

## RESEARCH ON TRANSLATIONS OF TESTS

# Development and Validation of the Inventory of Depression and Anxiety Symptoms–II—German Version

Robin Anno Wester<sup>1</sup>, Julian Rubel<sup>1</sup>, Johannes Zimmermann<sup>2</sup>, Mila Hall<sup>1</sup>,  
Leonie Kaven<sup>3</sup>, and David Watson<sup>4</sup>

<sup>1</sup> Department of Psychology, University of Giessen

<sup>2</sup> Department of Psychology, University of Kassel

<sup>3</sup> Department of Psychology, University of Greifswald

<sup>4</sup> Department of Psychology, University of Notre Dame

The expanded version of the Inventory of Depression and Anxiety Symptoms (IDAS-II) is a 99-item self-report measure containing 18 nonoverlapping dimensional scales assessing symptoms of depression, anxiety, and mania. The aim of this study was to develop and validate a German adaptation of the IDAS-II. Participants from a community sample ( $N = 1,054$ ) completed the IDAS-II (German version). In addition, part of the sample ( $N = 550$ ) completed a series of additional measures of depression (Patient Health Questionnaire–9, WHO–Five Well-Being Index, Symptom Checklist–90 Revised—Short Version) and anxiety disorders (Generalized Anxiety Disorder Scale–7, Fear of Negative Evaluation Scale–5, Dimensional Obsessive–Compulsive Scale, The International Trauma Questionnaire). We conducted item-level confirmatory factor analyses (CFA) separately for the 15 nonsaturated IDAS-II scales, which confirmed unidimensionality. McDonald's  $\omega$  indicated good internal consistency ( $\omega > .8$ ) of all symptom scales except Euphoria ( $\omega = .751$ ) and Ordering ( $\omega = .728$ ). A CFA of the IDAS-II scales, based on exploratory results by Watson et al. (2012), confirmed a three-factor model of “Distress,” “Obsessions/Fear,” and “Positive Mood.” Correlational analyses with additional symptom measures supported the convergent and discriminant validity of the IDAS-II scales. The IDAS-II (German version) allows for a reliable assessment of the severity of depression, anxiety, and bipolar symptoms and is one of the first clinical measures for German-speaking samples that is consistent with the Hierarchical Taxonomy of Psychopathology (HiTOP).

### Public Significance Statement

This study describes the validation of the German translation of the Inventory of Depression and Anxiety Symptoms–II (IDAS-II). The IDAS-II is designed to assess symptoms of anxiety and depression (e.g., insomnia, panic). It is freely available at <https://www.testarchiv.eu/de/test/9008301> and can be used by practitioners and scientists. Given that inaccessible psychometric tools hinder reproducibility, the translation of this questionnaire represents an important step toward improving open science practices in German psychological research.

**Keywords:** scale construction, depression, anxiety, obsessive–compulsive disorder, HiTOP

**Supplemental materials:** <https://doi.org/10.1037/pas0001185.supp>

Robin Anno Wester  <https://orcid.org/0000-0002-5288-2098>

Julian Rubel  <https://orcid.org/0000-0002-9625-6611>

Johannes Zimmermann  <https://orcid.org/0000-0001-6975-2356>

Mila Hall  <https://orcid.org/0000-0002-0421-4781>

David Watson  <https://orcid.org/0000-0002-9605-0576>

This study was not preregistered. Data and analysis scripts are available at <https://osf.io/x3rqm/>. The IDAS-II (German version) is freely available at <https://www.testarchiv.eu/de/test/9008301>. David Watson is the copyright holder of the IDAS-II. No conflict of interest exists for any of the authors.

Robin Anno Wester played lead role in methodology and writing of

original draft. Julian Rubel played supporting role in writing of review and editing and equal role in conceptualization. Johannes Zimmermann played supporting role in writing of review and editing and equal role in methodology. Mila Hall played equal role in translator. Leonie Kaven played supporting role in visualization and equal role in translator. David Watson played supporting role in writing of review and editing and equal role in translation.

Correspondence concerning this article should be addressed to Robin Anno Wester, Department of Psychology, University of Giessen, Otto-Behaghel-Straße 10, Gießen D-35394, Germany. Email: [robin.wester-baumgartner@psychol.uni-giessen.de](mailto:robin.wester-baumgartner@psychol.uni-giessen.de)

Symptoms of mood and anxiety disorders are not only highly prevalent but they also often co-occur, thereby increasing the burden for those affected (Brown et al., 2001; McDermut et al., 2001). Traditionally, diagnostic manuals like the *Diagnostic and Statistical Manual of Mental Disorders (DSM)* or the *International Classification of Diseases (ICD)* create a categorical system in which these problems are assigned to single mental disorders (or groups of disorders), which have contributed to the development of separate measurement instruments for single syndromes. However, several shortcomings of categorical nosologic systems (such as a deficiency to account for comorbidity) have led to different initiatives to replace these systems. One example is the Hierarchical Taxonomy of Psychopathology (HiTOP; Kotov et al., 2017), which assumes a hierarchical organization of psychopathological dimensions. Models like HiTOP not only aim to provide a novel approach to nosology but also call for new measurement instruments that adhere to the models' assumptions. The expanded version of the Inventory of Depression and Anxiety Symptoms (IDAS-II; Watson et al., 2012) is a self-report instrument for the internalizing spectrum within the HiTOP framework. The aim of the current article is to present the German version of IDAS-II and analyze its psychometric properties.

### Dimensional Models of Psychopathology

HiTOP builds a nosologic system from (empirical) bottom to top. At the lowest level, symptom components and maladaptive traits are constructed based on the empirical covariation of signs, which can be observed objectively, and symptoms, which are experienced subjectively. According to their co-occurrence (i.e., comorbidity), those components are combined into empirical syndromes and subfactors, which are grouped in spectra. This empirically derived hierarchical structure allows HiTOP to overcome two main shortcomings of diagnostic manuals following a rational approach, like the *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5, American Psychiatric Association, 2013)*: First, the *DSM-5* leads to high heterogeneity within diagnostic categories, both theoretically (Galatzer-Levy & Bryant, 2013) and in actual diagnoses. For example, in their analysis of data from depressed patients in the STAR\*D trial, Fried and Nesse (2015) identified 1,030 unique symptom profiles, of which almost half were reported only by a single patient. Thus, two patients sharing the same diagnosis (e.g., depression) might have little in common clinically. Second, the *DSM-5* treats mental disorders as independent entities whose high rates of comorbidity indicate that they co-occur beyond chance. While this problem has been discussed for decades (Mineka et al., 1998), the *DSM-5* still treats depression and anxiety as distinct clinical entities—two groups of disorders whose comorbidity rates reach up to 70% (Kaufman & Charney, 2000; Lamers et al., 2011).

HiTOP's hierarchical structure, in which different levels are characterized by varying degrees of specificity, takes into account both shared and specific components of anxiety and depression. This enables researchers and practitioners to use different levels of analysis, in accordance with their practical or theoretical utility. While for some treatment decisions or research questions, the presence or absence of specific symptoms might be of utmost importance (e.g., insomnia), other decisions/questions aim at higher order levels (e.g., the emotional dysfunction common to disorders from the internalizing spectrum; Conway et al., 2022). Thus, HiTOP enables researchers and clinicians to account for both heterogeneity

within, as well as similarities across, classical diagnostic categories when describing individual patients.

In addition to this hierarchical structure, HiTOP is distinct from traditional nosologic approaches by describing psychopathology dimensionally. While the *DSM-5* holds a categorical view and posits that a person either does or does not have a mental disorder, HiTOP follows recent evidence that differences in psychopathology are latently continuous (Haslam et al., 2020) by assuming a continuum ranging from normal psychological functioning to psychopathology. More specifically, several studies support this assumption of a continuous latent structure for unipolar depression (Hankin et al., 2005; Ruscio & Ruscio, 2000), mania (Prisciandaro & Roberts, 2011; Prisciandaro & Tolliver, 2015), and various anxiety disorders (Broman-Fulks et al., 2010; Kertz et al., 2014; Kollman et al., 2006). By taking this dimensional perspective, HiTOP concurs with another framework for psychopathology, the National Institute of Mental Health's Research Domain Criteria (RDoC; Clark et al., 2017). In concert with HiTOP's account for comorbidity, this might solve issues of low interrater reliability associated with the *DSM-5* (Chmielewski et al., 2015; Regier et al., 2013), since treating a continuous phenomenon categorically leads to a loss of information (Markon et al., 2011).

### IDAS-II

To date, the IDAS-II is one of only two self-report instruments consistent with the HiTOP structure of internalizing disorders. The construction of the IDAS-II differs from that of other widely used measures of mood and anxiety symptoms, such as the Beck Depression Inventory–II (BDI-II) or the generalized anxiety disorder module (GAD-7) from the patient health questionnaire (PHQ). The aforementioned disorder-specific questionnaires were constructed with the aim of yielding general scores of depression/GAD by combining items measuring correlated, but distinct symptoms (e.g., insomnia and appetite loss), which contributes to their multidimensional structure (Fried et al., 2022; Stochl et al., 2022). In contrast, in the development of the IDAS, special attention has been given to differentiating between specific symptom dimensions of depression (e.g., suicidality, lassitude) and anxiety (e.g., panic, social anxiety, claustrophobia), meaning the lowest level within the HiTOP hierarchy. Therefore, Watson et al. (2007, 2012) conducted extensive structural analyses on a large pool of symptoms to construct homogeneous symptom components (i.e., unidimensional, symptom-specific scales), which not only led to the construction of the IDAS-II but also provided insight into the close association between DSM criteria assumed to be distinct symptom dimensions (see below).

As a result, the IDAS-II has several advantages: First, the IDAS is better able to differentiate between symptoms of depression and anxiety than many disorder-specific self-report measures, which are supposed to measure either depression or anxiety. Second, “global score measures,” like the BDI-II, do not allow for scale scores of single symptoms, because they are measured by a limited set of indicators (i.e., many symptoms are assessed with only one item). Since global scores “hide” an abundance of possible symptom profiles, they lead to a reduction in information (Fried & Nesse, 2015). While researchers have mostly ignored data on specific symptoms, clinicians usually do not base their decisions solely on overall symptom severity, but likely adjust their treatment

decisions to the specific problems patients present. Third, the IDAS-II permits the assessment of all nine diagnostic symptom criteria for major depression as outlined in the *DSM-5*. In contrast, the Center for Epidemiologic Studies Depression Scale (CES-D, Radloff, 1977) omits indicators for appetite gain or hypersomnia, which limits its applicability.

