	9.5
25–34	20.3	25–34	20.2
35–44	20.1	35–44	19.9
45–54	18.8	45–54	18.8
55+	31.7	55+	31.7
<b>Place of birth</b>		<b>Place of birth</b>	
Ireland	76.9	UK	94.6
<b>Living location</b>		<b>Living location</b>	
City	20.5	City	23.6
Suburb	22.8	Suburb	25.6
Town	25.5	Town	32.4
Rural area	31.2	Rural area	18.4
<b>Income</b>		<b>Income</b>	
0–€19,999	24.9	0–£19,999	36.5
€20,000–€39,999	38.9	£20,000–£39,999	44.0
€40,000–€59,999	21.3	£40,000–£59,999	13.7
€60,000–€79,000	9.6	£60,000–£79,000	4.2
€80,000 or more	5.3	£80,000 or more	1.7
<b>Ethnicity</b>		<b>Ethnicity</b>	
Irish	76.5	British	79.8
British/Irish	11.5	British/Irish	12.0
Indian	0.8	Indian	1.1
Pakistani	1.7	Pakistani	1.7
Chinese	2.0	Chinese	0.8
Other Asian	0.4	Other Asian	0.3
African	1.6	African	0.7
Other ethnic group	8.4	Afro-Caribbean	0.1
Arab	0.2	Other ethnic group	3.4
Bangladeshi	0.1	Arab	0.2
<b>Education</b>		<b>Education</b>	
No qualifications	0.7	No qualifications	3.7
Finished mandatory schooling	6.3	GCSE or similar	23.0
Finished secondary school	22.2	A-level or similar	20.3
Technical qualification	15.2	Technical qualification	22.8
Undergraduate degree	20.4	Undergraduate degree	22.8
Diploma	12.3	Diploma	4.2
Postgraduate degree	21.7	Postgraduate degree	14.4
Other qualifications	1.3	Other qualification	2.2
<b>Employment</b>		<b>Employment</b>	
Full-time	52.3	Full-time	49.5
Part-time	16.3	Part-time	15.6
Unemployed	10.9	Unemployed	12.4

TABLE 1 (Continued)

Ireland (N = 1011)	%	UK (N = 1012)	%
Retired	13.0	Retired	15.9
Student	4.3	Student	2.4
Disabled	3.2	Disabled	4.2
<b>Religion</b>		<b>Religion</b>	
Christian	71.7	Christian	55.0
Muslim	1.8	Muslim	3.2
Jewish	0.3	Jewish	0.6
Hindu	0.7	Hindu	1.0
Buddhist	0.8	Buddhist	0.4
Atheist	13.2	Atheist	24.8
Agnostic	7.1	Agnostic	8.6
Other religion	4.5	Other religion	6.3
		Sikh	0.1

Second, a confirmatory factor analysis (CFA) model of the IAQ and the IDQ indicators was estimated to establish the fit of a baseline model for the United Kingdom and Ireland separately. The model specified two correlated latent variables, with the IAQ items loading on an “Anxiety” latent variable and the IDQ items loading on a “Depression” latent variable. The data from both countries were then combined and tests of configural and metric invariance were conducted: Configural invariance tests that the latent structure (i.e., a correlated two-factor model) is consistent across the groups, and metric invariance adds constraints to assess for the equality of factor loadings across the groups. Scalar invariance was not assessed as differences in the intercepts were assessed as part of the DIF analysis.

Third, a MIMIC model based on the exogenous predictor variables of country, age, and sex was specified to test for DIF on the IAQ/IDQ items. The presence of any direct effects between predictor variables and the IAQ/IDQ items after adjusting for variability on latent variables is indicative of DIF. As the country and sex variables are binary in nature the regression coefficients for these are reported as “y-standardized,” where the latent variable (y) is standardized, but the predictor variables retain their binary scoring. For age, the fully standardized effects are reported.

Whether a direct effect was to be included was determined using modification indices (MIs) (Sörbom, 1989) and the standardized expected parameter change (SEPCs) (Chou & Bentler, 1993; Saris et al., 1987) values. MIs indicate which path, if added to the model, would significantly improve model fit should it be freely estimated. Improved model fit is indicated by a reduction of 3.84 or more in the chi-square (this is the critical value for the chi-square for one degree of freedom,  $p < .05$ ). In the present study, a more conservative value of 10 was used to avoid the addition of small insignificant parameters, and this is reflected in Mplus only reporting MIs greater than 10. The SEPC indicates the estimated value of a fixed parameter (in this case fixed to zero) if it were estimated, that

is, the expected standardized regression coefficient. The MIs are influenced by sample size (Saris et al., 1987), and therefore, it has been recommended that the selection of which parameters should be added to the model be based on a combination of MIs and SEPCs (Kaplan, 1989). Thus, in this study, a direct effect from the predictor to an IAQ/IDQ item would be added if the MI was greater than 10 and the SEPC was greater than .20. A process followed whereby the path with the largest MI/SEPC was freely estimated in the model and the model was re-estimated. This continued until there were no MIs/SEPCs greater than 10/.20.

All analyses were conducted in Mplus 8.1 (Muthén & Muthén, 2017) and all models were estimated using robust maximum likelihood estimation (MLR) (Yuan & Bentler, 2000). Numerous fit statistics were used to evaluate the goodness of fit for each model: the Chi-square, the comparative fit index (CFI) (Bentler, 1990), and the Tucker-Lewis Index (TLI) (Tucker & Lewis, 1973). A nonsignificant chi-square and CFI and TLI values  $\geq .90$  and  $\geq .95$  were considered as good and excellent model fit. Additionally, the Root Mean Square Error of Approximation (RMSEA) (Steiger, 1990) was reported, where a value less than .05 indicated close fit and values up to .08 indicated reasonable errors of approximation. The same cut-off values can be used for the standardized root mean square residual (SRMR) (Jöreskog & Sörbom, 1996). To compare the configural and metric models of invariance the criteria proposed by Chen (2007) were used: less than .010 change in CFI, less than .015 in RMSEA, and less than .030 for the SRMR.

