Development and Initial Validation of a Self-Report Measure to Assess Eating Disorder-Specific Interoceptive Perception

Julie Ortmann[1], Annika P. C. Lutz[1], Gitta Rose[2], Christian Happ[3], Claus Vögele[1],
André Schulz[1], and Zoé van Dyck[1, 4]
[1] Department of Behavioural and Cognitive Sciences, University of Luxembourg
[2] Independent Psychotherapist/Clinical Psychologist, Luxembourg City, Luxembourg
[3] Yolande Asbl, Luxembourg City, Luxembourg
[4] Centre Hospitalier Neuro-Psychiatrique, Ettelbruck, Luxembourg

Interoceptive deficits—particularly with respect to the perception of emotions, hunger, and satiety—constitute important targets for intervention in eating disorders (EDs). Suitable self-report measures to identify these deficits, however, are lacking. We, therefore, developed and validated a multidimensional questionnaire to assess eating disorder-specific interoceptive perception (EDIP) in terms of the ability to perceive and discriminate between emotions, hunger, and satiety. In two independent samples with a total of 2058 individuals (22.74% with self-reported EDs), exploratory and confirmatory factor analyses revealed a four-factor solution of the EDIP Questionnaire (EDIP-Q) with the subscales Emotions, Hunger, Satiety, and Discrimination. The EDIP-Q has sound psychometric properties and was related to convergent questionnaires but unrelated to divergent self-report measures, supporting its construct validity. Participants with self-reported EDs had significantly lower EDIP-Q scores compared to participants without self-reported ED diagnosis. While individuals with self-reported anorexia nervosa (AN), bulimia nervosa (BN), and binge-eating disorder (BED) report similar difficulties in perceiving emotions, participants with BN and BED report greater difficulties in perceiving satiety and differentiating between hunger and emotional states compared to participants with AN. In contrast, individuals with AN report higher sensibility to satiety but lower sensibility to hunger compared to individuals with BN and BED. The EDIP-Q is a valuable clinical tool to establish profiles of deficits in EDIP that provide the basis for developing more targeted treatment approaches for EDs.

Public Significance Statement
The present studies report on the development and validation of the Eating Disorder-Specific Interoceptive Perception Questionnaire (EDIP-Q) in a large sample of individuals with and without a self-reported eating disorder diagnosis. The results suggest that individuals with anorexia nervosa, bulimia nervosa, and binge-eating disorder differ in their perception and discrimination of emotions, hunger, and satiety, highlighting the need to develop new and more targeted interventions for eating disorders.

Keywords: eating disorders, interoception, scale development, self-report

Supplemental materials: https://doi.org/10.1037/pas0001283.supp

Eating disorders (EDs) are defined by severe disturbances in eating and eating-related behaviors, as well as an unhealthy relationship with food, resulting in impaired physical health, psychosocial functioning, and quality of life (American Psychiatric Association [APA], 2013). Despite the availability of empirically supported interventions, relapse and mortality rates are still high in the long term (Fichter et al., 2017; Quadflieg & Fichter, 2019; Fichter & Quadflieg, 2016). This underlines the importance of improving our understanding of the factors that contribute to the development and maintenance of ED symptoms. Several lines of research highlight that interoceptive deficits constitute a key transdiagnostic feature in EDs (e.g., Monteleone & Cascino, 2021) and may, therefore, represent important targets for intervention.

Interoception refers to the ability to perceive the internal state of the body and comprises the processing of afferent signals from visceral organs and their integration and representation at the cortical level (Cameron, 2001). Established self-report instruments to assess interoceptive processes in EDs (e.g., Eating Disorder Inventory [EDI]; Garner et al., 1983), however, typically focus on the awareness of *emotional* states, thereby omitting a core aspect of interoceptive deficits in EDs, that is, impaired perception of *hunger* and *satiety* cues. Although the literature points to a differential contribution of emotions, hunger, and satiety signals to ED symptoms (Poovey et al., 2022; van Dyck et al., 2016), there is no instrument that allows for its differential assessment. The aims of the present studies are, therefore, (a) to develop and validate a new self-report measure to assess eating disorder-specific interoceptive perception (EDIP) in terms of the ability to perceive emotions, hunger, and satiety and to discriminate between those physiological states and (b) to investigate differences in EDIP in individuals with a self-reported diagnosis of anorexia nervosa (AN), bulimia nervosa (BN), and binge-eating disorder (BED).

Interoception encompasses signal processing originating from multiple organ systems involved in physical homeostasis, such as the cardiovascular, pulmonary, and gastrointestinal systems (Khalsa et al., 2018). Contemporary models distinguish several dimensions of interoception (Forkmann et al., 2016; S. N. Garfinkel et al., 2015; Khalsa et al., 2018; Murphy et al., 2019; Suksasilp & Garfinkel, 2022) based on the level of processing in the peripheral and central nervous system and its assessment approach, including psycho-physiological, behavioral, and self-report measures. Using different methodological approaches, altered interoceptive processes have been found in eating and weight disorders (Martin et al., 2019) and may contribute to the heterogeneous clinical picture, ranging from severe restrictive eating to uncontrolled overeating. Particularly deficits in emotional processing and in the perception of hunger and satiety have been described. A recent meta-analysis showed that individuals with EDs exhibited higher levels of alexithymia, that is, greater difficulties in identifying or describing their emotions (Westwood et al., 2017). Although this emphasizes a general impairment in emotional processing across the ED spectrum, a more heterogeneous picture emerges in the perception of hunger and satiety signals. Following a standardized meal, patients with AN report earlier satiety and postprandial fullness compared to healthy controls (HC; P. E. Garfinkel, 1974; P. E. Garfinkel et al., 1978, 1979; Klastrup et al., 2020). In contrast, individuals with BN and BED are less sensitive to satiety and gastric distention compared to HC (van Dyck et al., 2021).

In the self-report domain, empirical findings on interoceptive processing in EDs were primarily obtained using the interoceptive awareness (IA) subscale of the EDI (EDI-IA), which measures self-reported difficulties in "recognizing and accurately identifying emotions and sensations of hunger or satiety" (Garner et al., 1983, p.18). Results of a recent meta-analysis by Jenkinson et al. (2018) emphasize the presence of important, self-reported interoceptive deficits in individuals with EDs and the improvement of these deficits after recovery. Nevertheless, the use of EDI-IA as a measure of interoception in EDs is questionable for two reasons: First, several items of the subscale are more likely to reflect ED-specific symptomatology rather than interoceptive abilities (e.g., "When I am upset, I worry that I will start eating."). Second, most items measure interoceptive deficits in relation to emotional states (e.g., "When I am upset, I don't know if I am sad, frightened, or angry."), whereas only one item addresses the impaired perception of interoceptive signals of hunger and satiety ("I get confused as to whether or not I am hungry."). The EDI-IA, therefore, does not distinguish between interoceptive facets, although these may contribute in different ways to the development and maintenance of ED symptoms.

To reveal the role of potentially altered interoceptive processing in the etiology and maintenance of EDs, a novel instrument is required that is ED-specific and allows for the differential assessment of interoceptive deficits in EDs. Study 1 concerned the development of the EDIP Questionnaire (EDIP-Q), including the investigation of its factor structure using exploratory factor analysis (EFA) and psychometric properties (i.e., item analyses, internal consistency, and construct validity). Study 2 aimed at confirming the factor structure derived from Study 1 using confirmatory factor analysis (CFA) and to provide further evidence of its construct validity. As a second objective, we examined differences in EDIP, including sex differences and differences between participants with a self-reported current and past ED diagnosis (Diagnostic Status I), as well as between different types of self-reported EDs (Diagnostic Status II), compared with healthy participants.

Consistent with the definition of EDIP, we expected the items to load on four independent but interrelated factors forming the subscales of the EDIP-Q. Furthermore, we expected the EDIP-Q to have sound psychometric properties and to be related to other scales measuring similar or related constructs. We hypothesized that EDIP-Q scores would be inversely related to ED symptom severity, maladaptive eating behavior, and poor awareness of bodily states, particularly in relation to emotions, hunger, and satiety. Regarding group differences, we hypothesized that the EDIP-Q would discriminate well between individuals with and without a self-reported ED diagnosis. More specifically, individuals with a self-reported ED diagnosis would report lower subscale scores compared to HC, indicating greater deficits in EDIP. To the best of our knowledge, there are no studies that systematically examine differences in emotional processing and the perception of hunger and satiety *between* different ED types. Given the available evidence, however, we assumed no differences in emotional processing between ED types but differences in the perception of hunger and satiety. We, therefore, hypothesized that individuals with AN would report higher sensibility to satiety than to hunger, and that individuals with BN and BED would report higher sensibility to hunger than to satiety, as well as greater difficulty in distinguishing between physical states.