The IDAS was first developed as a 64-item questionnaire comprising 10 specific symptom scales (Well-Being, Suicidality, Lassitude, Insomnia, Appetite Loss, Appetite Gain, Ill Temper, Panic, Social Anxiety, Traumatic Intrusions) as well as a nonspecific scale that assesses core emotional and cognitive symptoms of internalizing disorders (Dysphoria; Watson et al., 2007). More specifically, Dysphoria contains items tapping into several *DSM-5* criteria of depression that are highly correlated (Watson et al., 2007): depressed mood (Criterion 1), anhedonia (Criterion 2), worthlessness and guilt (Criterion 7), psychomotor problems (Criterion 5), and cognitive problems (Criterion 8). To permit comparability with traditional measures of depression like the BDI-II, the authors also created a 20-item General Depression scale, which contains all items of the Dysphoria scale as well as items from Suicidality, Lassitude, Insomnia, Appetite Loss, and (low) Well-Being. The IDAS-II contains seven additional scales that measure further symptom dimensions of anxiety disorders along with bipolar disorder (Traumatic Avoidance, Checking, Ordering, Cleaning, Claustrophobia, Mania, Euphoria; Watson et al., 2012). Concerning the internal structure of the IDAS-II, an exploratory factor analysis (EFA) by Watson et al. (2012) yielded three interpretable factors that vary in their size and scope: Distress, Obsessions/Fear, and Positive Mood. Distress is a broad factor, which is mainly defined by Dysphoria and covers all symptoms of depression, mania, some anxiety disorders (panic and posttraumatic stress disorder [PTSD]), and secondary loadings of social anxiety and (low) Well-Being. The Obsessions/Fear factor contains symptoms of obsessive-compulsive behavior (OCD) as well as claustrophobia and a secondary loading of social anxiety. Positive Mood is a rather narrow factor and essentially defined by Well-Being and Euphoria, but with a secondary loading of Mania.

Several studies support the psychometric qualities of the IDAS-II and its clinical utility. Concerning reliability, IDAS-II scales exhibited (a) high internal consistencies (with the coefficient  $\alpha$  falling between .80 and .90 for all scales, except Euphoria, which has values falling between .72 and .79; Watson et al., 2012) and (b) 1-week retest correlations of .72 (Ill Temper) to .84 (General Depression) in a sample of psychiatric patients (Watson et al., 2007). Two studies have demonstrated good convergent and discriminant validity of the IDAS-II using different self-report and interview measures (Watson et al., 2008, 2012). Furthermore, Stasik-O'Brien et al. (2019) showed that the IDAS-II scales are good predictors of their associated *DSM-IV* as well as *DSM-5* diagnoses, and the authors provided cutoff scores for different uses.

Recently, Turkish (Irak & Albayrak, 2020) and Spanish (De la Rosa-Cáceres et al., 2020) versions of the IDAS-II have been developed. Regarding internal consistency, Cronbach's  $\alpha$  for several scales of the Turkish version fell below .8 (Appetite Loss, Appetite Gain, Mania, Euphoria, Checking, Ordering, Lassitude, Insomnia), with Cronbach's  $\alpha$  for Suicidality falling below .5. For the Spanish translation, Lassitude, Appetite Gain, Mania, Checking, and Ordering exhibited a Cronbach's  $\alpha$  below .8, but above .7 (De la Rosa-Cáceres et al., 2020). De la Rosa-Cáceres et al. (2020) also conducted a confirmatory factor analyses (CFA) based on the three-factor structure described above, but without any cross-loadings. This model fit well

in both a nonclinical and clinical sample. We are unaware of any study that assessed the dimensionality of scales in the Turkish and Spanish versions.

The objective of the present study was to adapt the IDAS-II to German-speaking populations. To this end, we developed a German translation of the IDAS-II and examined its psychometric properties in a community sample. We conducted a series of single-factor CFAs to investigate the unidimensionality of the IDAS-II scales and to estimate their reliability through internal consistencies. By conducting a multiple-factor CFA of scale scores, we examined whether the German version conformed to the internal structure of the IDAS-II proposed by Watson et al. (2012). Convergent and discriminant validity regarding established self-report measures of corresponding constructs was tested using correlational analyses. We did not expect meaningful differences between German-speaking and English-speaking populations in the way they conceptualize anxiety and depression. Therefore, we expected the German translation of the IDAS-II to have similar psychometric properties as the English version.

## Method

### Sample

Participants were recruited online via the panel provider Respondi. Respondi is based on a pool of participants with verified identities and complies with the international norm ISO 2052:2019, the quality standard for market, opinion, and social research. In successive waves, participants were randomly invited to participate in one of two simultaneously running surveys, until our prespecified sample size was reached. Survey 1 focused on research questions unrelated to this study but included the IDAS-II nonetheless in order to increase power for factor analyses. In contrast, Survey 2 focused exclusively on the validation of the IDAS-II. Since both surveys ran simultaneously, participants could only take part in one of them. To assess the reliability and internal structure of the IDAS-II, data from both surveys were combined. As mentioned above, the data from survey one were unrelated to IDAS-II content; hence, only data obtained in survey two were used to assess the relationship between the IDAS-II scales and other questionnaires. Participants were compensated by Respondi with reward points, which can be exchanged for money (hourly minimum wage in Germany at the time of the surveys was 9.50€), vouchers, or donations. Since individuals had to answer every item to proceed, there were no missing data.

### Sample 1

Overall, 741 participants started the survey, of which 237 were excluded because of indications of careless responding. Careless responding was assessed by including an attention check (i.e., two instructed response items explicitly asking for a specific answer) and a speeder test (i.e., a minimum amount of time was required for answering one of the questionnaires). Thus, the final sample consisted of 504 participants. Participants were approximately representative of the German general population in terms of age and gender: 51% of participants identified as female, 49% as male, and none as nonbinary. Age ranged from 18 to 74 ( $M = 46.6$ , interquartile range [IQR] = 25). The educational level varied widely: 0.2% did not attain a high-school diploma, 8.3% attained a high-school diploma after 9 years of school (German "Hauptschulabschluss"), 37.9% attained a high-school diploma after 10 years of

school (German “Realschulabschluss”), 21.8% attained a high-school diploma after 12 years of school (general qualification for university entrance, German “Abitur”), 11.3% attained a bachelor’s degree, 18.7% attained a master’s degree, and 1.2% attained a PhD. Approximately one-third of participants (33.1%) stated that they received a psychotherapeutic or psychiatric treatment in the past, while 13.1% of participants reported that they currently received a psychotherapeutic or psychiatric treatment.

## Sample 2

Overall, 774 participants started the survey, of which 224 were excluded because of indications of careless responding. Careless responding was assessed by including an attention check (i.e., one instructed response item explicitly asking for a specific answer) and a speeder test (i.e., a minimum amount of time was required for answering one of the questionnaires). Thus, the final sample consisted of 550 participants. Participants were approximately representative of the German general population in terms of age and gender: 51.1% of participants identified as female, 48.2% as male, 0.5% as nonbinary, and 0.2% did not indicate their gender. Age ranged from 18 to 74 ( $M = 46.2$ ,  $IQR = 26$ ). The educational level varied widely: 0.2% did not attain a high-school diploma, 11.5% attained a high-school diploma after 9 years of school (German “Hauptschulabschluss”), 36.2% attained a high-school diploma after 10 years of school (German “Realschulabschluss”), 20.9% attained a high-school diploma after 12 years of school (general qualification for university entrance, German “Abitur”), 10.2% attained a bachelor’s degree, 18.0% attained a master’s degree, and 2.9% attained a PhD. Just under one-third of participants (31.5%) stated that they received a psychotherapeutic or psychiatric treatment in the past, while 11.1% of participants reported that they currently received a psychotherapeutic or psychiatric treatment.

Concerning factor analyses reported in this article, the sample size exceeds the traditional (but often criticized) rule of thumb of 10 subjects per variable (Everitt, 1975) and the minimum necessary sample sizes in most conditions in the simulation study by Mundfrom et al. (2005). In addition, the size of the sample used in correlational analyses is twice that recommended for stable estimates by Schönbrodt and Perugini (2013).

## Instruments

### IDAS-II—German Version

IDAS-II instructions and items were translated to German separately by three authors (RAW, JZ, LK), who then compared translations. For 16 items, all translators came up with the same translation. For 19 items, only two translators agreed on a translation. For the remaining 64 items, translations by all three translators differed. In cases of divergence, translators discussed their different versions and compared dictionary entries of hard-to-translate words and phrases. For all items, translators were able to agree on a consensus version. This German version was back-translated to English by an English native speaker (MH) and sent to the author of the original version (DW) for feedback. Using this feedback, some items were reformulated and again back-translated to English. During this process, it became apparent that two-item formulations of the Ill Temper scale were redundant in German, which is why one item was

removed, reducing the length of this scale to four items and the overall length of IDAS-II to 98 items. Participants are asked to determine how well each item (e.g., “I had little interest in my usual hobbies and activities”) describes their feelings and experiences over the last 2 weeks on a 5-point scale ranging from 1 (*not at all*) to 5 (*extremely*).

## Measures in Sample 2

### Patient Health Questionnaire–9 (Kroenke et al., 2001)

The Patient Health Questionnaire–9 (PHQ-9) is the nine-item depression module of the Patient Health Questionnaire (PHQ) for common mental disorders (Spitzer, 1999). Items (e.g., “Little interest or pleasure in doing things?”) are rated on a 4-point scale ranging from 0 = *not at all* to 3 = *nearly every day*. The German version of the PHQ-9 has been demonstrated to have good sensitivity and specificity as a screener for depression in a large psychiatric sample (Gräfe et al., 2004). McDonald’s  $\omega$  of the PHQ-9 in this sample was .91.