Fourth, bivariate associations between the IAQ and IDQ summed scores and summed scores on the PHQ-4 and the IPGDS were assessed using a Pearson correlation test. Finally, factorial between groups analysis of variance (ANOVA) tests, with Bonferroni post-hoc tests, were used to compare mean IAQ and IDQ scores across the two-level country variable (UK/Ireland) and three-level mental health treatment variable (Never, Past, Current). Effect sizes are reported as eta-squared values ( $\eta^2$ ) where values up to .06



**TABLE 2** Mean and endorsement rates (%) for IDQ and IAQ.

Item	IDQ				Item	IAQ			
	UK	Ire	t (df)	p		UK	Ire	t (df)	p
1	1.25 (1.20)	1.05 (1.12)	3.93 (2012.30)	<.001	1	1.61 (1.28)	1.47 (1.22)	2.55 (2015.68)	.011
2	1.28 (1.19)	1.03 (1.08)	4.83 (2001.79)	<.001	2	1.77 (1.28)	1.68 (1.24)	1.63 (2018.81)	.103
3	1.39 (1.27)	1.24 (1.22)	2.81 (2017.66)	.002	3	1.50 (1.28)	1.35 (1.21)	2.82 (2014.84)	.005
4	1.20 (1.31)	0.95 (1.20)	4.47 (2004.19)	<.001	4	1.22 (1.28)	1.01 (1.18)	3.72 (2007.44)	<.001
5	1.22 (1.34)	0.98 (1.21)	4.20 (2002.67)	<.001	5	1.46 (1.29)	1.20 (1.21)	4.75 (2012.55)	<.001
6	0.79 (1.19)	0.52 (0.97)	5.61 (1942.11)	<.001	6	1.51 (1.30)	1.40 (1.24)	1.94 (2015.58)	.053
7	1.24 (1.29)	0.95 (1.13)	5.46 (1984.33)	<.001	7	1.55 (1.27)	1.45 (1.16)	1.96 (2004.84)	.050
8	1.26 (1.30)	1.07 (1.15)	3.44 (1990.84)	<.001	8	1.69 (1.35)	1.42 (1.22)	4.63 (2002.27)	<.001
9	1.55 (1.36)	1.40 (1.24)	2.71 (2003.95)	.003					
Total	11.19 (9.89)	9.19 (8.61)	4.85 (1983.71)	<.001	Total	12.32 (8.99)	10.99 (9.19)	3.48 (2008.74)	<.001
FI	.38 (.49)	.35 (.48)	1.46 (2020.41)	.072	FI	.39 (.48)	.36 (.48)	1.13 (2020.70)	.131

Abbreviations: IAQ, International Anxiety Questionnaire; IDQ, International Depression Questionnaire; FI, functional impairment.

indicate a “small effect,” values from .06 to .13 indicate a “medium” effect, and values of .14 and above indicate a “large” effect (Cohen, 1988).

### 3 | RESULTS

#### 3.1 | Endorsement rates results

Mean item and scale scores for the IDQ and IAQ are presented in Table 2. The IDQ items with the highest mean scores in the UK sample were item 9 *experienced reduced energy or fatigue* ( $M = 1.55$ ,  $SD = 1.36$ ), item 3 *Have difficulty concentrating* ( $M = 1.39$ ,  $SD = 1.27$ ), and item 2 *Experienced less interest or pleasure from normal activities* ( $M = 1.28$ ,  $SD = 1.19$ ). The IDQ items with the highest mean scores in the Irish sample were item 9 ( $M = 1.40$ ,  $SD = 1.24$ ), item 3 ( $M = 1.24$ ,  $SD = 1.22$ ), and item 8 *moved slower or felt more restless* ( $M = 1.07$ ,  $SD = 1.15$ ). The UK sample had higher mean scores on all IDQ items as compared to the Irish sample.

The IAQ items with the highest mean scores in the UK sample were item 2 *Worried a lot about different things* ( $M = 1.77$ ,  $SD = 1.28$ ), item 8 *Experienced sleep disturbances* ( $M = 1.69$ ,  $SD = 1.35$ ), and item 1 *Felt nervous or anxious* ( $M = 1.61$ ,  $SD = 1.28$ ). The IAQ items with the highest mean scores for the Irish sample were item 2 ( $M = 1.68$ ,  $SD = 1.24$ ), item 1 ( $M = 1.47$ ,  $SD = 1.22$ ), and item 7 *Been easily annoyed by different things* ( $M = 1.45$ ,  $SD = 1.16$ ). The UK sample had higher mean scores on all IAQ items as compared to the Irish sample except for item 2 and item 6 *Had difficulty concentrating*.

The mean summed IDQ and IAQ scores were significantly higher for the UK sample (IDQ = 11.19,  $SD = 9.89$ ; IAQ = 12.32,  $SD = 8.99$ ) as compared to the Irish sample (IDQ = 9.19,  $SD = 8.61$ ; IAQ = 10.99,  $SD = 9.19$ ), although there was no difference in the levels of functional impairment.

#### 3.2 | CFA and reliability results

The CFA fit statistics in Table 2 show that the correlated two-factor model was acceptable in both samples<sup>1</sup> on all fit statistics except the chi-square. The chi-square was significant for all models however this should not lead to rejection of these models as the power of chi-square tests is positively related to sample size (Tanaka, 1987). The standardized factor loadings were all positive, high, and statistically significant ( $p < .001$ ), and these are reported in Table 3. The configural and metric models of invariance also indicated adequate model fit based on the differences in the CFI, RMSEA and SRMR ( $\Delta CFI = .002$ ,  $\Delta RMSEA = .001$ ,  $\Delta SRMR = .003$ ).

The composite reliability ( $\omega$ ) estimates for the IAQ and IDQ were high (UK IAQ  $\omega = .95$ , UK IDQ  $\omega = .96$ , Ireland IAQ  $\omega = .95$ , Ireland IDQ  $\omega = .95$ ).