Study 1

Method

Scale Development and Expert Review

The item pool was generated in several consecutive steps. First, to account for its multifaceted nature, EDIP was defined as the ability to perceive emotions, hunger, and satiety and to discriminate between those physiological states. Based on this definition, a collection of existing items sourced from self-report measures focusing on the perception of bodily sensations related to emotions, hunger, and satiety was compiled. An expert panel (consisting of three clinical psychologists and a psychology student) with extensive experience in ED research or treatment, reviewed the items for quality and relevance to the content domain. Items that adequately assessed either the accuracy in perceiving emotions, hunger, or satiety signals were directly included in the item pool. Items that were only partially applicable were reformulated or removed. In a next step, the item pool was complemented by generating new items to address the facets of EDIP. A 7-point Likert scale was chosen as the response format (1 = *strongly disagree*, 7 = *strongly agree*), with higher scores indicating better EDIP. A pilot study was then conducted in which seven psychology students completed the questionnaire and provided feedback on which items were difficult to understand or misleadingly worded. The expert panel revised or removed difficult items, leaving a final item pool of 76 items. This version was worded in German.

Procedure

German-speaking participants were recruited through online advertisements and flyers distributed at the University of Luxembourg, as well as by sending invitation emails to students and employees. Moreover, student councils of several universities in Germany, Austria, and Switzerland were asked to send the link to the online survey to their students. To recruit a representative sample of individuals with ED, the link was published in various ED community groups. For snowball sampling, participants were asked to forward the link to the survey to their friends and acquaintances and to publish it on social media websites. The introduction to the survey comprised an information sheet and consent form. Participants were requested to confirm that they were at least 18 years of age. The survey itself consisted of a series of sociodemographic questions and a set of eating- and body-related measures, including the item pool to assess EDIP. Completing the survey took approximately 30–40 min. As an incentive for participation, gift vouchers of a total value of 250€ were raffled among participants. The survey was established and assembled using the platform SoSci Survey (https://www.soscisurvey.de/). Ethics approval was sought from the study program of psychology at the University of Luxembourg.

Participants

The online survey was completed by 1,088 participants. The data set was inspected carefully for ensuring data integrity. First, we checked for multiple respondents with the same start and end time for participation, which was not the case. Second, we identified and removed eight participants with high rates of missing or low-item responses. Third, we inspected the response pattern of outliers in response times, none of which were found to be fraudulent.

In addition, seven participants were excluded from statistical data analyses because they met the exclusion criteria (bariatric surgery: $n = 2$, current pregnancy: $n = 1$, being underage: $n = 1$, ED not officially diagnosed: $n = 3$). The remaining sample consisted of 1,073 participants aged between 18 and 75 years, with a mean age of 29.9 years ($SD = 10.8$) for women ($n = 888$) and 31.5 years ($SD = 10.8$) for men. The sample comprised 735 German, 234 Luxembourgish, 59 Austrian, 24 French, and 19 Swiss participants, as well as 76 participants of other nationalities. Four hundred forty-six participants were employed, whereas 479 were university students. The remaining participants stated that they were not in employment ($n = 20$ homeworkers, $n = 26$ pensioners, $n = 27$ jobseekers, $n = 25$ high school students, and $n = 50$ "others"). The body mass index (BMI) was calculated based on self-reported height and body weight (kg/m^2). The mean BMI was 23.5 kg/m^2 ($SD = 6.4$ kg/m^2) in women and 24.0 kg/m^2 ($SD = 3.9$ kg/m^2) in men.

ED diagnoses were assessed via self-report. Only participants who were diagnosed by a mental health professional, for example, clinical psychologist or psychiatrist, were included in the ED subsample. Participants were requested to indicate whether they had a history of EDs, and if so, who established the diagnosis. Furthermore, they were asked to specify whether they had recovered from or whether they were currently affected by their ED. Possible diagnoses were AN, BN, BED, and eating disorders not otherwise specified (EDNOS). As only seven males reported an ED, we decided to discard their data from further analysis. The ED subsample consisted of 212 females, of which 78 individuals reported to have been diagnosed with AN, 53 with BN, 34 with BED, and 34 with EDNOS. A total of 135 participants indicated to have a current ED, of which 43 individuals were diagnosed with AN, 36 with BN, 21 with BED, and 35 with EDNOS. Due to the heterogeneity of the EDNOS group, individuals with EDNOS diagnoses were excluded from statistical analyses.

Measures

Eating Disorder Inventory–2. The German version of the Eating Disorder Inventory–2 (EDI-2; Paul & Thiel, 2005) is a 91-item questionnaire to assess ED psychopathology. In the present study, only the first three subscales, Drive for Thinness (seven items), Bulimia/Bulimic Tendencies (seven items), and Body Dissatisfaction (nine items), were used to assess core ED psychopathology. Items are answered using a 6-point Likert scale ranging from 1 (*never*) to 6 (*always*), with higher scores indicating higher severity of ED symptoms. Cronbach's α (mean interitem correlation [MIIC]) in the current sample was .93 (.65), .91 (.59), and .93 (.61) for the three subscales.

Intuitive Eating Scale–2. The German version of the Intuitive Eating Scale–2 (IES-2; van Dyck et al., 2016) is a 23-item questionnaire to assess intuitive eating, which is characterized by eating in response to physiological hunger and satiety signals rather than for situational or emotional reasons. The IES-2 is composed of four subscales, that is, Unconditional Permission to Eat, Eating for Physical Reasons, Reliance on Hunger/Satiety Cues, and Body–Food Choice Congruence. Responses are given on a 5-point Likert scale from 1 (*strongly disagree*) to 5 (*strongly agree*), with higher scores indicating a higher level of intuitive eating. Cronbach's α (MIIC) in the current sample was .85 (.52),

.86 (.61), .94 (.72) and .79 (.55) for the four subscales, and .93 (.37) for the total score (TS).

Multidimensional Assessment of Interoceptive Awareness. The German version of the MAIA (Bornemann et al., 2015) is a 32-item questionnaire and composed of eight subscales to assess multiple facets of self-reported interoceptive abilities, that is, Noticing, Not-Distracting, Not-Worrying, Attention Regulation, Emotional Awareness, Self-Regulation, Body Listening, and Trusting. Each item is answered on a scale ranging from 1 (*never*) to 5 (*always*), with higher values indicating better interoceptive awareness. Cronbach's α (MIIC) in the current sample ranged from .58 to .92 (.29–.79).

Body Consciousness Questionnaire. In the present study, we used the Private Body Consciousness subscale (PBC; five items) of the Body Consciousness Questionnaire (Miller et al., 1981) to assess one's disposition to focus on internal physical state. Items were translated from English into German and back-translated into English by independent bilingual researchers. The back-translation was compared with the original English version. In case of discrepancies, the item formulations were discussed, and necessary adjustments were made. Items are answered on a 5-point Likert scale ranging from 0 (*extremely uncharacteristic*) to 4 (*extremely characteristic*), with higher scores indicating a stronger tendency to focus on internal bodily sensations. Cronbach's α (MIIC) in the current sample was .60 (.24).

Satisfaction With Life Scale. The German version of the Satisfaction with Life Scale (Glaesmer et al., 2011) is a five-item questionnaire to assess overall life satisfaction. Cronbach's α (MIIC) in the current sample was .92 (.71).

Interpersonal Reactivity Index. The German version of the interpersonal reactivity index (Paulus, 2009) is a self-report measure of affective and perspective empathetic tendencies. In the present study, we used the subscales Empathic Concern (EC; four items) and Perspective Taking (PT; four items) to assess the respondents' tendency to experience feelings of concern or compassion for others and their ability to adopt another person's perspective or point of view. Items are rated on a 5-point Likert scale from 1 (*does not apply at all*) to 5 (*applies very well*), with higher scores indicating greater perspective taking and empathic concern. Cronbach's α (MIIC) in the current sample was .74 (.42) for EC and .80 (.52) for PT.