### Generalized Anxiety Disorder Scale–7 (Spitzer et al., 2006)

The Generalized Anxiety Disorder Scale–7 (GAD-7) is a seven-item measure for generalized anxiety disorder (GAD) and part of the PHQ. Items (e.g., “Not being able to stop or control worrying”) are answered on a 4-point scale ranging from 0 = *not at all* to 3 = *nearly every day*. Factor analyses in a large community sample confirmed the unidimensionality and high internal consistency of the German version of the GAD-7 (Löwe et al., 2008). McDonald’s  $\omega$  of the GAD-7 in this sample was .92.

### WHO–Five Well-Being Index (WHO, 1998)

The WHO–Five Well-Being Index (WHO-5) is a self-report measure of positive affect, consisting of five items (e.g., “Over the last 2 weeks I have felt cheerful and in good spirits”). Items are rated on a 6-point Likert scale that ranges from 0 (*at no time*) to 5 (*all of the time*). The German translation of the WHO-5 has been demonstrated to exhibit good psychometric properties in a large representative sample (Brähler et al., 2007). McDonald’s  $\omega$  of the WHO-5 in this sample was .94.

### Fear of Negative Evaluation Scale–5 (Kemper et al., 2011)

The Fear of Negative Evaluation Scale–5 (SANB-5) is a five-item German short version of the Fear of Negative Evaluation Scale (FNE; Watson & Friend, 1969) and measures the fear of negative evaluation. Items (e.g., “I fear doing or saying the wrong thing.”) are answered on a 4-point scale ranging from 1 = *almost never true* to 4 = *almost always true*. McDonald’s  $\omega$  of the SANB-5 in this sample was .93.

### Dimensional Obsessive–Compulsive Scale (Abramowitz et al., 2010)

The Dimensional Obsessive–Compulsive Scale (DOCS) is a 20-item measure that assesses four dimensions of obsessive–compulsive disorder: Contamination, Responsibility, Unacceptable Thoughts, and Symmetry. Each of these dimensions is described by a series of bullet points (e.g., “Thoughts or feelings that you are contaminated because you came into contact with (or were nearby) a certain object or person.”). Respondents are then asked five questions regarding their



experiences of the described behaviors, thoughts, and feelings (e.g., “About how much time have you spent each day thinking about contamination and engaging in washing or cleaning behaviors because of contamination?”). These questions are answered on a 5-point scale ranging from 0 (e.g., *none at all*) to 4 (e.g., *8 hr or more each day*, with category descriptions depending on the question). Due to a technical error, the last item of the dimension Responsibility was not assessed, which is why scale scores were computed via means. The German version of the DOCS exhibited good internal consistency and convergent validity (Fink-Lamotte et al., 2021). McDonald’s  $\omega$  of the four dimension scales in this sample ranged from .87 to .93.

### **Symptom Checklist–90 Revised—Short Version** (Lutz et al., 2006)

The Symptom Checklist–90 Revised—Short Version (HSCL-11) is an 11-item German short version of the Symptom Checklist–90 Revised (Derogatis, 1986) that was developed for the evaluation of psychotherapeutic progress and focuses on symptoms of anxiety and depression. Items (e.g., “a feeling of hopelessness regarding the future”) are answered on a 5-point scale (0 = *not at all* to 4 = *very much*). McDonald’s  $\omega$  of the HSCL-11 in this sample was .94.

### **The International Trauma Questionnaire** (Cloitre et al., 2018)

The International Trauma Questionnaire (ITQ) is a self-report measure of ICD-11 PTSD and complex PTSD. For this study, only items measuring PTSD were included. Participants are first asked to name the experience that troubles them the most and to state the time point of its occurrence. Then, participants rate how much they have experienced a set of nine symptoms on a 5-point scale from 0 = *not at all* to 4 = *extremely*. These symptoms include avoidance, reexperiencing, sense of threat (each 2 items), and functional impairment (3 items). The ITQ German adult version exhibited good psychometric properties and discriminant validity in a sample of Austrian foster children (Age 10–18; Haselgruber et al., 2020). Sum scores for each of these symptom scales were computed, for which McDonald’s  $\omega$  ranged from .85 to .93.

### **Data Analysis**

A growing number of applied researchers are aware of the limitations of Cronbach’s  $\alpha$  as a measure of internal consistency due to its relatively strict assumption of a  $\tau$ -equivalent measurement model and has thus adopted McDonald’s  $\omega$ , which requires a  $\tau$ -congeneric model. In contrast, the assumption of unidimensionality (made by both Cronbach’s  $\alpha$  and McDonald’s  $\omega$ ), meaning that a single factor can sufficiently explain covariation among the indicators, mostly goes unnoticed and untested (Reise et al., 2013; Savalei & Reise, 2019). Strong violation of this assumption can lead to biased reliability estimates, which is why we started our analysis by conducting single-factor CFAs for the 15 nonsaturated IDAS-II scales<sup>1</sup> and assessing model fit. We then computed McDonald’s  $\omega$  following recommendations by Flora (2020). Since the IDAS-II employs Likert-type items, we treated them as ordered, categorical variables, and estimated all models using the weighted least-squares mean and variance adjusted (WLSMV) estimator that is applied to the polychoric correlation matrix (Flora et al., 2012).

We tested whether the internal structure of our translation is consistent with the original IDAS-II by conducting two sets of analyses: First, we tested whether scale correlations of our translation significantly differed from scale correlations reported in Watson et al. (2012) by using Jennrich’s test of the equality of two correlation matrices. Second, we conducted a multiple-factor CFA based on the EFA results reported by Watson et al. (2012). We specified all scales with  $|\lambda| > 0.2$  as indicators for the above-mentioned factors (Distress, Obsessions/Fear, and Positive Mood), thereby allowing for some cross-loadings (see Table 4 in Watson et al., 2012). Because the indicators in this analysis consisted of sum scores that showed a nonnormal distribution, this model was estimated using the robust maximum-likelihood estimator (MLR). The aforementioned analyses all were conducted on the combined sample.

Dynamic fit index (DFI; McNeish & Wolf, 2020) cutoffs for approximate fit indices were used to evaluate model fit for all CFAs. DFI cutoffs address two shortcomings of current interpretation practices of fit indices: First, general cutoff recommendations do not necessarily apply to the specific model to be evaluated. For example, the misspecifications defined by Hu and Bentler (1999) to develop cutoffs only apply to multifactor models. Nevertheless, the cutoffs derived from their simulation study are commonly applied to single-factor models (McNeish & Wolf, 2022). Second, the use of a single cutoff reduces approximate fit indices to quasihypothesis tests, whereas their purpose is to indicate the *amount* of misfit. In contrast to general cutoff recommendations, the simulation-based calculation of dynamic cutoff values is matched to the characteristics of the specific model to be evaluated. In addition, DFI cutoffs are calculated over several well-defined levels of misspecification (with higher level indicating greater misfit), thereby providing a better understanding of the *extent* of model misfit. Of particular benefit to our application, DFI cutoffs for single-factor models are specifically designed to detect deviations from unidimensionality by simulating residual correlations of  $\rho = .30$  among items. In this case, higher levels of misfit imply a larger number of residual correlations, starting with residual correlations among approximately one-third of the items at DFI Level 1 (the exact number of residual correlations depends on the number of indicators). For multifactor CFAs, simulation of misfit is based on adding one cross-loading to the model at Level 1 (and higher levels adding additional cross-loadings), similar to the approach used by Hu and Bentler (1999). Thus, DFI cutoffs represent a more reliable interpretation of goodness of fit (McNeish & Wolf, 2020). The DFI Shiny application (Wolf & McNeish, 2020) was used to calculate cutoffs for each model.<sup>2</sup>

We assessed the convergent and discriminant validity of the IDAS-II scales by calculating Pearson’s correlations of IDAS-II scales with all additional questionnaires from Sample 2. To this end, we collected data on symptom-specific questionnaires (e.g., DOCS) and well-established nonspecific questionnaires (i.e., PHQ-9, GAD-7, and HSCL-11). For the IDAS-II scales for which a corresponding

<sup>1</sup> For saturated one-factor models (consisting of only three items), fit indices cannot be computed. This includes IDAS-II scales Appetite Gain, Appetite Loss, and Checking.

<sup>2</sup> Calculations of dynamic cutoffs by the DFI Shiny application are currently based on maximum likelihood estimation. This introduces some uncertainty to the application of dynamic cutoffs to models estimated using WLSMV. Nevertheless, we believe this approach is more convincing than “traditional” fixed cutoffs, which were derived from simulation studies that do not necessarily apply to our analyses.

symptom-specific questionnaire was included in this study (e.g., SANB-5 corresponding to the Social Anxiety scale), we tested whether convergent correlations significantly differed from discriminant correlations using Williams's (1959) procedure for comparing correlation coefficients in dependent samples, which has been shown to exhibit better statistical properties than other procedures for comparing dependent correlations (Hittner et al., 2003). Since Dysphoria is considered a core dimension of depression and anxiety (Watson et al., 2007), we treated nonspecific questionnaires of depression and anxiety as corresponding measures to examine the convergent validity of this scale. To correct for multiple testing, we used the procedure proposed by Holm (1979), treating every set of comparisons for a single IDAS-II scale as a family of hypothesis tests.

All analyses were conducted using R Version 4.1.0 (R Core Team, 2021). Structural equation models and internal consistencies were computed using R packages lavaan (Rosseel, 2012) and semTools (Jorgensen et al., 2021). Significance tests for comparisons of correlations were computed using package cocor (Diedenhofen & Musch, 2015). An R script containing all analyses reported in this article (except computation of DFI cutoffs) as well as our data are available on the Open Science Framework (OSF, <https://osf.io/x3rqm/>). IDAS-II (German version) is freely available at <https://www.testarchiv.eu/de/test/9008301>. This study was approved by the ethics committee of the University of Kassel, and all participants gave written informed consent. This study was not preregistered.