The data from the United Kingdom and Ireland were combined and binary variables representing country and gender, and a continuous variable representing age, were added as predictors of the anxiety and depression latent variables. The standardized regression coefficients from the country variable to the latent variable indicated that there was significant difference (UK higher than Ireland) in the factor means for depression ( $\beta = .217$ ,  $p < .001$ ) and anxiety ( $\beta = .158$ ,  $p < .001$ ). There was a significant negative effect for age on depression ( $\beta = -.022$ ,  $p < .001$ ) and anxiety ( $\beta = -.024$ ,  $p < .001$ ). Females had significantly higher latent mean scores for anxiety ( $\beta = .187$ ,  $p < .001$ ) than males. Collectively, these variables explained 14.1% ( $p < .001$ ) of the variance in the depression latent variable and 18.1% ( $p < .001$ ) for the anxiety latent variable.

<sup>1</sup>A one-factor model fitted the data poorly of the Irish ( $\chi^2 = 1400.07$ ,  $df = 119$ ,  $p < .001$ ; RMSEA = .103 (90% CI .098–.108); CFI = .881; TLI = .864; SRMR = .044) and the UK ( $\chi^2 = 1334.28$ ,  $df = 119$ ,  $p < .001$ ; RMSEA = .101 (90% CI .096–.105); CFI = .873; TLI = .855; SRMR = .046) sample.

**TABLE 3** Fit statistics for the correlated two-factor model and tests of invariance for the International Anxiety Questionnaire (IAQ)/International Depression Questionnaire (IDQ) items.

Model	$\chi^2$	Df	<i>p</i>	CFI	TLI	RMSEA (90% CI)	SRMR
CFA							
UK	749.974	118	< .001	.941	.932	.073 (.068–.078)	.029
Ireland	842.560	118	< .001	.924	.913	.078 (.073–.083)	.035
Invariance							
Configural	1593.539	236	< .001	.933	.923	.075 (.072–.079)	.033
Metric	1657.362	251	< .001	.931	.925	.074 (.071–.078)	.036
Difference	40.175	15	< .001	.002		.001	.003

Note:  $\chi^2$  = chi-square goodness of fit statistic; *p* = statistical significance.  $\omega$  = omega reliability coefficient.

Abbreviations: CFI, comparative fit index; *df*, degrees of freedom; TLI, Tucker Lewis index; RMSEA (90% CI), root-mean-square error of approximation with 90% confidence intervals; SRMR, standardized root-mean square residual.

**TABLE 4** Standardized factor loadings for International Anxiety Questionnaire (IAQ)/International Depression Questionnaire (IDQ) confirmatory factor analysis for each country.

Item	Ireland	UK
Depression		
1. Felt down or depressed <i>for most of the day</i> ?	.849	.836
2. Experienced less interest or pleasure from normal activities <i>for most of the day</i> ?	.853	.848
3. Have had difficulty concentrating?	.831	.854
4. Had feelings of worthlessness or guilt?	.861	.858
5. Felt hopeless?	.863	.871
6. Had recurrent thoughts of death or suicide?	.704	.660
7. Have had changes in appetite or sleep?	.820	.815
8. Moved slower or felt more restless?	.865	.853
9. Experienced reduced energy or fatigue?	.829	.826
Anxiety		
1. Felt nervous or anxious?	.869	.847
2. Worried a lot about different things?	.862	.852
3. Felt physically tense or agitated?	.899	.885
4. Felt your heart racing, difficulty breathing, stomach discomfort, or dry mouth?	.838	.815
5. Felt “on edge”?	.863	.882
6. Had difficulty concentrating?	.855	.878
7. Been easily annoyed by different things?	.840	.824
8. Experienced sleep disturbances?	.759	.741
Factor correlation	.897	.876

The largest MI and SEPC was a direct effect between the variable representing gender and item 6 of the IDQ item (*Had recurrent thoughts of death or suicide?*: MI = 31.743, SEPC = -.251). This direct effect was added, and the model was re-estimated. No other MI/SPEC met the criteria for adding additional direct effects. The final model estimates show that the standardized direct effect from gender to IDQ item 6 was  $\beta = -.230$  ( $p < .001$ ) indicating that males scored higher on this item than females, when the level of depression was constant. The overall effect was small, increasing the percentage of

**TABLE 5** Bivariate correlations for the International Anxiety Questionnaire (IAQ) and International Depression Questionnaire (IDQ).

	PHQ-4	IPGDS
UK sample (N = 1012)		
IAQ summed scores	.82***	.60***
IDQ summed scores	.82***	.69***
Irish sample (N = 1011)		
IAQ summed scores	.83***	.54***
IDQ summed scores	.82***	.61***

\* $p < .05$ , \*\* $p < .01$ , and \*\*\* $p < .001$ .