Statistical Analyses

A series of item analyses were performed to determine which items should be included in EFAs. Frequency analyses were conducted to identify and remove items with high rates of missing data or no variance. Corrected item-to-total correlations $r_{i(t-1)}$ were calculated to determine and remove items with weak correlations to the total score ($r_{i(t-1)} < .3$). Item difficulties p_m were calculated, retaining items with an intermediate item difficulty (.2 < p_m < .8). Point-biserial correlations r_{pb} were calculated between items and diagnostic group (1 = individuals with self-reported EDs; 0 = individuals without self-reported EDs), to remove items that did not differentiate between-group modality ($r_{pb} < .2$). All thresholds were set in accordance with recommendations on scale development (Bortz & Döring, 2016; Bühner, 2011). To examine the factorial structure of the EDIP-Q, we proceeded as follows: First, the number of factors to be extracted was determined by visual inspection of the scree plot as well as Horn's parallel analysis (Horn, 1965). Subsequently, a principal component analysis (PCA) was performed for the total sample (N = 1,073) with oblique (promax) rotation

to allow for intercorrelation between the dimensions of EDIP. Item reductions were based on poor factor loadings (<.5) and cross-loadings (>.3). This procedure, from selecting the appropriate number of factors to item reduction, was repeated until the remaining items could be clearly assigned to one component. To examine the similarity of factors between different subsamples, we conducted PCAs with oblique (promax) rotation separately for sex and diagnostic group and calculated Tucker's (1951) congruence coefficients (φ) as index of the similarity between factors, whereas φ = .85–.94 indicate fair similarity, and φ ≥ .95 good similarity (Lorenzo-Seva & ten Berge, 2006). To evaluate internal consistencies, Cronbach's α and MIICs were calculated for the identified EDIP-Q subscales. Given the large sample size, distributions of all variables were examined by visually inspecting stem-and-leaf plots, boxplots, and Q–Q plots, and by checking values of the skewness and kurtosis statistics (Field, 2013). As no obvious signs of deviation from a normal distribution were found, Pearson correlation coefficients were calculated to determine convergent and divergent validity between EDIP-Q and related and unrelated concepts, whereas r = .10 is considered as "small," r = .30 a "medium," and r = .50 "large" (Cohen, 1992). Statistical analyses were performed with IBM SPSS Statistics (Version 27; IBM Corp., Armonk, NY, USA). Critical α-level was set to .05 in all analyses. This study was not preregistered. Materials and analysis code for this study are available by emailing the corresponding author.

Results

Exploratory Factor Analysis

The Kaiser–Meyer–Olkin coefficient yielded a value of .98, and Bartlett's test of sphericity was significant ($p < .001$), both indicating that the data set was adequate for factor analysis (Tabachnik & Fidell, 2007). Parallel analysis and the inspection of the scree plot suggested a four-factor solution. PCA also revealed a four-factor solution, explaining 65.34% of the total variance. Table 1 contains item-factor loadings for each factor obtained from PCA. These results led to the specification of a 25-item instrument with four subscales named Emotions (seven items assessing the perception of emotions), Hunger (five items assessing the perception of hunger signals), Satiety (seven items assessing the perception of satiety signals), and Discrimination (six items assessing the ability to discriminate between hunger and emotional states). The subscale scores are computed by taking the average of the respective item scores, thus ranging from 1 to 7, with lower scores indicating greater difficulty in EDIP.

Factor Structure Congruence

For the subsamples of men, women, and participants with and without a self-reported ED diagnosis, the Kaiser–Meyer–Olkin coefficients ranged from .87 to .96, and Bartlett's tests for sphericity were significant ($p < .001$). The data sets were, therefore, adequate for factor analysis (Tabachnik & Fidell, 2007). Item factor loadings for Tucker's congruence coefficients (φ) between the two sexes and between participants with and without self-reported ED diagnoses are summarized in the Supplemental Tables S1 and S2. Congruence coefficients were in a high range (.95 ≤ φ ≤ .98), indicating a high degree of factor similarity.

Table 1

Item Factor Loadings and Item Parameters of the EDIP-Q (n = 1,073)

EDIP-Q item	Factor loading				Item parameter		
	1	2	3	4	$r_{i(t-1)}$	p_m	r_{pb}
Factor 1: Satiety							
Item 16	**.88**	−.05	−.02	−.02	.63	.59	.28**
Item 31	**.86**	.08	−.05	.02	.72	.59	.41**
Item 65	**.85**	.02	.07	−.04	.73	.63	.38**
Item 36^R	**.81**	−.10	.20	−.08	.65	.61	.28**
Item 37	**.75**	.12	−.15	.05	.61	.51	.33**
Item 43	**.75**	.12	−.18	.04	.58	.56	.24**
Item 24^R	**.70**	−.12	.26	.10	.72	.69	.38**
Factor 2: Emotions							
Item 49	.02	**.86**	.00	−.02	.70	.65	.38**
Item 58	.02	**.85**	.00	.02	.73	.72	.44**
Item 76	−.05	**.84**	.01	−.02	.63	.67	.38**
Item 67	−.05	**.80**	.17	−.03	.71	.76	.40**
Item 10	.05	**.78**	.04	.00	.70	.71	.41**
Item 33	.00	**.76**	.00	.05	.64	.80	.38**
Item 30	.13	**.64**	−.06	.05	.61	.62	.35**
Factor 3: Discrimination							
Item 42^R	−.06	.01	**.84**	.03	.61	.84	.31**
Item 23^R	−06	.04	**.81**	.81	.67	.80	.38**
Item 53^R	.04	.12	**.78**	−.08	.66	.80	.37**
Item 61^R	.08	−.03	**.76**	−.02	.58	.77	.27**
Item 21^R	.09	.03	**.76**	−.02	.64	.74	.34**
Item 15^R	−.18	−.01	**.71**	.10	.43	.85	.27**
Factor 4: Hunger							
Item 32^R	−.08	−.10	.06	**.84**	.41	.71	.32**
Item 01	.04	.03	−.09	**.83**	.49	.71	.35**
Item 08^R	−.05	.01	.22	**.71**	.58	.79	.38**
Item 70^R	.02	.09	.03	**.68**	.48	.66	.37**
Item 74	.17	.03	−.13	**.63**	.43	.67	.24**

Note. EDIP-Q = Eating Disorder-Specific Interoceptive Perception Questionnaire; extraction method: principal component analysis; rotation method: Promax with Kaiser normalization; factor loadings above .30 are in bold; $r_{i(t-1)}$ = corrected item-total correlations; p_m = item difficulty; r_{pb} = point-biserial correlations.
^{R} Reversed coded items.
** $p < .001$.

Item Analyses

Corrected item-total correlation, item difficulty, and point-biserial correlation for each item are summarized in Table 1. Corrected item-total correlations were in the medium-to-high range. Item difficulties were in the medium range. Point-biserial correlations between the items and group (1 = ED group, 0 = control group) were in the medium range, indicating that items were moderately related to the group modality.

Internal Consistencies

Internal consistencies of the EDIP-Q subscales were high, with α (MIIC) = .91 (.61) for Emotions, .81 (.47) for Hunger, .92 (.61) for Satiety, and .89 (.56) for Discrimination.

Construct Validity

Correlations between the EDIP-Q subscales and other variables are presented in Table 2. Regarding convergent validity,

intercorrelations of the EDIP-Q subscales were in the medium-to-high range. EDIP-Q subscales were moderately to strongly inversely related to the EDI-2 subscales and moderately to strongly related to IES-2 total score and subscales, except for a small correlation between EDIP-Q Hunger and Eating for Physical Reasons. Moderate-to-strong relationships were found between EDIP-Q subscales and the MAIA mean score and subscales Noticing, Attention Regulation, Emotional Awareness, Self-Regulation, Body Listening, and Trusting, and weak-to-moderate relationships for Not-Distracting and Not-Worrying. In addition, moderate relationships were found between the EDIP-Q subscales and PBC and SLWS. Regarding divergent validity, weak correlations were found between the EDIP-Q subscales and EC and PT.

Study 2

Method

Procedure

The procedure was the same as in Study 1.

Participants

The online survey was completed by 976 participants. To ensure data integrity, we followed the same procedure as in Study 1. A total of 16 participants were excluded from statistical analyses (high rates of missing or low-item responses: $n = 2$; extreme response style: $n = 3$; bariatric surgery: $n = 3$; current pregnancy: $n = 1$; ED not officially diagnosed: $n = 7$). The remaining sample, therefore, consists of 958 participants aged between 18 and 76 years. Women ($n = 838$) had a mean age of 29.0 years ($SD = 10.8$) and mean BMI of 23.9 kg/m^2 ($SD = 7.2$ kg/m^2), whereas men had a mean age of 32.8 years ($SD = 13.7$) and mean BMI of 25.2 kg/m^2 ($SD = 6.9$ kg/m^2). The sample included 692 German, 192 Luxembourgish, 32 Swiss, 27 Austrian, and 15 French participants, whereas 65 participants indicated another nationality. Three hundred fifty-one participants were currently working, and 343 were university students. The remaining participants stated that they were not in employment ($n = 25$ homeworkers, $n = 34$ pensioners, $n = 29$ jobseekers, $n = 63$ trainees, $n = 43$ high school students, and $n = 70$ "others").