## Results

### Assessment of Unidimensionality and Internal Consistency

Table 1 shows fit indices from single-factor CFAs and McDonald's  $\omega$  in the combined sample. All models failed the exact-fit hypothesis (significant  $\chi^2$ ), but approximate fit indices mostly did not indicate

severe misspecification according to dynamic cutoffs (see Table S1 in the Supplemental Material for dynamic cutoff values of each model): Six scales met at least two dynamic cutoffs at Level 1 (Ill Temper, Mania, Euphoria, Claustrophobia, Traumatic Intrusions, Traumatic Avoidance). This means that two fit indices indicate no residual correlations, while one fit index indicates possible residual correlations among approximately one-third of all scale items. Five scales met at least two DFI cutoffs at Level 2 (Dysphoria, Lassitude, Social Anxiety, Cleaning, Panic), which means that two fit indices indicate possible residual correlations among approximately one-third of all items, while one fit index indicates possible residual correlations among approximately two-thirds of all scale items. Suicidality met dynamic cutoff criteria for RMSEA at Level 2 and for CFI at Level 3, thus not providing a clear picture of model misfit. Three scales exhibited clear signs of misspecification: For scales Insomnia and Well-Being, approximate fit indices met dynamic cutoffs for CFI at Level 3, but not for RMSEA and SRMR. For scale Ordering, no approximate fit index met dynamic cutoffs at Level 3.

To examine sources of misfit and evaluate the dimensionality of the scales that did not meet at least two DFI cutoffs at Level 1, we examined standardized factor loadings, correlation residuals, and modification indices and conducted exploratory factor analyses (see Lines 26–464 of the commented R-script on OSF for a detailed report). Based on these analyses, we respecified several models, which lead to a significant improvement in fit (see Table 1, for approximate fit indices): For the Lassitude scale, we allowed a residual correlation between two items that specifically referred to experiences on mornings rather than describing general exhaustion/sleepiness (i.e., Item 43 “I had trouble waking up in the morning,” and Item 54 “I felt much worse in the morning than later in the day”). For scales Suicidality, Insomnia, Well-Being, Cleaning, and Ordering results from two-factor EFAs suggested that meaningful item parcels could be constructed. For example, Well-Being contains

**Table 1**  
*Descriptive Statistics, CFA Results, and Internal Consistencies for IDAS-II Scales*

Scale	<i>M</i> ( <i>SD</i> )	$\omega$	$\omega_{re}^a$	$\chi^2$	<i>df</i>	RMSEA [90% CI]	RMSEA <sub>re</sub> <sup>a</sup> [90% CI]	CFI	CFI <sub>re</sub> <sup>a</sup>	SRMR	SRMR <sub>re</sub> <sup>a</sup>
Dysphoria (10)	20.06 (9.17)	.936	—	138.1	35	.053 [.044, .062]	—	.995	—	.023	—
Lassitude (6)	13.09 (5.74)	.884	.872	103.9	9	.100 [.083, .118]	.082 [.064, .101]	.991	.994	.034	.028
Insomnia (6)	13.11 (6.25)	.911	.830	421.9	9	.209 [.192, .226]	.084 [.065, .104]	.967	.996	.063	.019
Suicidality (6)	7.83 (3.42)	.889	.809	120.0	9	.108 [.091, .126]	.028 [.000, .054]	.986	.999	.096	.018
Appetite gain (3)	5.66 (2.93)	.802	—	—	0	—	—	—	—	—	—
Appetite loss (3)	4.89 (2.74)	.893	—	—	0	—	—	—	—	—	—
Well-being (8)	20.18 (7.16)	.906	.837	481.8	20	.148 [.137, .160]	.092 [.079, .105]	.967	.989	.045	.024
Ill temper (4)	6.25 (3.14)	.839	—	4.5	2	.034 [.000, .078]	—	.999	—	.013	—
Mania (5)	8.27 (3.89)	.849	—	7.5	5	.022 [.000, .052]	—	1.000	—	.013	—
Euphoria (5)	7.48 (3.06)	.751	—	25.0	5	.062 [.039, .087]	—	.988	—	.033	—
Social anxiety (6)	10.22 (5.19)	.886	—	91.1	9	.093 [.076, .111]	—	.989	—	.028	—
Claustrophobia (5)	7.08 (3.67)	.907	—	31.0	5	.070 [.048, .095]	—	.998	—	.022	—
Traum. intrus. (4)	7.04 (3.94)	.890	—	14.1	2	.076 [.042, .115]	—	.998	—	.013	—
Traum. avoid. (4)	8.16 (4.22)	.896	—	5.9	2	.043 [.000, .085]	—	1.000	—	.007	—
Ordering (5)	9.20 (3.84)	.781	.728	153.9	5	.168 [.146, .192]	.025 [.000, .057]	.941	.999	.067	.013
Cleaning (7)	10.89 (4.97)	.865	.832	128.0	14	.088 [.074, .102]	.022 [.000, .042]	.985	.999	.046	.017
Checking (3)	5.33 (2.80)	.840	—	—	0	—	—	—	—	—	—
Panic (8)	12.23 (5.79)	.900	—	85.0	20	.056 [.044, .068]	—	.992	—	.029	—

*Note.* CFA = confirmatory factor analyses; IDAS-II = Inventory of Depression and Anxiety Symptoms; Traum. intrus. = Traumatic Intrusions; Traum. avoid. = Traumatic Avoidance; CFI = comparative fit index; RMSEA = root-mean-square error of approximation; SRMR = standardized root-mean-square residual. *N* = 1,054. The number of items in each scale is shown in parentheses.

<sup>a</sup>These values pertain to respecified models.

items with a future (“I looked forward to things with enjoyment,” “I felt that I had a lot to look forward to”) and a past perspective (“I felt that I had accomplished a lot”; “I was proud of myself”). Therefore, we specified higher order models with one second- and two first-order factors, which lead to significant improvements in model fit. For these models, we calculated omega higher order which “represents the reliability of a total score for measuring a single construct that influences all items, despite the multidimensional nature of the test.” (Flora, 2020, p. 496).

McDonald's  $\omega$  of all scales was above .8 except for Euphoria (.751) and Ordering (.728), with a median of .857, indicating good to excellent internal consistency of most scales. We report results from item response theory-graded response models (IRT GRM) in the [Supplemental Material](#) for readers interested in M2-based fit indices, marginal reliability, and reliability as a function of theta for each scale (see [Tables S2–S20](#)).

## Internal Structure

### Scale Correlations

[Table 2](#) shows the Pearson's correlation coefficients between all scales of the IDAS-II in the combined sample. Based on Jennrich's test, the null hypothesis stating that the correlation matrices of our translation and the original IDAS-II are equal could not be retained ( $p < .001$ ). However, differences between single correlations tended to be rather small: Differences between both IDAS-II versions ranged from  $-0.18$  to  $0.16$ , with a median absolute difference of  $0.07$ . Overall, absolute scale correlations of our translation tended to be a little larger than those reported by [Watson et al. \(2012\)](#).

Dysphoria showed higher correlations with the other IDAS-II scales than any other scale, with a median of  $r = .61$ . Similar to results from [Watson et al. \(2012\)](#), the seven symptom scales that were added in the development of IDAS-II showed somewhat lower correlations with Dysphoria (median  $r = .49$ ) than the 10 “original” IDAS scales (median  $r = .68$ ).

Correlations of the 17 specific symptom scales ranged from  $r = -.38$  (Well-being and Lassitude) to  $r = .69$  (Panic and Traumatic Intrusion), with a median value of  $r = .42$ . Scales assessing specific symptoms of depression (Lassitude, Insomnia, Suicidality, Appetite Gain, Appetite Loss) were moderately associated with each other, with correlations ranging from  $r = .30$  (Appetite Gain and Suicidality) to  $r = .61$  (Lassitude and Insomnia; ignoring the correlation between Appetite Gain and Appetite Loss of  $r = .03$ ) and a median value of  $r = .44$ . Scales assessing specific symptoms of anxiety (Social Anxiety, Claustrophobia, Traumatic Intrusions, Traumatic Avoidance, Ordering, Cleaning, Checking, Panic) were also moderately associated with each other, with correlations ranging from  $r = .69$  (Panic and Traumatic Intrusions) to  $r = .34$  (Cleaning and Traumatic Intrusions) and a median value of  $r = .5$ .

### Scale-Level CFA

We tested the three-factor structure proposed by [Watson et al. \(2012\)](#) by conducting a CFA. In contrast to validation studies of the Turkish (Irak & Albayrak, 2020) and Spanish (De la Rosa-Cáceres et al., 2020) adaptations of IDAS-II, we specified cross-loadings of several symptom scales that exhibited  $\lambda > |0.2|$  in the study by

[Watson et al. \(2012\)](#): Traumatic Avoidance, Social Anxiety, Well-Being, and Mania ([Figure 1](#)).<sup>3</sup>

The model fit indices,  $\chi^2(128) = 778.915$ ,  $p < .001$ ; comparative fit index [CFI] = .916; root-mean-square error of approximation [RMSEA] = .069, 90% confidence interval [0.066, 0.073]; standardized root-mean-square residual [SRMR] = .045, suggest a moderate to good model fit, since dynamic cutoffs in this sample resulted in  $CFI \geq .943$ ;  $RMSEA \leq .075$ ;  $SRMR \leq 0.62$ . Absolute standardized loadings ranged from  $\lambda' = .235$  (Social Anxiety on factor Obsession/Fear) to  $\lambda' = .948$  (Dysphoria on factor Distress), with a median of  $.653$ .  $R^2$  of the symptom scales ranged from .90 (Dysphoria) to .239 (Appetite Gain), with a median of .54, indicating that the hypothesized factors can explain the majority of individual differences in most indicators. The Obsessions/Fear factor exhibited a correlation of  $r = .75$  with Dysphoria and of  $r = .04$  with Positive Mood; the latter two factors were unrelated to one another ( $r = .04$ ).