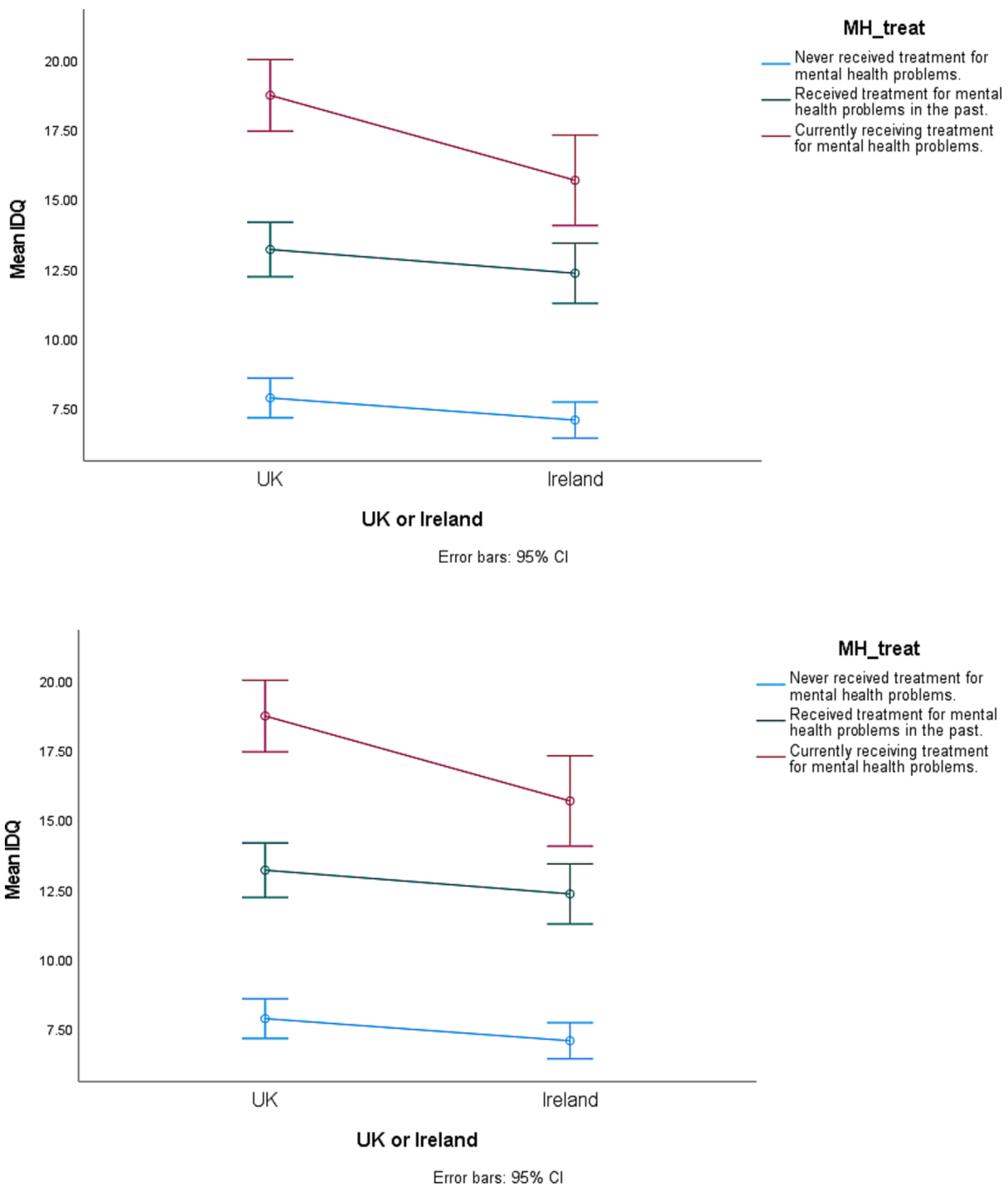
variance explained in that item from 48.0% to 49.3%. Thus, this direct effect explained 1.3% of the variation of that item (Table 4).

### 3.3 | Associations with external variables

As shown in Table 5, the IAQ and IDQ summed scores were strongly, positively, and significantly ( $ps < .001$ ) correlated with the summed PHQ-4 scores, and the IPGDS (prolonged grief disorder symptoms) scores in the UK and Irish samples.

A three-way factorial analysis of variance (ANOVA) was conducted to examine the influence of country and mental health treatment-seeking on mean depression and anxiety scores. For depression scores, there was a significant main effect for country ( $F [1, 2022] = 11.57, p < .001, \eta^2 = .00$ ), and mental health help seeking ( $F [2, 2022] = 173.92, p < .001, \eta^2 = .14$ ) and no significant interaction ( $F [2, 2022] = 285.069, p = .142$ ). For anxiety scores, there was no main effect for country ( $F [1, 2022] = 198.428, p = .07$ ), a significant main effect for mental health help seeking ( $F [2, 2022] = 26222.34, p < .001$ ), and the interaction was not significant ( $F [2, 2022] = 231.65, p = .152$ ). The mean plots are shown in Figure 1, and the mean IDQ and IAQ scores are highest for those who reported currently seeking mental health treatment, lowest for those who never sought treatment, and the intermediate group were those who reported having previously sought help.





**FIGURE 1** Mean plots for International Depression Questionnaire (IDQ) and International Anxiety Questionnaire (IAQ) scores by country and mental health treatment status.

### 3.4 | Prevalence estimates

In the UK sample, 18.6% (95% CI = 16.2%, 21.0%) met criteria for ICD-11 GAD, and 13.8% (95% CI = 11.7%, 16.0%) met criteria for ICD-11 DD. Of those who met requirements for ICD-11 GAD or ICD-11 DD, 47.7% (95% CI = 41.1%, 54.4%) met requirements for both disorders, 15.3% (95% CI = 10.5%, 20.1%) met requirements for

ICD-11 DD only, and 36.9% (95% CI = 30.5%, 43.3%) met requirements for ICD-11 GAD only. Significantly more females than males met criteria for ICD-11 GAD (23.7% vs. 12.8%,  $\chi^2 [1] = 19.88$ ,  $p < .001$ ; OR = 2.12 [95% CI = 1.52, 2.96]) but no gender differences were observed for ICD-11 DD (14.8% vs. 12.0%,  $\chi^2 [1] = 1.78$ ,  $p = .182$ ; OR = 1.28 [95% CI = .89, 1.85]). Those who met the diagnostic requirements for ICD-11 GAD were also significantly younger

than those that did not ( $M = 37.46$  [ $SD = 13.06$ ] years vs.  $M = 47.33$  [ $SD = 15.87$ ] years:  $t(325.81) = 8.95$ ,  $p < .001$ ,  $d = .64$ ). Likewise, those who met the diagnostic requirements for ICD-11 DD were significantly younger than those that did not ( $M = 38.09$  [ $SD = 12.78$ ] years vs.  $M = 46.68$  [ $SD = 15.99$ ] years:  $t(215.463) = 7.12$ ,  $p < .001$ ,  $d = .55$ ).