The ED subsample was established in the same way as in Study 1 and consisted of 256 participants. As only seven males were diagnosed with an ED, we decided to discard their data from further analysis. The ED subsample finally consisted of 249 female participants, of which 156 individuals reported to have been diagnosed with AN, 116 with BN, 45 with BED, and 41 with EDNOS during their lifetime (multiple answers were possible). A total of 175 participants indicated to have a current ED (70.28% of the ED subsample), of which 49 individuals were diagnosed with AN, 47 with BN, 21 with BED, and 58 with EDNOS. As in Study 1, individuals with EDNOS diagnoses were excluded from statistical analyses for reasons of heterogeneity.

Measures

Eating Disorder-Specific Interoceptive Perception–Questionnaire. The 25-item version of the EDIP-Q derived from Study 1 to assess EDIP consists of four subscales: Emotions (seven items), Hunger (five items), Satiety (seven items), and

Table 2
Intercorrelations of Measured Variables (n = 1,073)

Measures	Convergent validity				Discriminant validity			
	S	E	D	H	S	E	D	H
EDIP-Q								
Satiety								
Emotions	.62**							
Discrimination	.56**	.59**						
Hunger	.40**	.53**	.42**					
EDI-2								
DT	−.55**	−.53**	−.50**	−.42**				
BT	−.69**	−.56**	−.63**	−.35**				
BD	−.53**	−.50**	−.44**	−.36**				
IES-2								
Mean score	.75**	.63**	.62**	.50**				
UPE	.40**	.44**	.41**	.41**				
EPR	.60**	.44**	.52**	.16**				
RHSC	.75**	.64**	.56**	.50**				
B-FCC	.34**	.36**	.25**	.34**				
MAIA								
Mean score	.56**	.73**	.45**	.46**				
Noticing	.40**	.52**	.29**	.32**				
Not-Distracting	.27**	.34**	.24**	.30**				
Not-Worrying	.21**	.24**	.22**	.09**				
AR	.44**	.57**	.37**	.31**				
EA	.34**	.49**	.24**	.32**				
Self-Regulation	.40**	.51**	.27**	.32**				
Body Listening	.41**	.54**	.30**	.35**				
Trusting	.57**	.71**	.50**	.49**				
BCQ-PBC	.30**	.23**	.32**	.30**				
SWLS	.49**	.36**	.37**	.49**				
IRI								
EC					.01	.08**	−.04	.02
PT					.13**	.17**	.07**	.03

Note. EDIP-Q = Eating Disorder-Specific Interoceptive Perception Questionnaire; EDI-2 = Eating Disorder Inventory–2; DT = Drive for Thinness; BT = Bulimic Tendencies; BD = Body Dissatisfaction; IES-2 = Intuitive Eating Scale–2; UPE = Unconditional Permission to Eat; EPR = Eating for Physical Rather Than Emotional Reasons; RHSC = Reliance on Hunger and Satiety Cues, B-FCC = Body–Food Choice Congruence; MAIA = Multidimensional Assessment of Interoceptive Awareness; AR = Attention Regulation; EA = Emotional Awareness; BCQ-PBC = Private Body Consciousness subscale of the Body Consciousness Questionnaire; SWLS = Satisfaction With Life Scale; IRI = Interpersonal Reactivity Index; EC = Empathetic Concern; PT = Perspective Taking.
$**p < .01$.

Discrimination (six items). A 7-point Likert scale is used to assess the degree of agreement, whereby lower scores indicate greater difficulties in EDIP.

Eating Disorder Examination–Questionnaire. The German version of the Eating Disorder Examination–Questionnaire (Hilbert & Tuschen-Caffier, 2007) is a 28-item questionnaire for the assessment of ED psychopathology. In the present study, only six items were used to assess diagnostically relevant core ED behaviors, that is, the frequency of overeating episodes, loss of control over eating, binge eating, self-induced vomiting, laxative abuse, and excessive exercise. Participants were asked to indicate how often these behaviors occurred within the past 28 days.

Dutch Eating Behavior Questionnaire. The German version of the Dutch Eating Behavior Questionnaire (Nagl et al., 2016) is a 33-item questionnaire to assess three distinct maladaptive eating behaviors, that is, emotional eating, external eating, and restrained eating. Each item is answered on a scale ranging from 1 (*never*) to 5 (*very often*), with higher scores indicating more frequent

endorsement of maladaptive eating behavior. Cronbach's α (MIIC) in the current sample was .96 (.65), .77 (.25), and .93 (.56) for the three subscales.

Multidimensional Assessment of Interoceptive Awareness. The German version of the MAIA (Bornemann et al., 2015) is described in detail in the method section of Study 1. Cronbach's α (MIIC) in the current sample ranged from .62 to .93 (.31–.82).

Emotional Competency Questionnaire. The Emotional Competency Questionnaire (German: Emotionale-Kompetenz-Fragebogen, EKF; Rindermann, 2009) is a German 62-item questionnaire for the assessment of emotional competencies. In the present study, only the subscale Perception and Acknowledgement of one's own Emotions (EKF-Emotion; 15 items) was used to assess the ability to perceive and understand one's own emotions. Items were answered on a scale ranging from 1 (*not at all true*) to 5 (*definitely true*), with higher scores indicating a better ability to perceive and recognize emotions. Cronbach's α (MIIC) in the current sample was .94 (.51).

Three-Factor Eating Questionnaire. The German version of the Three-Factor Eating Questionnaire (TFEQ; Pudel & Westenhöfer, 1989) consists of 51 items to measure three dimensions of eating behavior, that is, Cognitive Restraint of Eating, Disinhibition, and Hunger. In the present study, only the subscale Hunger (TFEQ-Hunger; 14 items) was used to measure experienced feelings of hunger, often perceived as disturbing. Higher sum scores indicate a stronger endorsement of hunger feelings. Cronbach's α (MIIC) in the current sample was .83 (.47).

Mindful Eating Questionnaire. The Mindful Eating Questionnaire (MEQ; Framson et al., 2009) is a 28-item questionnaire to measure mindful eating, which is defined as "nonjudgmental awareness of physical and emotional sensations while eating or in a food-related environment." In the present study, only the subscale Disinhibition (eight items) was used, which assesses the inability to stop eating despite being full, with lower scores indicating greater inability. The items were translated from English into German and back-translated into English by independent bilingual researchers. The back-translation was compared with the original English version. In case of discrepancies, the item formulations were discussed, and necessary adjustments were made. Cronbach's α (MIIC) in the current sample was .88 (.47).

Statistical Analyses

To test the four-factor structure of the EDIP-Q established in Study 1, we conducted CFA with maximum likelihood estimation and the covariance matrix as input using MPLUS (Version 8.7). Goodness-of-fit for the models was evaluated with the χ^2-statistic. As the χ^2-test is sensitive to sample size and usually significant in samples with more than 200 cases (Browne & Cudeck, 1993), we also used the comparative fit index (CFI) and normed fit index, the root-mean-square error of approximation (RMSEA), and the standardized root-mean-square residual (SRMR) as guides in assessing fit. Levels of 0.90 or higher for CFI and normed fit index and 0.08 or lower for RMSEA and SRMR indicate a good fit of the model to the data (Bentler & Bonett, 1980; Browne & Cudeck, 1993; Hu & Bentler, 1999). Following the same procedure as in Study 1, we visually inspected the distribution of each variable. As no obvious signs of deviation from a normal distribution were found, Pearson's rs were calculated to determine convergent and divergent validity between EDIP-Q and related and unrelated concepts according to Cohen's (1992) criteria. This study was not preregistered. Materials and analysis code for this study are available by emailing the corresponding author.

Results

Confirmatory Factor Analysis

In the first step, we established a first-order model (based on the results from Study 1) and examined the magnitude and pattern of correlations among the hypothesized factors in the first-order solution. The EDIP-Q items were specified to load on their respective latent first-order factor. Relationships between the four latent factors were estimated. The four-factor solution provided a good fit for the data, CFI = 0.95, TLI = 0.94, SRMR = 0.047, RMSEA = 0.052, 90% CI [0.049, 0.056], $\chi^2(1, 269) = 978.122$, $p < .001$. Factor loadings of all items were high (>.50) and the first-order factors were significantly

interrelated (rs range from .45 to .64, $p < .001$), indicating the viability of a postulated second-order model (Brown, 2015).