A closer inspection of correlation residuals revealed that eight exceeded an absolute value of .1 (correlation residuals in brackets): Appetite Gain and (a) Appetite Loss ( $-.239$ ), (b) Well-being (.113), and (c) Euphoria (.111); Claustrophobia and (a) Social Anxiety (.160), (b) Panic (.133), and (c) Cleaning (.115); Traumatic Intrusion and Traumatic avoidance (.127); and Mania and Checking (.104). Especially for Appetite Gain and Appetite Loss, it would make sense to allow for residual correlations since both are hypothesized to be influenced by Distress, yet are mutually exclusive ([Watson et al., 2013](#)). Standardized loadings of indicators that did not load on the same factor and exhibited high correlation residuals (Panic and Claustrophobia; Mania and Checking) did not suggest a change in the specified factor structure. Nevertheless, we provide results of a three-factor EFA in the [Supplemental Material](#) (see [Tables S21](#)).

### Convergent and Discriminant Validity

[Table 3](#) shows the Pearson's correlation coefficients between IDAS-II scales and all additional questionnaires used in Sample 2. Corresponding to its definition as a core symptom of both depression and anxiety disorders, Dysphoria exhibited a higher median correlation with additional instruments than any other IDAS-II scale ( $Mdn(|r|) = .64$ , with a range of  $.38 > |r| > .86$ ). Dysphoria correlated particularly strongly with the PHQ-9 ( $r = .86$ ), the HSCL-11 ( $r = .86$ ), and the GAD-7 ( $r = .84$ ); these very strong coefficients indicate that these scales all assess the same basic symptom dimension.<sup>4</sup>

Most IDAS-II scales, for which corresponding instruments were included in the study, exhibited the highest correlation with their corresponding instrument: Correlations of (a) Well-Being with WHO-5 ( $r = .65$ ), (b) Social Anxiety with SANB-5 ( $r = .75$ ), (c) Ordering with DOCS-Symmetry ( $r = .58$ ), and (d) Cleaning with DOCS-Contamination ( $r = .66$ ) were significantly higher than with any other additional questionnaire. Traumatic Avoidance correlated strongest with ITQ-Avoidance, but this correlation did not

<sup>3</sup> We tried estimating a multidimensional CFA consisting of three second-order factors, 18 first-order factors, and 98 items as indicators. However, despite our relatively large sample, this model did not converge.

<sup>4</sup> PHQ-9 and GAD-7 ( $r = .84$ ), PHQ-9 and HSCL-11 ( $r = .84$ ), and PHQ-9 and HSCL-11 ( $r = .89$ ) correlated strongly as well.

**Table 2**  
*IDAS-II Intercorrelations*

Scale	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
1. Dysphoria	—																
2. Lassitude	<b>.79</b>	—															
3. Insomnia	<b>.68</b>	<b>.61</b>	—														
4. Suicidality	<b>.63</b>	.49	.45	—													
5. Appetite Gain	.47	.44	.41	.30	—												
6. Appetite Loss	.53	.46	.45	.44	.03	—											
7. Well-Being	−.45	−.38	−.31	−.27	−.10	−.21	—										
8. Ill Temper	<b>.67</b>	.57	.47	.54	.41	.36	−.25	—									
9. Mania	<b>.63</b>	.52	.48	.48	.39	.38	−.12	.54	—								
10. Euphoria	.00	.00	.02	.12	.13	.10	.48	.12	.33	—							
11. Social Anxiety	<b>.74</b>	.59	.52	.57	.42	.41	−.30	.56	.57	.07	—						
12. Claustrophobia	.49	.40	.36	.37	.25	.32	−.14	.39	.43	.15	<b>.60</b>	—					
13. Traumatic Intrusions	<b>.77</b>	<b>.62</b>	<b>.60</b>	.58	.36	.44	−.29	.59	.56	.06	<b>.65</b>	.48	—				
14. Traumatic Avoidance	<b>.61</b>	.52	.47	.45	.34	.37	−.23	.45	.53	.13	<b>.61</b>	.50	<b>.67</b>	—			
15. Ordering	.49	.42	.36	.39	.30	.31	−.07	.45	.58	.28	.45	.39	.45	.50	—		
16. Cleaning	.37	.36	.30	.27	.23	.28	−.05	.33	.41	.22	.41	.52	.34	.41	.49	—	
17. Checking	.57	.47	.38	.43	.36	.33	−.13	.46	<b>.61</b>	.25	.52	.43	.49	.51	<b>.65</b>	.48	—
18. Panic	<b>.76</b>	<b>.67</b>	.59	<b>.62</b>	.36	.49	−.29	<b>.60</b>	.55	.08	<b>.63</b>	.52	<b>.69</b>	.51	.40	.35	.43

*Note.* IDAS-II = Inventory of Depression and Anxiety Symptoms.  $N = 1,054$ . All correlations greater than  $|.10|$  were significant at  $p < .05$ . Large correlations ( $r \geq .6$ ) are in boldface.

significantly differ from correlations with HSCL-11, PHQ-9, DOCS-Unacceptable Thoughts, GAD, and ITQ-Reexperiencing. The correlation of Checking with DOCS-Responsibility ( $r = .53$ ) was smaller than with GAD-7 ( $r = .55$ ); however, this difference was not significant. Traumatic Intrusions correlated less with its corresponding instrument (ITQ-Reexperiencing,  $r = .68$ ) than with HSCL-11 ( $r = .74$ ), GAD-7 ( $r = .72$ ), and PHQ-9 ( $r = .70$ ), but this difference was only significant for HSCL-11.

## Discussion

The aim of this study was to develop a German translation of the IDAS-II and test its psychometric properties in a community sample. Results indicated that the IDAS-II scales (German version) are mostly unidimensional and exhibit high reliability (estimated through internal consistency). The internal structure of the German adaptation corresponded to the internal structure of the original IDAS-II, with minor differences regarding scale correlations. However, in the community sample that was investigated in this study, a CFA based on the factor structure hypothesized by Watson et al. (2012) suggested some potentially necessary model respecifications. Furthermore, correlation analyses of IDAS-II scales with other instruments of psychopathology suggested good convergent and discriminant validity for most scales that we investigated.

One of the major advantages of the IDAS-II is its ability to assess single symptoms of depression and anxiety in contrast to solely assessing global symptom severity. To investigate the reliability of these symptom scales, we not only computed a measure of internal consistency (i.e., McDonald's  $\omega$ ) but we also tested a crucial assumption (unidimensionality) for this measure through a series of single-factor CFAs. While all models failed the exact-fit hypothesis, model misfit was only moderate for most scales and a closer inspection of models showing a larger misfit suggested that the respective scales are essentially unidimensional. Internal consistency was high (i.e.,  $\omega > .8$ ) for all scales, except Euphoria and Ordering. Concerning Euphoria, similar results were obtained using

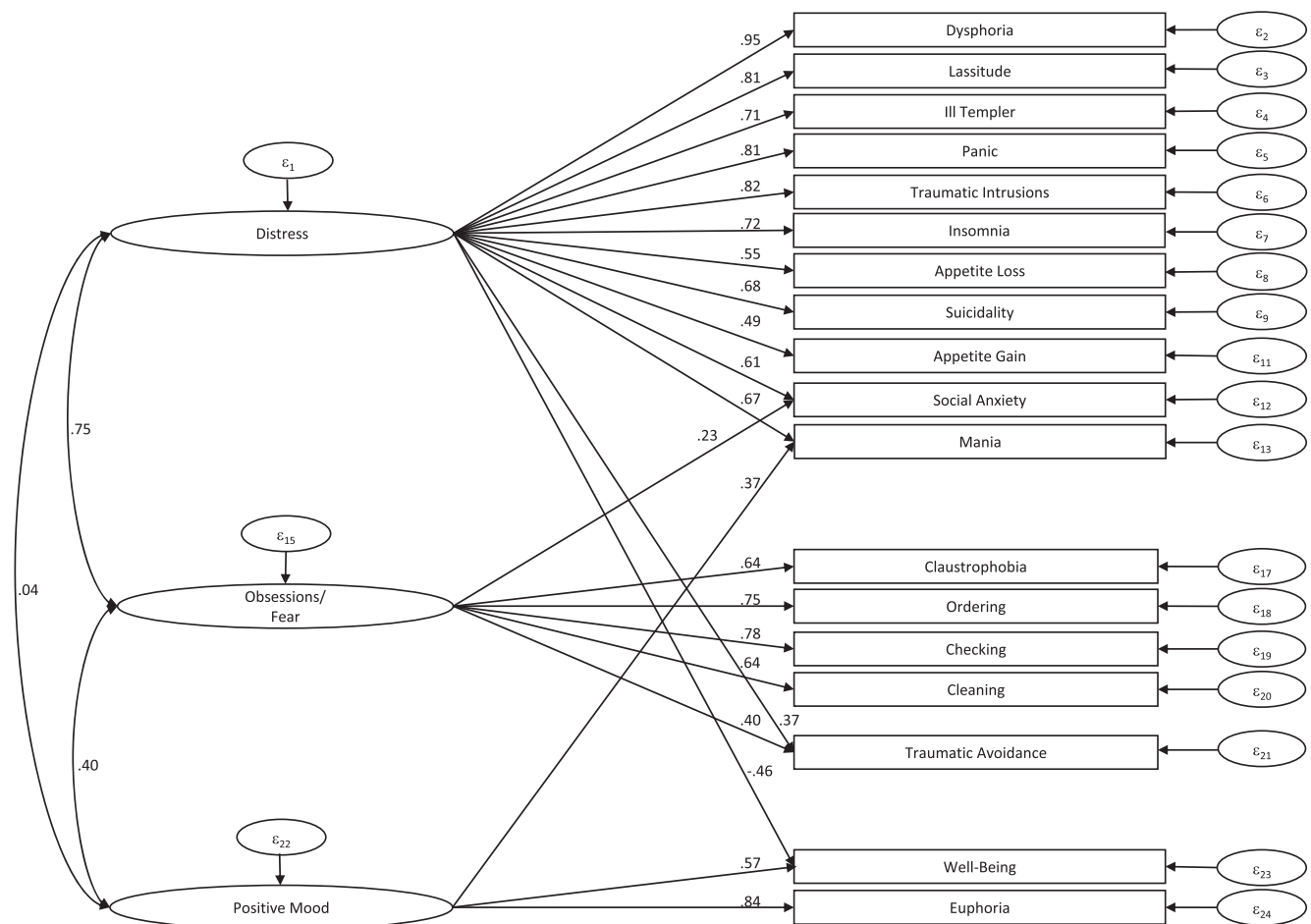
the original IDAS-II and existing translations (De la Rosa-Cáceres et al., 2020; Irak & Albayrak, 2020; Watson et al., 2012). This lower internal consistency might be due to the fact that Euphoria, as a symptom dimension of bipolar disorder, is broader than other scales and contains items tapping into elevated mood (e.g., "I felt elated for no special reason"), heightened energy (e.g., "I had so much energy it was hard for me to sit still"), and grandiosity (e.g., "I felt that I could do things that other people couldn't").