In the Irish sample, 16.1% (95% CI = 13.9%, 18.4%) of the sample met criteria for ICD-11 GAD, and 10.5% (95% CI = 8.6%, 12.4%) met criteria for ICD-11 DD. Of those who met requirements for ICD-11 GAD or ICD-11 DD, 49.4% (95% CI = 42.1%, 56.8%) met requirements for both disorders, 9.4% (95% CI = 5.1%, 13.8%) met requirements for ICD-11 DD only, and 41.1% (95% CI = 33.9%, 48.4%) met requirements for ICD-11 GAD only. Significantly more females than males met criteria for ICD-11 GAD (14.3% vs. 6.3%,  $\chi^2[1] = 17.28$ ,  $p < .001$ ; OR = 2.95 [95% CI = 1.60, 3.88]) and ICD-11 DD (22.6% vs. 9.0%,  $\chi^2[1] = 34.37$ ,  $p < .001$ ; OR = 2.49 [95% CI = 2.03, 4.29]). Those who met the diagnostic requirements for ICD-11 GAD were also significantly younger than those that did not ( $M = 37.07$  [ $SD = 11.97$ ] years vs.  $M = 46.63$  [ $SD = 15.64$ ] years:  $t(1009) = 7.39$ ,  $p < .001$ ,  $d = .63$ ). Likewise, those who met the diagnostic requirements for ICD-11 DD were significantly younger than those that did not ( $M = 37.95$  [ $SD = 12.71$ ] years vs.  $M = 45.92$  [ $SD = 15.59$ ] years:  $t(1009) = 5.07$ ,  $p < .001$ ,  $d = .52$ ).

## 4 | DISCUSSION

The current study's objectives were to examine the (1) combined dimensionality of the IAQ and the IDQ, (2) internal reliability of the IAQ and IDQ scores, (3) differential item functioning according to age, gender, and nationality, (4) convergent validity of the IAQ and IDQ scores, (5) ability of the IAQ and IDQ scores to distinguish between people who have and have not previously received treatment for mental health issues, and to (6) estimate prevalence rates of ICD-11 GAD and ICD-11 DD and determine whether these prevalence rates varied according to sex and age.

Extending findings from Shevlin, Butter, McBride, et al. (2022), the current study demonstrated that a correlated two-factor model, in which all IDQ items loaded on a "Depression" latent variable and all IAQ items loaded on an "Anxiety" latent variable provided the best fit to the data in both samples. All factor loadings were strong and statistically significant and both the IAQ and IDQ demonstrated excellent internal consistency. The UK sample generally reported higher IDQ/IAQ scores at the item and total scale level, and the MIMIC model also indicated that the latent means were higher. Recent research reported no difference in population levels of depression and anxiety between United Kingdom and Ireland (Shevlin, Butter, McBride, et al., 2022). The difference may be attributable to cultural differences in relation to bereavement and loss. In Ireland, it is customary for the community to come together to support the grieving individuals and share their burdens (Ronan, 2021). Specifically, it is common to have a wake in the family home, and this is typically followed by a funeral, a burial or cremation service, and then a repast.

Different mourning customs generally apply in the United Kingdom, where the deceased is kept at a mortuary until a funeral ceremony and a repast takes place. It could be argued that the traditions implemented in Ireland are somewhat more "community-focused," with it being well-established that social support is strongly linked to bereavement outcomes (e.g., Cao et al., 2020; Scott et al., 2020). Moreover, Ireland is more ethnically and religiously homogenous than the United Kingdom such that the Irish population is predominantly Christian, or influenced by Christian traditions related to death, whereas the United Kingdom is composed of more varied religious-based or religiously influenced traditions related to death. Future research may benefit from examining how cultural and ethnic differences in grieving and loss in the United Kingdom and Ireland relate to bereavement outcomes.

Regarding the DIF analysis, after controlling for the overall level of depression, males were found to score higher on the IDQ item *Had recurrent thoughts of death or suicide*; however, the size of the effects were small and unlikely to contribute to incorrect inferences surrounding gender differences in IDQ scores. Overall, our findings indicate that both the IDQ and IAQ are measuring depression and anxiety, respectively, in the same manner, across country, sex, and age among the bereaved population. These findings largely mirror a prior study examining the measurement invariance of the PHQ-9 and GAD-7 in the general populations of the United Kingdom and Ireland (Shevlin, Butter, McBride, et al., 2022). Supporting the convergent validity of the IAQ and IDQ, findings highlighted a significant association between the latent IAQ and IDQ scores and scores on the PHQ-4. These results reflect those of Shevlin, Butter, McBride, et al. (2022) who also found that summed scores on the IAQ and IDQ were strongly associated with scores on DSM-IV based measures of GAD and MDD, respectively. The strong association between the IAQ and IDQ scores and an independent measure of prolonged grief symptoms (i.e., IPGDS) aligns with research highlighting a high degree of co-occurrence among the symptoms of anxiety, depression, and prolonged grief disorder (Komicshke-Konnerup et al., 2021). Despite a high degree of co-occurrence between these disorders, previous factor analytic research has shown that these disorders represent similar yet distinct constructs (Boelen & van den Bout, 2005). That said, the ICD-11 stipulates that if a person satisfies the diagnostic requirements for both ICD-11 DD and ICD-11 PGD, then both diagnoses should be given.

Similar to Shevlin, Butter, McBride, et al. (2022), the current study found significant differences in IAQ and IDQ scores depending on mental health treatment-seeking status such that those who were currently receiving mental health treatment had the highest average IAQ and IDQ scores, while those who had never received mental health treatment had the lowest average IAQ and IDQ scores. There was no interaction effect between country and mental health treatment seeking status on average IAQ and IDQ scores, suggesting that the IAQ and IDQ effectively discriminates among different levels of anxiety and depression symptom severity irrespective of country of residence.

The final aim of the present study was to determine the prevalence of ICD-11 GAD and ICD-11 DD and whether prevalence rates