In a second step, we specified, therefore, a second-order model, in which the first-order factors loaded on a higher order EDIP factor. This model provided a good fit for the data, CFI = 0.95, TLI = 0.94, SRMR = 0.050, RMSEA = 0.053, 90% CI [0.050, 0.057], $\chi^2(1, 271) = 922.17$, $p < .001$. Both the item-factor loadings and the first-order factor loadings on the higher order EDIP factor were high (≥0.50; Backhaus et al., 2015). Table 3 includes the item-factor loadings as well as the first-order factor loadings on the higher order EDIP factor for the overall sample.

Internal Consistencies

Internal consistencies for the EDIP-Q were high, with α (MIIC) = .94 (.40) for the TS, .92 (.65) for Emotions, .86 (.53) for Hunger, .93 (.64) for Satiety, and .88 (.55) for Discrimination.

Construct Validity

Correlations between the EDIP-Q TS and subscales and other variables are presented in Table 4. Regarding convergent validity, the intercorrelations of the EDIP-Q TS and subscales were in the medium-to-high range. The EDIP-Q (TS and subscales) was inversely

Table 3

Standardized Item and Factor Loadings of the EDIP Questionnaire Obtained by Confirmatory Factor Analyses of the Overall Sample (n = 958)

Factor and EDIP-Q item	First order	Second order
Factor 1: Satiety		.80
Item 16	.79	
Item 31	.86	
Item 65	.82	
Item 36[R]	.82	
Item 37	.76	
Item 43	.61	
Item 24[R]	.80	
Factor 2: Emotions		.80
Item 49	.79	
Item 58	.83	
Item 76	.72	
Item 67	.77	
Item 10	.85	
Item 33	.68	
Item 30	.71	
Factor 3: Discrimination		.76
Item 42[R]	.74	
Item 23[R]	.82	
Item 53[R]	.82	
Item 61[R]	.71	
Item 21[R]	.72	
Item 15[R]	.56	
Factor 4: Hunger		.63
Item 32[R]	.66	
Item 01	.78	
Item 08[R]	.85	
Item 70[R]	.61	
Item 74	.60	

Note. EDIP-Q = Eating Disorder-Specific Interoceptive Perception Questionnaire.
[R] Reversed coded items.

Table 4
Intercorrelations of Measured Variables (n = 958)

Measures	Convergent validity				
	TS	S	E	D	H
EDIP-Q					
Total score					
Satiety	.84**				
Emotions	.84**	.60**			
Discrimination	.78**	.39**	.52**		
Hunger	.68**	.37**	.50**	.39**	
EDE-Q					
Overeating	−.06	−.03	−.04	−.07**	−.07**
Loss of control	−.46**	−.47**	−.35**	−.39**	−.20**
Binge eating	−.50**	−.51**	−.35**	−.44**	−.23**
Self-induced vomiting	−.32**	−.25**	−.28**	−.23**	−.26**
Laxative/diuretics abuse	−.18**	−.12**	−.18**	−.11**	−.17**
Excessive exercise	−.27**	−.20**	−.20**	−.16**	−.31**
DEBQ					
Restrained eating	−.49**	−.39**	−.40**	−.35**	−.43**
Emotional eating	−.56**	−.59**	−.39**	−.60**	−.15**
External eating	−.13	−.28**	−.04	−.18**	−.17**
BMI	−.13**	−.19**	−.04	−.16	.02
MAIA					
Mean score	.70**	.52**	.76**	.42**	.45**
Noticing	.50**	.40**	.58**	.28**	.30**
Not-Distracting	.44**	.32**	.44**	.26**	.37**
Not-Worrying	.21**	.19**	.20**	.19**	.09**
AR	.59**	.45**	.63**	.39**	.36**
EA	.39**	.30**	.49**	.15**	.28**
Self-regulation	.50**	.41**	.54**	.28**	.31**
Body listening	.51**	.42**	.55**	.27**	.32**
Trusting	.70**	.55**	.76**	.42**	.45**
EKF-Emotion	.67**	.49**	.76**	.48**	.41**
TFEQ-Hunger	−.41**	−.58**	−.26**	−.42 **	−.08**
MEQ-Disinhibition	−.02	−.62**	−.24**	−.35**	−.02

Note. EDIP-Q = Eating Disorder-Specific Interoceptive Perception Questionnaire; EDE-Q = Eating Disorder Examination–Questionnaire; DEBQ = Dutch Eating Behavior Questionnaire; BMI = body mass index; MAIA = Multidimensional Assessment of Interoceptive Awareness; AR = Attention Regulation; EA = Emotional Awareness; EKF = Emotionale Kompetenz Fragebogen; TFEQ = Three-Factor Eating Questionnaire; MEQ = Mindful Eating Questionnaire.
** $p < .01$.

related to core ED behaviors as assessed by the Eating Disorder Examination–Questionnaire. Specifically, moderate-to-strong negative relationships were found between the EDIP-Q (TS and subscales) and loss of control and binge eating. Weak-to-moderate negative relationships were found between the EDIP-Q (TS and subscales) and self-induced vomiting, laxative/diuretics abuse, and excessive exercise. Weak negative relationships were found between the EDIP-Q (TS and subscales) and overeating. The EDIP-Q (TS and subscales) was negatively related to Dutch Eating Behavior Questionnaire subscales. Specifically, moderate-to-large correlations were found for restrained and emotional eating, except for a small relationship between EDIP-Q hunger and emotional eating. Weak-to-moderate relationships were found between the EDIP-Q (TS and subscales) and external eating. There was only a weak correlation between the EDIP-Q (TS and subscales) and BMI. The correlations between the EDIP-Q subscales and the MAIA subscales were comparable to

those obtained in Study 1. Moderate-to-large correlations were identified between the EDIP-Q TS and MAIA mean score and subscales.

To further address the convergent validity of the EDIP-Q Emotions, Hunger, and Satiety, we examined their relationship to subscales designed to measure the same construct. EDIP-Q Emotions was strongly related to EKF-Emotion. EDIP-Q Satiety was significantly inversely related to MEQ-Disinhibition. EDIP-Q Hunger and TFEQ Hunger were weakly negatively correlated. In addition, the EDIP-Q subscales correlated with subscales measuring different constructs. Specifically, we found small-to-medium negative relationships between EDIP-Q Emotions and TFEQ-Hunger and MEQ-Disinhibition. EDIP-Q Hunger was moderately related to EKF-Emotion and weakly related to MEQ-Disinhibition. EDIP-Q Satiety was largely correlated with EKF-Emotion, and largely inversely with TFEQ-Hunger.

Studies 1 and 2: Group Comparisons

Method

Procedure and Participants

The procedure and samples are described in the methods sections of Studies 1 and 2.

Statistical Analyses

Sex differences in EDIP were examined using multivariate analyses of variance (MANOVA), with EDIP-Q TS and subscales as dependent variables and Sex as between-group factor. Compared to HC, differences between participants with self-reported current and past ED diagnoses and between different types of EDs were examined using MANOVA, with EDIP-Q TS and subscales as dependent variables and Diagnostic Status as between-subjects factor. Only the female subgroup of the nonclinical sample was included in the analysis because the ED subgroup included female participants only. Wilks' lamba (Λ) was used as a test statistic for F tests in all MANOVA models because it is more robust with unequal sample sizes and variances (Ateş et al., 2019). When F tests were significant for the comparison of more than two groups, post hoc analyses were conducted using the Games–Howell multiple comparison procedure for unequal sample sizes and variances (Games & Howell, 1976). Partial η squared (η_p^2) was calculated for estimation of effect sizes, where η_p^2 = .01 is considered as "small," η_p^2 = .06 "medium," and η_p^2 = .14 "large" (Cohen, 1988). Statistical analyses were performed with IBM SPSS Statistics (Version 27; IBM Corp., Armonk, NY, USA). Critical α level was set to .05 in all analyses.

Results

Sex Differences

EDIP-Q mean scores and *SD*s for men and women are presented in Table 5. In Sample 1, the MANOVA with EDIP-Q TS and subscales as dependent variables and Sex as between-subject factor revealed higher EDIP scores in men than in women, Λ = .98, $F(5, 1067) = 3.625$, $p < .01$, η_p^2 = .02. The same effect was found in Sample 2, Λ = .97, $F(5, 952) = 5.236$, $p < .01$, η_p^2 = .03.