IDAS-II (German version) scale intercorrelations showed only minor differences from the original IDAS-II scale intercorrelations. Replicating results from Watson et al. (2007) and Watson et al. (2012), Dysphoria showed higher correlations with the other IDAS-II scales than any other scale. Based on this finding, Watson et al. (2007) argued that this "non-specific scale assesses the core emotional and cognitive symptoms of depression and anxiety." (Watson et al., 2007, p. 22). This notion is in line with results from Stasik-O'Brien et al. (2019) showing the utility of the Dysphoria scale as a broad screening instrument for disorders from the internalizing spectrum. Thus, while the IDAS-II is suitable for measuring several disorder-specific symptoms, researchers and clinicians might want to limit their assessment to this 10-item scale if their interest lies in the Distress/Negative Affectivity component, which is common to internalizing disorders (to a varying degree) according to different theories (Bullis et al., 2019; Watson, 2009).

Concerning the internal structure of the IDAS-II, results of a CFA did not suggest excessive misfit of a model containing the factors Distress, Obsessions/Fear, and Positive Mood. Nevertheless, some noticeable differences to results from Watson et al. (2012) emerged: While in Watson et al. (2012), Social Anxiety seemed to be an interstitial symptom, consisting equally of both the Distress and the Obsessions/Fear component, Social Anxiety exhibited a standardized loading on Distress more than 2.5 times higher than on Obsessions/Fear in our study. In contrast, Traumatic Avoidance, which loaded higher on Distress than on Obsessions/Fear in Watson et al. (2012), had very similar loadings in our sample. Factor loadings in recent structural analyses of symptoms from the



**Figure 1**  
Path Diagram of the Scale Level CFA of IDAS-II



*Note.* IDAS-II = Inventory of Depression and Anxiety Symptoms; CFA = confirmatory factor analyses. The proposed three-factor structure of the scales (Watson et al., 2012) was examined using CFA. Standardized factor loadings are depicted.  $N = 1,054$ .

internalizing spectrum (Dornbach-Bender et al., 2017) that included IDAS-II coincide with results from Watson et al. (2012). Therefore, further studies should investigate whether the above-mentioned structural deviations are due to our translation or due to the characteristics of the sample we investigated.

This study provided evidence of convergent and discriminant validity for most IDAS-II scales. Medium-to-large correlations of Dysphoria with all other questionnaires underlines the potential of this scale as a screening instrument for internalizing disorders, as evidenced by Stasik-O'Brien et al. (Stasik-O'Brien et al., 2019). In addition, most specific symptom scales showed high correlations with corresponding instruments and lower correlations with non-corresponding instruments. Only two PTSD scales and Checking deviated from this pattern. More precisely, the PTSD scales correlated highly with measures of depression (PHQ-9, HSCL-11) and anxiety (GAD-7, HSCL-11), two constructs that are strongly connected to these specific symptom dimensions (Armenta et al., 2019; Kehle et al., 2011). It is therefore an open question whether IDAS PTSD and Checking scales are unable to discriminate between the specific symptoms they are supposed to measure and more general

symptoms of depression and anxiety (as measured by the PHQ-9, HSCL-11, and GAD-7), or whether high correlations simply reflect true associations between these constructs. In addition, correlations of these three scales with corresponding and noncorresponding instruments showed only minor differences.

Overall, results indicate that this German adaptation exhibits similar psychometric properties as the original IDAS-II. Thus, the IDAS-II (German version) provides researchers and clinicians with a comprehensive measure for disorders from the internalizing spectrum that covers a wide range of symptoms in an economic manner. The IDAS-II is now freely available for use in English-, Turkish-, Spanish-, and German-speaking populations, making it possible to assess symptoms dimensionally in a wide variety of samples. To support the clinical use of the IDAS-II (German version), we provide norms based on our data in the [Supplemental Material](#) available online.

### Limitations and Future Directions

This study was limited with respect to several factors, calling for further validation of the German adaptation of the IDAS-II. Most

**Table 3**  
*IDAS-II Discriminant and Convergent Validity*

Scale	PHQ-9	HSCL	GAD7	SANB5	WHO5	DOCS-C	DOCS-R	DOCS-UT	DOCS-S	ITQ-R	ITQ-A	ITQ-ST
Dysphoria	.86	.86	.84	.69	-.68	.38	.59	.68	.47	.57	.55	.57
Lassitude	.79	.71	.67	.54	-.65	.32	.50	.55	.41	.48	.48	.46
Insomnia	.66	.64	.60	.45	-.53	.22	.41	.48	.32	.51	.41	.46
Suicidality	.66	.72	.63	.49	-.40	.27	.46	.53	.41	.47	.45	.49
Appetite Gain	.39	.32	.39	.40	-.32	.08	.20	.24	.12	.26	.23	.26
Appetite Loss	.57	.57	.49	.38	-.37	.36	.45	.43	.40	.28	.36	.35
Well-Being	-.48	-.45	-.39	-.36	.65	-.05	-.22	-.29	-.15	-.21	-.26	-.21
Ill Temper	.64	.66	.71	.52	-.45	.30	.45	.52	.46	.43	.37	.48
Mania	.59	.58	.61	.49	-.38	.37	.52	.53	.47	.48	.46	.46
Euphoria	-.03	-.01	.03	-.01	.23	.18	.14	.08	.16	.08	.07	.14
Social Anxiety	.67	.69	.67	.75	-.49	.32	.50	.59	.45	.49	.55	.50
Claustrophobia	.45	.49	.46	.39	-.32	.36	.42	.48	.42	.40	.45	.50
Tr. Intrusions	.70	.74	.72	.59	-.49	.30	.60	.67	.46	.68	.57	.58
Tr. Avoidance	.57	.59	.56	.49	-.41	.31	.49	.56	.42	.55	.61	.47
Ordering	.47	.48	.48	.42	-.27	.46	.48	.50	.58	.39	.46	.41
Cleaning	.34	.35	.35	.30	-.19	.66	.36	.37	.42	.27	.35	.31
Checking	.50	.49	.54	.46	-.33	.39	.53	.50	.48	.40	.43	.43
Panic	.70	.74	.68	.46	-.48	.34	.54	.58	.41	.53	.51	.59

*Note.*  $N = 550$ ;  $N = 310$  for ITQ subscales since only those who had already experienced a traumatic experience answered the ITQ. All correlations greater than .141 were significant on  $p < .05$ . For ITQ subscales, all correlations greater than .121 were significant on  $p < .05$ . Convergent correlations are in italics. PHQ-9 = Patient Health Questionnaire-9; HSCL-11 = Symptom Checklist-90 Revised—Short Version; GAD-7 = Generalized Anxiety Disorder Scale-7; IDAS-II = Inventory of Depression and Anxiety Symptoms; DOCS = Dimensional Obsessive-Compulsive Scale; SANB-5 = Fear of Negative Evaluation Scale-5; WHO-5 = WHO-Five Well-Being Index; DOCS-C = DOCS subscale-Contamination; DOCS-R = DOCS subscale-Responsibility; DOCS-UT = DOCS subscale-Unacceptable Thoughts; DOCS-S = DOCS subscale-Symmetry; ITQ = International Trauma Questionnaire; ITQ-R = ITQ subscale-Reexperience; ITQ-A = ITQ subscale-Avoidance; ITQ-ST = ITQ subscale-Sense of Threat.

importantly, this validation study was based on a community sample, thereby limiting the generalizability of results to clinical settings. Based on the validation of the original English scale and related research, it seems probable that scales would be similarly correlated within clinical populations (Watson et al., 2007) and perhaps even within different demographic and diagnostic groups (Nelson et al., 2018; Watson et al., 2012). Nonetheless, future studies should empirically test this claim and specifically investigate psychometric properties of the IDAS-II (German version) in more heterogeneous samples, including psychiatric patients with diagnosable mental health conditions of varying severity and chronicity. It should be noted, however, that the symptom burden in our sample can be assumed to be somewhat larger than in other community samples, as more subjects reported being in psychotherapeutic or psychiatric treatment than the general population in Germany (Larisch et al., 2013).

Furthermore, convergent and discriminant validity was only investigated for a subset of IDAS-II scales and was restricted to self-report measures. A full validation of IDAS-II (German version) should include self-report as well as interview measures covering all symptom dimensions assessed by the instrument. This is not an easy task to achieve, since most measures of internalizing disorders are restricted to global scores of severity and do not assess specific symptoms (e.g., Lassitude) with sufficient reliability. The IDAS-II symptom-specific scales should also be examined regarding their incremental validity beyond the contributions of global scores of severity using external criteria like psychosocial functioning. An extended validation should also include the investigation of the temporal stability of IDAS-II scales along with an assessment of their sensitivity to change. Investigation of the temporal stability of IDAS-II scales was thus far limited to 1-week intervals. Future studies should extend this interval and ideally compare test-retest

correlations across different time frames. Finally, future studies might examine the specificity of the IDAS-II items using alternative modeling methods such as bifactor-IRT.