varied according to sex and age. Findings demonstrate that the prevalence of ICD-11 GAD was 18.6% and 16.1% in the UK and Irish samples, respectively, and the prevalence of ICD-11 DD was 13.8% and 10.5% in the UK and Irish samples, respectively. These rates are higher than those observed by Shevlin, Butter, McBride, et al. (2022) in their large community sample of adults from the United Kingdom where the prevalence of ICD-11 GAD and ICD-11 DD were 7.1% and 7.4%, respectively. It is well established that the loss of a loved one can trigger the onset or worsening of depression and anxiety (e.g., Kristiansen et al., 2019; Onrust & Cuijpers, 2006; Shear & Skritskaya, 2012; Zisook et al., 2014), and hence it is likely that the bereaved nature of the participants in the present study explains these higher rates. The co-occurrence of depression and anxiety disorders is well documented (e.g., Möller et al., 2016; Shevlin et al., 2022, b), and the results of the current study show that more people in the UK and Irish samples met the criteria for both ICD-11 GAD and ICD-11 DD rather than for either disorder alone. These findings support recommendations from Shevlin, Hyland, Nolan, et al. (2022) who suggested that the IAQ and IDQ are appropriate measures for the assessment of ICD-11 mixed depressive and anxiety disorder (MDAD; 6A73). The availability of such a diagnosis may reduce likelihood of overdiagnosis in the bereaved population, a particularly important point given the stigma that can be associated with mental disorders (Schomerus et al., 2012). Contrary to previous research (Shevlin, Butter, McBride, et al., 2022), which found that more people met the criteria for both ICD-11 DD and ICD-11 GAD than for either disorder alone, the current study demonstrated that the percentage of people who met the criteria for ICD-11 GAD alone was comparable to the percentage of people who met the criteria for both disorders. It would seem that among those who have experienced a loss, “pure” anxiety is equally as prevalent as combined anxiety and depression. Indeed, anxiety is a common response to bereavement due to the separation from a significant other, confrontation with one's mortality, and exposure to stressors such as financial adversity (Shear & Skritskaya, 2012).

Consistent with prior research (e.g., Boyd et al., 2015), findings highlighted a higher prevalence of ICD-11 GAD among females in both samples and a higher prevalence of ICD-11 DD among females in the Irish sample only. Different factors have been proposed to account for sex differences in internalizing disorders including genetic, neurobiological, neurodevelopmental, environmental, and psychological (for review, see Altemus et al., 2014). The absence of such an effect for ICD-11 DD in the UK sample is unsurprising given that prior research reported no sex differences in ICD-11 GAD nor ICD-11 DD in the UK general population (Shevlin, Hyland, Butter, et al., 2022). Consistent with existing research (e.g., Kessler et al., 2010; Hobbs et al., 2014), and our earlier findings regarding mean IAQ and IDQ scores, findings illustrated a higher prevalence of ICD-11 GAD and ICD-11 DD among younger adults as compared to older adults.

This study has several limitations. First, participants were recruited using a non-probability sampling method and hence the degree to which the samples are representative of the UK and Irish bereaved populations is uncertain. However, the composition of the final samples reflected the sex, age, and regional distributions of

the respective nations. Second, further research is required in clinical samples where the prevalence of ICD-11 GAD and ICD-11 DD are likely to be much higher. It should be noted that a clinician-administered measure of ICD-11 GAD and DD has not yet been developed. Third, despite these samples being composed of bereaved adults, it was not possible to ascertain whether ICD-11 GAD and ICD-11 DD stemmed from the bereavement specifically or from other factors. Finally, the samples used in the current study were drawn from affluent, English-speaking Western European nations. Given that one focus of the ICD-11 is on improving the clinical utility of psychiatric diagnoses globally (Reed et al., 2018), replication of this study's procedures across low- and middle-income countries is crucial.

In conclusion, this study provides further evidence that the IAQ and IDQ are psychometrically sound measures of ICD-11 GAD and DD. It should be noted that although the IDQ was measured with the intention of capturing ICD-11 single episode Depressive Disorder, that it may also be used at the clinician's discretion to screen for ICD-11 recurrent Depressive Disorder. With the exception of the requirement of prior episodes, recurrent Depressive Disorder (code 6A71) shares the same symptom profile as Single Episode Depressive Disorder. Our results show that the IAQ and IDQ generate reliable and valid scores irrespective of sex, age, and nationality. Our results also show that ICD-11 GAD and DD are prevalent in a substantial minority of bereaved people in the UK and Irish populations, underscoring the mental health effects of bereavement.

## ACKNOWLEDGEMENTS

Open access funding provided by IReL.

## CONFLICT OF INTEREST STATEMENT

The authors have no conflicts of interest to declare.

## DATA AVAILABILITY STATEMENT

Data for this study are not publicly available.

## ETHICS STATEMENT

Ethical approval for the collection of all data was provided by the research ethics committee at Ulster University (Reference number: FCPSY-22-026-A).

## ORCID

Thanos Karatzias  <https://orcid.org/0000-0002-3002-0630>

## REFERENCES

- Altemus, M., Sarvaiya, N., & Neill Epperson, C. (2014). Sex differences in anxiety and depression clinical perspectives. *Frontiers in Neuroendocrinology*, 35(3), 320–330. <https://doi.org/10.1016/j.yfrne.2014.05.004>
- Beck, A. T., Epstein, N., Brown, G., & Steer, R. (1988). Beck Anxiety Inventory [Databaserecord]. APA PsycTests. <https://doi.org/10.1037/t02025-000>
- Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin*, 107(2), 238–246. <https://doi.org/10.1037/0033-2909.107.2.238>
- Boelen, P. A., & van den Bout, J. (2005). Complicated grief, depression, and anxiety as distinct postloss syndromes: A confirmatory factor