Table 5
EDIP Mean Scores (M) and Standard Deviations (SD) for Men and Women in Studies 1 and 2

| | Study 1 | | | | | | | Study 2 | | | | | | |
| | Men (n = 185) | | Women (n = 888) | | | | | Men (n = 120) | | Women (n = 838) | | | | |
EDIP-Q	M	SD	M	SD	F(1, 1071)	p	η_p^2	M	SD	M	SD	F(1, 956)	p	η_p^2
TS	5.5	0.8	5.1	1.1	19.17	<.001	.02	5.4	0.8	4.9	1.2	21.75	<.001	.02
E	5.5	0.9	5.2	1.3	13.46	.003	.01	5.5	1.0	4.8	1.4	22.23	<.001	.02
H	5.5	1.1	5.2	1.4	10.27	.015	.01	5.3	1.3	4.9	1.5	7.68	.006	.01
S	5.0	1.3	4.5	1.6	31.36	<.001	.01	4.9	1.3	4.3	1.6	15.65	<.001	.02
D	6.1	1.0	5.7	1.4	23.06	<.001	.01	5.9	1.3	5.5	1.4	8.53	.004	.01

Note. Group differences were calculated using MANOVA. EDIP-Q = Eating Disorder-specific Interoceptive Perception Questionnaire; TS = total score; E = Emotion subscale; H = Hunger subscale; S = Satiety subscale; D = Discrimination subscale; MANOVA = multivariate analyses of variance.

Diagnostic Status I: HC Versus Self-Reported Past and Current ED

EDIP-Q mean scores and *SD*s for female healthy participants and participants with self-reported current and past ED diagnoses are presented in Table 6. In Sample 1, the MANOVA with EDIP-Q TS and subscales as dependent variables and Diagnostic Status I as between-subjects factor revealed an overall effect for Diagnostic Status I, $\Lambda = .60$, $F(10, 1762) = 51.47$, $p < .001$, $\eta_p^2 = .23$. The same effect was found in Sample 2, $\Lambda = .71$, $F(10, 1662) = 31.48$, $p < .001$, $\eta_p^2 = .16$. Group differences on the TS and all subscales were significant ($p < .05$), except for a marginally significant group difference ($p = .08$) on Emotions between HC and participants with self-reported past ED in Sample 2. Overall, HC had the highest scores, followed by participants with self-reported past and current EDs.

Diagnostic Status II: HC Versus Types of EDs

EDIP-Q mean scores and *SD*s for female healthy participants and participants reporting a current diagnosis of AN, BN, and BED are presented in Table 7. In Sample 1, the MANOVA with EDIP-Q TS and subscales as dependent variables and Diagnostic Status II as between-subject factor revealed an overall effect for Diagnostic Status II, $\Lambda = .57$, $F(15, 2120) = 31.81$, $p < .001$, $\eta_p^2 = .17$. Post hoc tests showed that all types of EDs scored significantly lower on the TS and subscale scores compared to HC. Among types of EDs, there were no significant group differences in the TS and Emotions. Individuals with BN and BED, however, scored higher on Hunger ($p = .044$ and $p = .017$, resp.) and lower on Satiety ($p = .05$ and $p = .09$, resp.) than participants with AN. Furthermore, participants with BN and BED scored lower on Discrimination compared to participants with AN ($p = .032$ and $p = .033$, resp.). No other post hoc comparisons reached statistical significance.

In Sample 2, the MANOVA with EDIP-Q TS and subscales as dependent variables and Diagnostic Status II as between-subjects factor showed an overall effect for Diagnostic Status II, $\Lambda = .71$, $F(15, 2131) = 19.18$, $p < .001$, $\eta_p^2 = .11$. Post hoc tests revealed that all types of EDs scored significantly lower on the TS and subscale scores compared to HC. Post hoc tests between ED subgroups yielded significant group differences on the TS and the subscales Satiety and Discrimination. In detail, individuals with AN had a higher TS compared to participants with BN ($p = .019$) and BED ($p = .051$). Participants with BN and BED scored significantly lower on Satiety ($p < .001$ and $p = .045$, resp.) and Discrimination ($p < .001$) compared to participants with AN. No other post hoc comparisons reached statistical significance.

Table 6
EDIP Mean Scores (M) and Standard Deviations (SD) for Female Healthy Participants and Participants With Past and Current Self-Reported Eating Disorder for Studies 1 and 2

| | Study 1 | | | | | | | | | Study 2 | | | | | | | | |
| | HC (n = 676) | | Past ED (n = 77) | | Current ED (n = 135) | | | | | HC (n = 589) | | Past ED (n = 74) | | Current ED (n = 175) | | | | |
EDIP-Q	M	SD	M	SD	M	SD	F(2, 885)	p	η_p^2	M	SD	M	SD	M	SD	F(2, 835)	p	η_p^2
TS	5.5ₐ	0.9	5.0_b	1.1	3.5_c	1.0	275.16	<.001	.38	5.2ₐ	1.1	4.8_b	1.1	3.7_c	0.9	143.29	<.001	.26
E	5.5ₐ	1.0	5.02_b	1.3	3.50_c	1.3	202.74	<.001	.31	5.2ₐ	1.3	4.9ₐ	1.4	3.5_b	1.2	122.06	<.001	.23
H	5.6ₐ	1.1	4.99_b	1.3	3.6_c	1.5	154.27	<.001	.26	5.3ₐ	1.3	4.6_b	1.5	3.7_c	1.5	96.85	<.001	.19
S	4.9ₐ	1.4	4.35_b	1.7	2.8_c	1.3	126.16	<.01	.22	4.6ₐ	1.5	4.1_b	1.5	3.2_c	1.4	59.97	<.001	.13
D	6.1ₐ	1.1	5.61_b	1.4	4.2_c	1.7	134.94	<.01	.23	5.8ₐ	1.3	5.6ₐ	1.3	4.5_b	1.5	62.03	<.001	.13

Note. Group differences were calculated using MANOVA, means with different subscripts differ at the $p = .05$ level. HC = healthy control participants; past ED = participants with past self-reported eating disorder; current ED = participants with current self-reported eating disorder; EDIP-Q = Eating Disorder-Specific Interoceptive Perception Questionnaire; TS = total score; E = Emotion subscale; H = Hunger subscale; S = Satiety subscale; D = Discrimination subscale; MANOVA = multivariate analyses of variance.

Table 7

EDIP Mean Scores (M) and Standard Deviations (SD) for Healthy Control Participants and Participants With Current Self-Reported AN, BN, and BED for Studies 1 and 2

	Study 1																	Study 2																
	HC (n = 676)		AN (n = 43)		BN (n = 36)		BED (n = 21)					HC (n = 589)		AN (n = 49)		BN (n = 47)		BED (n = 21)				
EDIP-Q	M	SD	M	SD	M	SD	M	SD	$F(3, 776)$	p	η_p^2	M	SD	M	SD	M	SD	M	SD	$F(3, 776)$	p	η_p^2
TS	5.5$_a$	0.9	3.6$_b$	1.0	3.3$_b$	1.0	3.5$_b$	0.9	148.87	<.001	.37	5.2$_a$	1.1	4.0$_b$	.9	3.3$_c$	0.7	3.3$_{bc}$	1.0	78.90	<.001	.23
E	5.2$_a$	1.0	3.4$_b$	1.4	3.4$_b$	1.3	3.8$_b$	1.0	111.98	<.001	.30	5.2$_a$	1.3	3.5$_b$	1.2	3.4$_b$	1.1	3.0$_b$	1.1	66.58	<.001	.21
H	5.6$_a$	1.1	3.1$_c$	1.3	4.0$_b$	1.5	4.4$_b$	1.6	83.77	<.001	.25	5.2$_a$	1.3	3.5$_b$	1.4	3.8$_b$	1.4	4.2$_b$	1.5	39.21	<.001	.13
S	4.9$_a$	1.4	3.3$_b$	1.5	2.5$_b$	1.2	2.5$_b$	1.0	70.24	<.001	.21	4.6$_a$	1.5	3.7$_b$	1.3	2.5$_b$	1.0	2.7$_c$	1.4	39.98	<.001	.13
D	6.1$_a$	1.1	4.8$_b$	1.5	3.8$_c$	1.7	3.6$_c$	1.5	83.37	<.001	.25	5.8$_a$	1.3	5.1$_b$	1.5	3.8$_c$	1.3	3.6$_c$	1.2	51.87	<.001	.17

Note. Group differences were calculated using MANOVA, means with different subscripts differ at the $p = .05$ level. HC = healthy control; AN = anorexia nervosa; BN = bulimia nervosa; BED = binge-eating disorder; EDIP-Q = Eating Disorder-Specific Interoceptive Perception Questionnaire; TS = total score; E = Emotion subscale; H = Hunger subscale; S = Satiety subscale; D = Discrimination subscale; MANOVA = multivariate analyses of variance.