## References

- Abramowitz, J. S., Deacon, B. J., Olatunji, B. O., Wheaton, M. G., Berman, N. C., Losardo, D., Timpano, K. R., McGrath, P. B., Riemann, B. C., Adams, T., Björgvinsson, T., Storch, E. A., & Hale, L. R. (2010). Assessment of obsessive-compulsive symptom dimensions: Development and evaluation of the Dimensional Obsessive-Compulsive Scale. *Psychological Assessment*, 22(1), 180–198. <https://doi.org/10.1037/a0018260>
- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). <https://doi.org/10.1176/appi.books.9780890425596>
- Armenta, R. F., Walter, K. H., Geronimo-Hara, T. R., Porter, B., Stander, V. A., LeardMann, C. A., & the Millennium Cohort Study Team. (2019). Longitudinal trajectories of comorbid PTSD and depression symptoms among U.S. service members and veterans. *BMC Psychiatry*, 19(1), Article 396. <https://doi.org/10.1186/s12888-019-2375-1>
- Brähler, E., Mühlen, H., Albani, C., & Schmidt, S. (2007). Teststatistische Prüfung und Normierung der deutschen Versionen des EUROHIS-QOL Lebensqualität-Index und des WHO-5 Wohlbefindens-Index. *Diagnostica*, 53(2), 83–96. <https://doi.org/10.1026/0012-1924.53.2.83>
- Broman-Fulks, J. J., Deacon, B. J., Olatunji, B. O., Bondy, C. L., Abramowitz, J. S., & Tolin, D. F. (2010). Categorical or dimensional: A reanalysis of the anxiety sensitivity construct. *Behavior Therapy*, 41(2), 154–171. <https://doi.org/10.1016/j.beth.2009.02.005>
- Brown, T. A., Campbell, L. A., Lehman, C. L., Grisham, J. R., & Mancill, R. B. (2001). Current and lifetime comorbidity of the DSM-IV anxiety and mood disorders in a large clinical sample. *Journal of Abnormal Psychology*, 110(4), 585–599. <https://doi.org/10.1037/0021-843X.110.4.585>
- Bullis, J. R., Boettcher, H., Sauer-Zavala, S., Farchione, T. J., & Barlow, D. H. (2019). What is an emotional disorder? A transdiagnostic mechanistic definition with implications for assessment, treatment, and

- prevention. *Clinical Psychology: Science and Practice*, 26(2), Article e12278. <https://doi.org/10.1111/cpsp.12278>
- Chmielewski, M., Clark, L. A., Bagby, R. M., & Watson, D. (2015). Method matters: Understanding diagnostic reliability in DSM-IV and DSM-5. *Journal of Abnormal Psychology*, 124(3), 764–769. <https://doi.org/10.1037/abn0000069>
- Clark, L. A., Cuthbert, B., Lewis-Fernández, R., Narrow, W. E., & Reed, G. M. (2017). Three approaches to understanding and classifying mental disorder: ICD-11, DSM-5, and the National Institute of Mental Health's Research Domain Criteria (RDoC). *Psychological Science in the Public Interest*, 18(2), 72–145. <https://doi.org/10.1177/1529100617727266>
- Cloitre, M., Shevlin, M., Brewin, C. R., Bisson, J. I., Roberts, N. P., Maercker, A., Karatzias, T., & Hyland, P. (2018). The International Trauma Questionnaire: Development of a self-report measure of ICD-11 PTSD and complex PTSD. *Acta Psychiatrica Scandinavica*, 138(6), 536–546. <https://doi.org/10.1111/acps.12956>
- Conway, C. C., Forbes, M. K., & South, S. C. (2022). A hierarchical taxonomy of psychopathology (HiTOP) primer for mental health researchers. *Clinical Psychopathology Science*, 10(2), 236–258. <https://doi.org/10.1177/21677026211017834>
- De la Rosa-Cáceres, A., Stasik-O'Brien, S., Rojas, A. J., Sanchez-Garcia, M., Lozano, O. M., & Díaz-Batanero, C. (2020). Spanish Adaptation of the Inventory of Depression and Anxiety Symptoms (IDAS-II) and a study of its psychometric properties. *Journal of Affective Disorders*, 271, 81–90. <https://doi.org/10.1016/j.jad.2020.03.187>
- Derogatis, L. R. (1986). *SCL 90 R Administration, Scoring and Procedures Manual II for the Revised Version and Other Instruments of the Psychopathology Rating Scale Series*. Clinical Psychometric Research.
- Diedenhofen, B., & Musch, J. (2015). cocor: A comprehensive solution for the statistical comparison of correlations. *PLOS ONE*, 10(3), Article e0121945. <https://doi.org/10.1371/journal.pone.0121945>
- Dornbach-Bender, A., Ruggero, C. J., Waszczuk, M. A., Gamez, W., Watson, D., & Kotov, R. (2017). Mapping emotional disorders at the finest level: Convergent validity and joint structure based on alternative measures. *Comprehensive Psychiatry*, 79, 31–39. <https://doi.org/10.1016/j.comppsy.2017.06.011>
- Everitt, B. S. (1975). Multivariate analysis: The need for data, and other problems. *The British Journal of Psychiatry*, 126(3), 237–240. <https://doi.org/10.1192/bjp.126.3.237>
- Fink-Lamotte, J., Jahn, I., Stierle, C., Kühne, F., Lincoln, T., Stengler, K., & Exner, C. (2021). Die Validierung der Dimensional Obsessive-Compulsive Scale (DOCS) an einer deutschsprachigen Stichprobe. *Verhaltenstherapie*, 31(2), 119–131. <https://doi.org/10.1159/000510093>
- Flora, D. B. (2020). Your coefficient alpha is probably wrong, but which coefficient omega is right? A tutorial on using R to obtain better reliability estimates. *Advances in Methods and Practices in Psychological Science*, 3(4), 484–501. <https://doi.org/10.1177/2515245920951747>
- Flora, D. B., Labrish, C., & Chalmers, R. P. (2012). Old and new ideas for data screening and assumption testing for exploratory and confirmatory factor analysis. *Frontiers in Psychology*, 3, Article 55. <https://doi.org/10.3389/fpsyg.2012.00055>
- Fried, E. I., Flake, J. K., & Robinaugh, D. J. (2022). Revisiting the theoretical and methodological foundations of depression measurement. *Nature Reviews Psychology*, 1(6), 358–368. <https://doi.org/10.1038/s44159-022-00050-2>
- Fried, E. I., & Nesse, R. M. (2015). Depression is not a consistent syndrome: An investigation of unique symptom patterns in the STAR\*D study. *Journal of Affective Disorders*, 172, 96–102. <https://doi.org/10.1016/j.jad.2014.10.010>
- Galatzer-Levy, I. R., & Bryant, R. A. (2013). 636,120 ways to have posttraumatic stress disorder. *Perspectives on Psychological Science*, 8(6), 651–662. <https://doi.org/10.1177/1745691613504115>
- Gräfe, K., Zipfel, S., Herzog, W., & Löwe, B. (2004). Screening psychischer Störungen mit dem "Gesundheitsfragebogen für Patienten (PHQ-D)". *Diagnostica*, 50(4), 171–181. <https://doi.org/10.1026/0012-1924.50.4.171>
- Hankin, B. L., Fraley, R. C., Lahey, B. B., & Waldman, I. D. (2005). Is depression best viewed as a continuum or discrete category? A taxometric analysis of childhood and adolescent depression in a population-based sample. *Journal of Abnormal Psychology*, 114(1), 96–110. <https://doi.org/10.1037/0021-843X.114.1.96>
- Haselgruber, A., Sölva, K., & Lueger-Schuster, B. (2020). Validation of ICD-11 PTSD and complex PTSD in foster children using the International Trauma Questionnaire. *Acta Psychiatrica Scandinavica*, 141(1), 60–73. <https://doi.org/10.1111/acps.13100>
- Haslam, N., McGrath, M. J., Viechtbauer, W., & Kuppens, P. (2020). Dimensions over categories: A meta-analysis of taxometric research. *Psychological Medicine*, 50(9), 1418–1432. <https://doi.org/10.1017/S003329172000183X>
- Hittner, J. B., May, K., & Silver, N. C. (2003). A Monte Carlo evaluation of tests for comparing dependent correlations. *The Journal of General Psychology*, 130(2), 149–168. <https://doi.org/10.1080/00221300309601282>
- Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, 6(2), 65–70.
- Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, 6(1), 1–55.
- Irak, M., & Albayrak, E. O. (2020). Psychometric properties of the expanded version of the inventory of depression and anxiety symptoms in a Turkish population. *Psychological Reports*, 123(2), 517–545. <https://doi.org/10.1177/0033294118813844>
- Jorgensen, T. D., Pornprasertmanit, S., Schoemann, A. M., & Rosseel, Y. (2021). *SemTools: Useful tools for structural equation modeling* [R package Version 0.5-5]. <https://CRAN.R-project.org/package=semTools>
- Kaufman, J., & Charney, D. (2000). Comorbidity of mood and anxiety disorders. *Depression and Anxiety*, 12(Suppl. 1), 69–76. [https://doi.org/10.1002/1520-6394\(2000\)12:1+<69::AID-DA9>3.0.CO;2-K](https://doi.org/10.1002/1520-6394(2000)12:1+<69::AID-DA9>3.0.CO;2-K)
- Kehle, S. M., Reddy, M. K., Ferrier-Auerbach, A. G., Erbes, C. R., Arbis, P. A., & Polusny, M. A. (2011). Psychiatric diagnoses, comorbidity, and functioning in National Guard troops deployed to Iraq. *Journal of Psychiatric Research*, 45(1), 126–132. <https://doi.org/10.1016/j.jpsychires.2010.05.013>
- Kemper, C. J., Lutz, J., & Neuser, J. (2011). SANB-5. Skala Angst vor negativer Bewertung-5 [Verfahrensdokumentation und Fragebogen]. In Leibniz-Institut für Psychologie (ZPID) (Ed.), *Open test archive*. Leibniz Institute for Psychology Information. <https://doi.org/10.23668/psycharchives.416>
- Kertz, S. J., McHugh, R. K., Lee, J., & Björgvinsson, T. (2014). Examining the latent structure of worry and generalized anxiety in a clinical sample. *Journal of Anxiety Disorders*, 28(1), 8–15. <https://doi.org/10.1016/j.janxdis.2013.11.003>
- Kollman, D. M., Brown, T. A., Liverant, G. I., & Hofmann, S. G. (2006). A taxometric investigation of the latent structure of social anxiety disorder in outpatients with anxiety and mood disorders. *Depression and Anxiety*, 23(4), 190–199. <https://doi.org/10.1002/da.20158>
- Kotov, R., Krueger, R. F., Watson, D., Achenbach, T. M., Althoff, R. R., Bagby, R. M., Brown, T. A., Carpenter, W. T., Caspi, A., Clark, L. A., Eaton, N. R., Forbes, M. K., Forbush, K. T., Goldberg, D., Hasin, D., Hyman, S. E., Ivanova, M. Y., Lynam, D. R., Markon, K., ... Zimmerman, M. (2017). The Hierarchical Taxonomy of Psychopathology (HiTOP): A dimensional alternative to traditional nosologies. *Journal of Abnormal Psychology*, 126(4), 454–477. <https://doi.org/10.1037/abn0000258>
- 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>
- Lamers, F., van Oppen, P., Comijs, H. C., Smit, J. H., Spinhoven, P., van Balkom, A. J., Nolen, W. A., Zitman, F. G., Beekman, A. T., & Penninx, B. W. (2011). Comorbidity patterns of anxiety and depressive disorders in a large cohort study: The Netherlands Study of Depression and Anxiety (NESDA). *The Journal of Clinical Psychiatry*, 72(3), 341–348. <https://doi.org/10.4088/JCP.10m06176blu>