- analysis study. *American Journal of Psychiatry*, 162(11), 2175–2177. <https://doi.org/10.1176/appi.ajp.162.11.2175>
- Boyd, A., van de Velde, S., Vilagut, G., de Graaf, R., O'Neill, S., Florescu, S., Alonso, J., & Kovess-Masfety, V. (2015). Gender differences in mental disorders and suicidality in Europe: Results from a large cross-sectional population-based study. *Journal of Affective Disorders*, 173, 245–254. <https://doi.org/10.1016/j.jad.2014.11.002>
- Cao, X., Yang, C., & Wang, D. (2020). The impact on mental health of losing an only child and the influence of social support and resilience. *OMEGA-Journal of Death and Dying*, 80(4), 666–684. <https://doi.org/10.1177/0030222818755284>
- Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling: a Multidisciplinary Journal*, 14(3), 464–504. <https://doi.org/10.1080/10705510701301834>
- Chou, C.-P., & Bentler, P. M. (1993). Invariant standardized estimated parameter change for model modification in covariance structure analysis. *Multivariate Behavioral Research*, 28(1), 97–110. [https://doi.org/10.1207/s15327906mbr2801\\_6](https://doi.org/10.1207/s15327906mbr2801_6)
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum Associates, Publishers.
- Hobbs, M. J., Anderson, T. M., Slade, T., & Andrews, G. (2014). Relationship between measurement invariance and age-related differences in the prevalence of generalized anxiety disorder. *Journal of Affective Disorders*, 152, 306–312. <https://doi.org/10.1016/j.jad.2013.09.030>
- Jöreskog, K. G., & Sörbom, D. (1996). *LISREL 8: User's reference guide*. Scientific Software.
- Kaplan, D. (1989). Model modification in covariance structure analysis: Application of the expected parameter change statistic. *Multivariate Behavioral Research*, 24(3), 285–305. [https://doi.org/10.1207/s15327906mbr2403\\_2](https://doi.org/10.1207/s15327906mbr2403_2)
- Kessler, R. C., McLaughlin, K. A., Green, J. G., Gruber, M. J., Sampson, N. A., Zaslavsky, A. M., Aguilar-Gaxiola, S., Alhamzawi, A. O., Alonso, J., Angermeyer, M., Benjet, C., Bromet, E., Chatterji, S., de Girolamo, G., Demyttenaere, K., Fayyad, J., Florescu, S., Gal, G., Gureje, O., ... Williams, D. R. (2010). Childhood adversities and adult psychopathology in the WHO World Mental Health Surveys. *The British Journal of Psychiatry*, 197(5), 378–385. <https://doi.org/10.1192/bjp.bp.110.080499>
- Killikelly, C., Merzhvynska, M., Zhou, N., Stelzer, E. M., Hyland, P., Rocha, J., Ben-Ezra, M., & Maercker, A. (2021). Examination of the new ICD-11 prolonged grief disorder guidelines across five international samples. *Clinical Psychology in Europe*, 3(1), e4159. <https://doi.org/10.32872/cpe.4159>
- Killikelly, C., Zhou, N., Merzhvynska, M., Stelzer, E. M., Dotschung, T., Rohner, S., Sun, L. H., & Maercker, A. (2020). Development of the international prolonged grief disorder scale for the ICD-11: Measurement of core symptoms and culture items adapted for chinese and german-speaking samples. *Journal of Affective Disorders*, 277, 568–576. <https://doi.org/10.1016/j.jad.2020.08.057>
- Komischke-Konnerup, K. B., Zachariae, R., Johannsen, M., Nielsen, L. D., & O'Connor, M. (2021). Co-occurrence of prolonged grief symptoms and symptoms of depression, anxiety, and posttraumatic stress in bereaved adults: A systematic review and meta-analysis. *Journal of Affective Disorders Reports*, 4, 100140. <https://doi.org/10.1016/j.jadr.2021.100140>
- Kristiansen, C. B., Kjær, J. N., Hjorth, P., Andersen, K., & Prina, A. M. (2019). Prevalence of common mental disorders in widowhood: A systematic review and meta-analysis. *Journal of Affective Disorders*, 245, 1016–1023. <https://doi.org/10.1016/j.jad.2018.11.088>
- Kroenke, K., Spitzer, R. L., & Williams, J. B. (2001). The PHQ-9: Validity of a brief depression severity measure. *Journal of General Internal Medicine*, 16(9), 606–613. <https://doi.org/10.1046/j.1525-1497.2001.016009606.x>
- Kroenke, K., Spitzer, R. L., Williams, J. B., & Löwe, B. (2009). An ultra-brief screening scale for anxiety and depression: The PHQ-4. *Psychosomatics*, 50(6), 613–621. <https://doi.org/10.1176/appi.psy.50.6.613>
- Löwe, B., Wahl, I., Rose, M., Spitzer, C., Glaesmer, H., Wingenfeld, K., Schneider, A., & Brähler, E. (2010). A 4-item measure of depression and anxiety: Validation and standardization of the Patient Health Questionnaire-4 (PHQ-4) in the general population. *Journal of Affective Disorders*, 122(1–2), 86–95. <https://doi.org/10.1016/j.jad.2009.06.019>
- Möller, H. J., Bandelow, B., Volz, H. P., Barnikol, U. B., Seifritz, E., & Kasper, S. (2016). The relevance of 'mixed anxiety and depression' as a diagnostic category in clinical practice. *European Archives of Psychiatry and Clinical Neuroscience*, 266(8), 725–736. <https://doi.org/10.1007/s00406-016-0684-7>
- Muthén, L. K., & Muthén, B. (2017). *Mplus user's guide: Statistical analysis with latent variables, user's guide*. Muthén & Muthén.
- Onrust, S. A., & Cuijpers, P. (2006). Mood and anxiety disorders in widowhood: A systematic review. *Aging & Mental Health*, 10(4), 327–334. <https://doi.org/10.1080/13607860600638529>
- Pham, S., Porta, G., Biernesser, C., Walker Payne, M., Iyengar, S., Melhem, N., & Brent, D. A. (2018). The burden of bereavement: Early-onset depression and impairment in youths bereaved by sudden parental death in a 7-year prospective study. *The American Journal of Psychiatry*, 175(9), 887–896. <https://doi.org/10.1176/appi.ajp.2018.17070792>
- Radloff, L. S. (1977). The CES-D scale: A self-report depression scale for research in the general population. *Applied Psychological Measurement*, 1, 385–401. <https://doi.org/10.1177/014662167700100306>
- Reed, G. M., Keeley, J. W., Rebello, T. J., First, M. B., Gureje, O., Ayuso-Mateos, J. L., Kanba, S., Khoury, B., Kogan, C. S., Krasnov, V. N., Maj, M., de Jesus Mari, J., Sharan, P., Stein, D. J., Zhao, M., Akiyama, T., Andrews, H. F., Asevedo, E., Cheour, M., ... Medina-Mora, M. E. (2018). Clinical utility of ICD-11 diagnostic guidelines for high-burden mental disorders: Results from mental health settings in 13 countries. *World Psychiatry*, 17(3), 306–315. <https://doi.org/10.1002/wps.20581>
- Ronan, M. (2021). Funerals in the time of coronavirus. *Irish Journal of Sociology*, 29(2), 236–240. <https://doi.org/10.1177/0791603520967617>
- Saris, W. E., Satorra, A., & Sorbom, D. (1987). The detection and correction of specification errors in structural equation models. *Sociological Methodology*, 17, 105. <https://doi.org/10.2307/271030>
- Scholten, S., Veltens, J., Bieda, A., Zhang, X. C., & Margraf, J. (2017). Testing measurement invariance of the Depression, Anxiety, and Stress Scales (DASS-21) across four countries. *Psychological Assessment*, 29(11), 1376–1390. <https://doi.org/10.1037/pas0000440>
- Schomerus, G., Schwahn, C., Holzinger, A., Corrigan, P. W., Grabe, H. J., Carta, M. G., & Angermeyer, M. C. (2012). Evolution of public attitudes about mental illness: A systematic review and meta-analysis. *Acta Psychiatrica Scandinavica*, 125(6), 440–452. <https://doi.org/10.1111/j.1600-0447.2012.01826.x>
- Scott, H. R., Pitman, A., Kozhuharova, P., & Lloyd-Evans, B. (2020). A systematic review of studies describing the influence of informal social support on psychological wellbeing in people bereaved by sudden or violent causes of death. *BMC Psychiatry*, 20, 1–20. <https://doi.org/10.1186/s12888-020-02639-4>
- Shear, M. K., & Skritskaya, N. A. (2012). Bereavement and anxiety. *Current Psychiatry Reports*, 14(3), 169–175. <https://doi.org/10.1007/s11920-012-0270-2>
- Shevlin, M., Butter, S., McBride, O., Murphy, J., Gibson-Miller, J., Hartman, T. K., Levita, L., Mason, L., Martinez, A. P., McKay, R., Stocks, T. V. A., Bennett, K. M., Hyland, P., Vallieres, F., Valiente, C., Vazquez, C., Contreras, A., Peinado, V., Trucharte, A., ... Bentall, R. P. (2022). Measurement invariance of the Patient Health Questionnaire (PHQ-9) and Generalized Anxiety Disorder Scale (GAD-7) across four European countries during the COVID-19 pandemic. *BMC Psychiatry*, 22(1), 154. <https://doi.org/10.1186/s12888-022-03787-5>
- Shevlin, M., Hyland, P., Butter, S., McBride, O., Hartman, T. K., Karatzias, T., & Bentall, R. P. (2022). The development and initial