General Discussion

Although the literature points to a differential contribution of emotions, hunger, and satiety signals to ED symptoms (e.g., Poovey et al., 2022; van Dyck et al., 2016), there is a lack of appropriate measures that distinguish between the perception of these visceral signals. The first objective of the present studies, therefore, concerned the development and initial validation of the EDIP-Q, a multifaceted assessment of EDIP. The second objective was to investigate differences in EDIP using the EDIP-Q, including sex differences and differences between participants with a self-reported current and past ED diagnosis (Diagnostic Status I), as well as between different types of EDs (Diagnostic Status II), compared with healthy participants.

Factor Structure

In Study 1, the factor structure of the original scale was examined using EFA in a large sample of individuals with and without self-reported EDs, and in different subsamples. With a high degree of factor similarity between the different subsamples, our results led to the specification of a 25-item instrument, with four factors forming the subscales Emotions (assessment of emotion perception), Hunger (assessment of the perception of hunger signals), Satiety (assessment of the perception of satiety signals), and Discrimination (assessment of the ability to discriminate between hunger and emotional states). In Study 2, the hypothesized four-factor structure was confirmed using CFA in an equally large sample of individuals with and without self-reported EDs. The moderate-to-strong intercorrelations of the first-order factors (i.e., Emotions, Hunger, Satiety, and Discrimination) indicated the presence of a higher order EDIP factor. Hence, we tested a second-order model, in which the first-order factors loaded on a higher order EDIP factor. The second-order model also provided a good fit for the data. These results provide initial empirical evidence that the first-order factors are independent of each other but represent important components of EDIP. Accordingly, EDIP-Q subscale scores and the TS should be calculated, the latter reflecting the general ability of EDIP.

Construct Validity

Regarding convergent validity, the EDIP-Q TS and subscales were inversely related to self-reports of ED symptoms and different maladaptive eating styles, including nonintuitive, emotional, external, and restrictive eating. The correlation coefficients fell predominantly within the moderate-to-high range, except for the small correlation between EDIP-Q Hunger and IES-Eating for Physical Reasons. This can be ascribed to the fact that the latter measures the individual tendency to eat due to physical hunger rather than to cope with emotional distress. Some items refer to emotional eating as a coping strategy (e.g., "I find myself eating when I'm feeling emotional [e.g., anxious, depressed, sad], even when I'm not physically hungry."). The EDIP-Q Hunger items refer, however, exclusively to the perception of physiological hunger, which explains the weak association between the two subscales. Overall, these findings indicate that impaired EDIP is associated with ED symptom severity and maladaptive eating behaviors. Although the BMI has been associated with both maladaptive eating and interoceptive deficits (e.g., Jenkinson et al., 2018; Nagl et al., 2016; Richard et al., 2019), only small correlations were found between the EDIP-Q (TS and subscales) and BMI. This weak association may be explained by other factors affecting the BMI, such as age, sex, metabolism, muscle mass, and bone density. In line with our hypothesis, the EDIP-Q TS and subscale scores were moderately positively related to self-report measures on the perception of internal bodily states (i.e., Body Consciousness Questionnaire, MAIA). This suggests that deficits in EDIP are associated with general difficulties in focusing on physical states. In addition, the correlation coefficients were in a higher range than those found in other studies investigating the relationship between different self-report measures of interoception (e.g., Murphy et al., 2020), which indicates a higher degree of convergence of the EDIP-Q. EDIP-Q subscale scores were moderately related to self-reported overall life satisfaction. In summary, our findings imply that the EDIP-Q is a suitable measure of interoceptive processing in the domain of maladaptive eating behaviors and EDs. Further research is, however, warranted to investigate the relationship between the EDIP-Q and other interoceptive dimensions.

To investigate further aspects of convergent validity of EDIP-Q subscales Emotions, Hunger, and Satiety, we investigated whether these subscales correlated with subscales of other questionnaires designed to measure the same construct. In line with our hypothesis, EDIP-Q Emotions was strongly related to the perception and acknowledgment of one's own emotions (assessed by EKF-Emotion), indicating a high degree of convergence between the two subscales. As expected, EDIP-Q Satiety was strongly inversely related to the inability to stop eating despite being full (assessed by MEQ-Disinhibition), providing evidence that EDIP-Q Satiety is an appropriate measure of the perception of satiety signals. Contrary to our expectations, we found a small correlation between EDIP-Q Hunger and TFEQ-Hunger. Comparing the content of the items, it becomes apparent that TFEQ-Hunger assesses experienced appetite as a response to external stimuli (e.g., "When I see a real delicacy, I often get so hungry that I have to eat right away"), whereas EDIP-Q Hunger was designed to directly assess the perception of physiological hunger signals. Interestingly, we also found small to large correlations between the EDIP-Q subscales Emotion, Hunger, and Satiety with subscales designed to measure another physical state (e.g., moderate relationships between EDIP-Q Hunger and Satiety with EKF-Emotion). These associations can be attributed to the fact that similar bodily processes are involved in the perception of emotions, hunger, and satiety. Together with the strong intercorrelations found between the EDIP-Q subscales, these findings provide additional evidence that the EDIP-Q factors are independent but interrelated.

To address divergent validity, we examined the relationship between the EDIP-Q and self-reports on empathetic concern and perspective taking. As expected, EDIP-Q subscales were unrelated to EC and PT, supporting the divergent validity of the EDIP-Q.

Group Comparisons

In both study samples, women had significantly lower EDIP-Q scores than men, indicating a sex-specific trend toward difficulties in perceiving and discriminating between emotions, hunger, and satiety signals. Given the higher prevalence of EDs among females (Galmiche et al., 2019), it is reasonable to assume that this sex-specific trend represents a possible risk factor for the development of ED symptoms in women.

We examined group differences in EDIP based on diagnostic status. In both study samples, we observed the lowest EDIP-Q scores in female participants with self-reported current ED diagnosis, followed by participants with self-reported past ED diagnosis and HCs. These results suggest that EDIP is especially impaired in individuals currently experiencing ED symptoms and improves during recovery, although these are not comparable to scores of participants without a self-reported ED history. Together with the high correlations found between the EDIP-Q and self-reports of ED severity, these findings support etiological models, in which impaired interoceptive processing plays a crucial role in the development and maintenance of EDs (e.g., Martin et al., 2019).

To determine whether EDIP plays a differential role in the development and maintenance of different EDs, we examined differences between women with self-reported AN, BN, and BED compared to HC. As expected, individuals with EDs had significantly lower total and subscale scores than HC, indicating an overall impaired ability of EDIP in EDs. As expected, individuals with AN, BN, and BED have equal difficulty in perceiving emotions

and differ in their perception of hunger and satiety signals and in their ability to distinguish between emotional states and hunger. Compared to AN, female participants with BN or BED reported higher sensibility to hunger, lower sensibility to satiety, and greater difficulty distinguishing between emotional states and feelings of hunger. These results are consistent with previous findings that show binge eating in response to unpleasant emotional states in individuals with BN and BED (e.g., Meule et al., 2021; Schaefer et al., 2020) and reflect the clinical picture of BN and BED, which is characterized by recurrent episodes of binge eating, that is, eating large amounts of food in a short period of time (APA, 2013). The opposite picture was observed in individuals with AN, who reported greater difficulty perceiving hunger signals but lesser difficulty perceiving satiety signals and distinguishing between emotional states and hunger compared with BN and BED.

Although profiles of deficits in EDIP were evident in both study samples, not all differences reached statistical significance. A possible explanation could be that ED diagnoses were self-reported, resulting in a certain degree of heterogeneity within groups and, thus, less meaningful group differences in EDIP. Although only participants who indicated being diagnosed by a qualified mental health professional were included in the clinical subsample, it was not possible to verify the validity of the diagnoses and to determine whether the participants fulfilled the diagnostic criteria. Future research should, therefore, consider using the EDIP-Q in a controlled clinical setting and examine group differences in individuals diagnosed with a standardized clinical interview. Notwithstanding the former, together with EFA and CFA results, these differences between ED types point to the need to differentiate the components of EDIP and to develop more targeted treatment approaches for EDs.