- Larisch, A., Heuft, G., Engbrink, S., Brähler, E., Herzog, W., & Kruse, J. (2013). Behandlung Psychischer und Psychosomatischer Beschwerden—Inanspruchnahme, Erwartungen und Kenntnisse der Allgemeinbevölkerung in Deutschland. *Zeitschrift für Psychosomatische Medizin und Psychotherapie*, 59(2), 153–169. <https://doi.org/10.13109/zptm.2013.59.2.153>
- Löwe, B., Decker, O., Müller, S., Brähler, E., Schellberg, D., Herzog, W., & Herzberg, P. Y. (2008). Validation and standardization of the Generalized Anxiety Disorder Screener (GAD-7) in the general population. *Medical Care*, 46(3), 266–274. <https://doi.org/10.1097/MLR.0b013e318160d093>
- Lutz, W., Tholen, S., Schürch, E., & Berking, M. (2006). Die Entwicklung, Validierung und Reliabilität von Kurzformen gängiger psychometrischer Instrumente zur Evaluation destherapeutischen Fortschritts in Psychotherapie und Psychiatrie. *Diagnostica*, 52(1), 11–25. <https://doi.org/10.1026/0012-1924.52.1.11>
- Markon, K. E., Chmielewski, M., & Miller, C. J. (2011). The reliability and validity of discrete and continuous measures of psychopathology: A quantitative review. *Psychological Bulletin*, 137(5), 856–879. <https://doi.org/10.1037/a0023678>
- McDermut, W., Mattia, J., & Zimmerman, M. (2001). Comorbidity burden and its impact on psychosocial morbidity in depressed outpatients. *Journal of Affective Disorders*, 65(3), 289–295. [https://doi.org/10.1016/S0165-0327\(00\)00220-2](https://doi.org/10.1016/S0165-0327(00)00220-2)
- McNeish, D., & Wolf, M. G. (2020). Dynamic fit index cutoffs for confirmatory factor analysis models. *Psychological Methods*. Advance online publication. <https://doi.org/10.1037/met0000425>
- McNeish, D., & Wolf, M. G. (2022). Dynamic fit index cutoffs for one-factor models. *Behavior Research Methods*. Advance online publication. <https://doi.org/10.3758/s13428-022-01847-y>
- Mineka, S., Watson, D., & Clark, L. A. (1998). Comorbidity of anxiety and unipolar mood disorders. *Annual Review of Clinical Psychology*, 49(1), 377–412. <https://doi.org/10.1146/annurev.psych.49.1.377>
- Mundfrom, D. J., Shaw, D. G., & Ke, T. L. (2005). Minimum sample size recommendations for conducting factor analyses. *International Journal of Testing*, 5(2), 159–168. [https://doi.org/10.1207/s15327574ijt0502\\_4](https://doi.org/10.1207/s15327574ijt0502_4)
- Nelson, G. H., O'Hara, M. W., & Watson, D. (2018). National norms for the expanded version of the inventory of depression and anxiety symptoms (IDAS-II). *Journal of Clinical Psychology*, 74(6), 953–968. <https://doi.org/10.1002/jclp.22560>
- Prisciandaro, J. J., & Roberts, J. E. (2011). Evidence for the continuous latent structure of mania in the Epidemiologic Catchment Area from multiple latent structure and construct validation methodologies. *Psychological Medicine*, 41(3), 575–588. <https://doi.org/10.1017/S0033291710001078>
- Prisciandaro, J. J., & Tolliver, B. K. (2015). Evidence for the continuous latent structure of mania and depression in out-patients with bipolar disorder: Results from the Systematic Treatment Enhancement Program for Bipolar Disorder (STEP-BD). *Psychological Medicine*, 45(12), 2595–2603. <https://doi.org/10.1017/S0033291715000513>
- Radloff, L. S. (1977). The CES-D Scale: A Self-Report Depression Scale for research in the general population. *Applied Psychological Measurement*, 1(3), 385–401. <https://doi.org/10.1177/014662167700100306>
- R Core Team. (2021). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. <https://www.R-project.org/>
- Regier, D. A., Narrow, W. E., Clarke, D. E., Kraemer, H. C., Kuramoto, S. J., Kuhl, E. A., & Kupfer, D. J. (2013). DSM-5 field trials in the United States and Canada, Part II: Test–retest reliability of selected categorical diagnoses. *The American Journal of Psychiatry*, 170(1), 59–70. <https://doi.org/10.1176/appi.ajp.2012.12070999>
- Reise, S. P., Bonifay, W. E., & Haviland, M. G. (2013). Scoring and modeling psychological measures in the presence of multidimensionality. *Journal of Personality Assessment*, 95(2), 129–140. <https://doi.org/10.1080/00223891.2012.725437>
- Rosseel, Y. (2012). lavaan: An R Package for Structural Equation Modeling. *Journal of Statistical Software*, 48(2), 1–36. <https://doi.org/10.18637/jss.v048.i02>
- Ruscio, J., & Ruscio, A. M. (2000). Informing the continuity controversy: A taxometric analysis of depression. *Journal of Abnormal Psychology*, 109(3), 473–487. <https://doi.org/10.1037/0021-843X.109.3.473>
- Savalei, V., & Reise, S. P. (2019). Don't forget the model in your model-based reliability coefficients: A reply to McNeish. *Collabra. Psychology*, 5(1), Article 36. <https://doi.org/10.1525/collabra.247>
- Schönbrodt, F. D., & Perugini, M. (2013). At what sample size do correlations stabilize? *Journal of Research in Personality*, 47(5), 609–612. <https://doi.org/10.1016/j.jrp.2013.05.009>
- Spitzer, R. L. (1999). Validation and utility of a self-report version of PRIME-MD The PHQ Primary Care Study. *JAMA*, 282(18), 1737–1744. <https://doi.org/10.1001/jama.282.18.1737>
- Spitzer, R. L., Kroenke, K., Williams, J. B. W., & 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>
- Stasik-O'Brien, S. M., Brock, R. L., Chmielewski, M., Naragon-Gainey, K., Koffel, E., McDade-Montez, E., O'Hara, M. W., & Watson, D. (2019). Clinical utility of the inventory of depression and anxiety symptoms (IDAS). *Assessment*, 26(5), 944–960. <https://doi.org/10.1177/1073191118790036>
- Stochl, J., Fried, E. I., Fritz, J., Croudace, T. J., Russo, D. A., Knight, C., Jones, P. B., & Perez, J. (2022). On dimensionality, measurement invariance, and suitability of sum scores for the PHQ-9 and the GAD-7. *Assessment*, 29(3), 355–366. <https://doi.org/10.1177/1073191120976863>
- Watson, D. (2009). Differentiating the mood and anxiety disorders: A quadripartite model. *Annual Review of Clinical Psychology*, 5(1), 221–247. <https://doi.org/10.1146/annurev.clinpsy.032408.153510>
- Watson, D., Clark, L. A., Chmielewski, M., & Kotov, R. (2013). The value of suppressor effects in explicating the construct validity of symptom measures. *Psychological Assessment*, 25(3), 929–941. <https://doi.org/10.1037/a0032781>
- Watson, D., & Friend, R. (1969). Measurement of social-evaluative anxiety. *Journal of Consulting and Clinical Psychology*, 33(4), 448–457. <https://doi.org/10.1037/h0027806>
- Watson, D., O'Hara, M. W., & Chmielewski, M. (2008). Further validation of the IDAS: Evidence of convergent, discriminant, criterion, and incremental validity. *Psychological Assessment*, 20(3), 248–259. <https://doi.org/10.1037/a0012570>
- Watson, D., O'Hara, M. W., Naragon-Gainey, K., Koffel, E., Chmielewski, M., Kotov, R., Stasik, S. M., & Ruggero, C. J. (2012). Development and validation of new anxiety and bipolar symptom scales for an expanded version of the IDAS (the IDAS-II). *Assessment*, 19(4), 399–420. <https://doi.org/10.1177/1073191112449857>
- Watson, D., O'Hara, M. W., Simms, L. J., Kotov, R., Chmielewski, M., McDade-Montez, E. A., Gamez, W., & Stuart, S. (2007). Development and validation of the Inventory of Depression and Anxiety Symptoms (IDAS). *Psychological Assessment*, 19(3), 253–268. <https://doi.org/10.1037/1040-3590.19.3.253>
- Williams, E. J. (1959). The comparison of regression variables. *Journal of the Royal Statistical Society. Series B. Methodological*, 21(2), 396–399. <https://doi.org/10.1111/j.2517-6161.1959.tb00346.x>
- Wolf, M. G., & McNeish, D. (2020). *Dynamic model fit* [R Shiny application Version 1.1.0].
- World Health Organization. (1998). *Wellbeing measures in primary health care/The depcare project*. WHO Regional Office for Europe.

Received December 22, 2021

Revision received August 29, 2022

Accepted September 6, 2022 ■