- validation of self-report measures of ICD-11 depressive episode and generalized anxiety disorder: The International Depression Questionnaire (IDQ) and the International Anxiety Questionnaire (IAQ). *Journal of Clinical Psychology*, 79, 854–870. <https://doi.org/10.1002/jclp.23446>
- Shevlin, M., Hyland, P., Nolan, E., Owczarek, M., Ben-Ezra, M., & Karatzias, T. (2022). ICD-11 'mixed depressive and anxiety disorder' is clinical rather than sub-clinical and more common than anxiety and depression in the general population. *British Journal of Clinical Psychology*, 61(1), 18–36. <https://doi.org/10.1111/bjc.12321>
- Shevlin, M., Redican, E., Hyland, P., Murphy, J., Karatzias, T., McBride, O., Bennett, K., Butter, S., Hartman, T. K., Vallières, F., & Bentall, R. P. (2023). Symptoms and levels of ICD-11 Prolonged Grief Disorder in a representative community sample of UK adults. *Social Psychiatry and Psychiatric Epidemiology*, 58, 1535–1547. <https://doi.org/10.1007/s00127-023-02469-1>
- Sörbom, D. (1989). Model modification. *Psychometrika*, 54(3), 371–384. <https://doi.org/10.1007/bf02294623>
- Spitzer, R. L., Kroenke, K., Williams, J. B., & Löwe, B. (2006). A brief measure for assessing generalized anxiety disorder: The GAD-7. *Archives of Internal Medicine*, 166(10), 1092–1097. <https://doi.org/10.1001/archinte.166.10.1092>
- Steiger, J. H. (1990). Structural model evaluation and modification: An interval estimation approach. *Multivariate Behavioral Research*, 25(2), 173–180. [https://doi.org/10.1207/s15327906mbr2502\\_4](https://doi.org/10.1207/s15327906mbr2502_4)
- Tanaka, J. S. (1987). "How big is big enough?": Sample size and goodness of fit in structural equation models with latent variables. *Child Development*, 58(1), 134–146. <https://doi.org/10.2307/1130296>
- Thimm, J. C., Kristoffersen, A. E., & Ringberg, U. (2020). The prevalence of severe grief reactions after bereavement and their associations with mental health, physical health, and health service utilization: A population-based study. *European Journal of Psychotraumatology*, 11(1), 1844440. <https://doi.org/10.1080/20008198.2020.1844440>
- Tucker, L. R., & Lewis, C. (1973). A reliability coefficient for maximum likelihood factor analysis. *Psychometrika*, 38(1), 1–10. <https://doi.org/10.1007/BF02291170>
- World Health Organisation. (2022). ICD-11 for mortality and morbidity statistics, (version: 02/2022). <https://icd.who.int/browse11/l-m/en>
- Yuan, K.-H., & Bentler, P. M. (2000). 5. Three likelihood-based methods for mean and covariance structure analysis with nonnormal missing data. *Sociological Methodology*, 30(1), 165–200. <https://doi.org/10.1111/0081-1750.00078>
- Zisook, S., Iglewicz, A., Avanzino, J., Maglione, J., Glorioso, D., Zetumer, S., Seay, K., Vahia, I., Young, I., Lebowitz, B., Pies, R., Reynolds, C., Simon, N., & Shear, M. K. (2014). Bereavement: Course, consequences, and care. *Current Psychiatry Reports*, 16(10), 1–10. <https://doi.org/10.1007/s11920-014-0482-8>
- Zisook, S., & Kendler, K. S. (2007). Is bereavement-related depression different than non bereavement-related depression? *Psychological Medicine*, 37(6), 779–794. <https://doi.org/10.1017/S0033291707009865>

**How to cite this article:** Hyland, P., Redican, E., Karatzias, T., & Shevlin, M. (2023). Assessing the validity and reliability of the International Anxiety Questionnaire and the International Depression Questionnaire in two bereaved national samples. *Clinical Psychology & Psychotherapy*, 1–12. <https://doi.org/10.1002/cpp.2917>