Limitations

Selection and response bias are common in web-based data collection. Both samples consisted primarily of female participants (82.80% in Study 1 and 87.47% in Study 2), since women are more interested in, and seek treatment for, EDs, which may have led to a greater willingness to participate. Previous research on EDs using a similar study design observed a similar gender ratio in their sample population (e.g., van Dyck et al., 2016). Although the risk of response bias in the data cannot be excluded, a number of studies have shown that the study format does not affect psychometric properties of self-report measures and leads to comparable questionnaire results (e.g., Mayr et al., 2012). Moreover, the use of anonymous web-based surveys can be even more beneficial for collecting unbiased data on sensitive issues, such as eating- and body-related information (Kays et al., 2012). Cronbach's α coefficients for some subscales (e.g., Body Consciousness Questionnaire) were below the acceptable range ($\alpha < .60$). Since the α coefficient depends on the number of items in the scale, the MIIC was additionally considered as a direct indicator of internal consistency. All MIIC were above the lower cutoff value of .15 (Clark & Watson, 2016), indicating a sufficient level of internal consistency. All subscales were therefore retained.

Conclusion

In summary, our findings imply that the EDIP-Q is a suitable measure of interoceptive processing in the domain of maladaptive

eating behaviors and EDs. The EDIP-Q has sound psychometric properties and excellent construct validity. Using the EDIP-Q, we were able to show (a) that deficits in EDIP are significantly present in EDs and improve with recovery, and (b) that the ability to perceive and distinguish between emotions, hunger, and satiety is differentially impaired in AN, BN, and BED. The EDIP-Q provides a differential indication of deficits regarding EDIP, making it a useful clinical assessment tool that offers the basis for developing more targeted and personalized treatment approaches and can be used throughout the course of psychotherapy to track treatment progress.

References

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). https://doi.org/10.1176/appi.books.97808 90425596

Ateş, C., Kaymaz, Ö., Kale, H. E., & Tekindal, M. A. (2019). Comparison of test statistics of nonnormal and unbalanced samples for multivariate analysis of variance in terms of type-I error rates. *Computational and Mathematical Methods in Medicine*, *2019*, Article 2173638. https://doi.org/10.1155/2019/2173638

Backhaus, K., Erichson, B., & Weiber, R. (2015). *Fortgeschrittene multivariate analysemethoden* [Advanced multivariate analysis methods] (3rd ed.). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-662-46087-0

Bentler, P. M., & Bonett, D. G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin*, *88*(3), 588–606. https://doi.org/10.1037/0033-2909.88.3.588

Bornemann, B., Herbert, B. M., Mehling, W. E., & Singer, T. (2015). Differential changes in self-reported aspects of interoceptive awareness through 3 months of contemplative training. *Frontiers in Psychology*, *5*, Article 1504. https://doi.org/10.3389/fpsyg.2014.01504

Bortz, J., & Döring, N. (2016). *Forschungsmethoden und evaluation* [Research methods and evaluation]. Springerverlag.

Brown, T. A. (2015). *Confirmatory factor analysis for applied research* (2nd ed.). Guilford Press.

Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 136–162). Sage Publications.

Bühner, M. (2011). *Einführung in die Test-und Fragebogenkonstruktion* [Introduction to test and questionnaire construction]. Pearson Deutschland GmbH.

Cameron, O. G. (2001). Interoception: The inside story—A model for psychosomatic processes. *Psychosomatic Medicine*, *63*(5), 697–710. https://doi.org/10.1097/00006842-200109000-00001

Clark, L. A., & Watson, D. (2016). Constructing validity: Basic issues in objective scale development. In A. E. Kazdin (Ed.), *Methodological issues and strategies in clinical research* (4th ed., pp. 187–203). American Psychological Association. https://doi.org/10.1037/14805-012

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawerence Erlbaum Associates.

Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112*(1), 155–159. https://doi.org/10.1037/0033-2909.112.1.155

Fichter, M. M., & Quadflieg, N. (2016). Mortality in eating disorders—Results of a large prospective clinical longitudinal study. *International Journal of Eating Disorders*, *49*(4), 391–401. https://doi.org/10.1002/eat.22501

Fichter, M. M., Quadflieg, N., Crosby, R. D., & Koch, S. (2017). Long-term outcome of anorexia nervosa: Results from a large clinical longitudinal study. *International Journal of Eating Disorders*, *50*(9), 1018–1030. https://doi.org/10.1002/eat.22736

Field, A. (2013). *Discovering statistics using IBM SPSS statistics*. Sage Publications.

Forkmann, T., Scherer, A., Meessen, J., Michal, M., Schächinger, H., Vögele, C., & Schulz, A. (2016). Making sense of what you sense: Disentangling interoceptive awareness, sensibility and accuracy. *International Journal of Psychophysiology*, *109*, 71–80. https://doi.org/10.1016/j.ijpsycho.2016.09.019

Framson, C., Kristal, A. R., Schenk, J. M., Littman, A. J., Zeliadt, S., & Benitez, D. (2009). Development and validation of the mindful eating questionnaire. *Journal of the American Dietetic Association*, *109*(8), 1439–1444. https://doi.org/10.1016/j.jada.2009.05.006

Galmiche, M., Déchelotte, P., Lambert, G., & Tavolacci, M. P. (2019). Prevalence of eating disorders over the 2000–2018 period: A systematic literature review. *The American Journal of Clinical Nutrition*, *109*(5), 1402–1413. https://doi.org/10.1093/ajcn/nqy342

Games, P. A., & Howell, J. F. (1976). Pairwise multiple comparison procedures with unequal n's and/or variances: A Monte Carlo study. *The Journal of Educational Research*, *1*(2), 113–125. https://doi.org/10.3102/10769986001002113

Garfinkel, P. E. (1974). Perception of hunger and satiety in anorexia nervosa. *Psychological Medicine*, *4*(3), 309–315. https://doi.org/10.1017/S0033291700042999

Garfinkel, P. E., Moldofsky, H., & Garner, D. M. (1979). The stability of perceptual disturbances in anorexia nervosa. *Psychological Medicine*, *9*(4), 703–708. https://doi.org/10.1017/S0033291700034012

Garfinkel, P. E., Moldofsky, H., Garner, D. M., Stancer, H. C., & Coscina, D. V. (1978). Body awareness in anorexia nervosa: Disturbances in "body image" and "satiety." *Psychosomatic Medicine*, *40*(6), 487–498. https://doi.org/10.1097/00006842-197810000-00004

Garfinkel, S. N., Seth, A. K., Barrett, A. B., Suzuki, K., & Critchley, H. D. (2015). Knowing your own heart: Distinguishing interoceptive accuracy from interoceptive awareness. *Biological Psychology*, *104*, 65–74. https://doi.org/10.1016/j.biopsycho.2014.11.004

Garner, D. M., Olmstead, M. P., & Polivy, J. (1983). Development and validation of a multidimensional eating disorder inventory for anorexia nervosa and bulimia. *International Journal of Eating Disorders*, *2*(2), 15–34. https://doi.org/10.1002/1098-108X(198321)2:2<15::AID-EAT2260020203>3.0.CO;2-6

Glaesmer, H., Grande, G., Braehler, E., & Roth, M. (2011). The German version of the satisfactionwith life scale (SWLS). *European Journal of Psychological Assessment*, *27*(2), 127–132. https://doi.org/10.1027/1015-5759/a000058

Hilbert, A., & Tuschen-Caffier, B. (2007). Eating disorder examination-questionnaire. *Diagnostica*, *53*(3), 144–1541. https://doi.org/10.1026/0012-1924.53.3.144

Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika*, *30*(2), 179–185. https://doi.org/10.1007/BF02289447

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, *6*(1), 1–55. https://doi.org/10.1080/10705519909540118

Jenkinson, P. M., Taylor, L., & Laws, K. R. (2018). Self-reported interoceptive deficits in eating disorders: A meta-analysis of studies using the eating disorder inventory. *Journal of Psychosomatic Research*, *110*, 38–45. https://doi.org/10.1016/j.jpsychores.2018.04.005

Kays, K., Gathercoal, K., & Buhrow, W. (2012). Does survey format influence self-disclosure on sensitive question items? *Computers in Human Behavior*, *28*(1), 251–256. https://doi.org/10.1016/j.chb.2011.09.007

Khalsa, S. S., Adolphs, R., Cameron, O. G., Critchley, H. D., Davenport, P. W., Feinstein, J. S., Feusner, J. D., Garfinkel, S. N., Lane, R. D., Mehling, W. E., Meuret, A. E., Nemeroff, C. B., Oppenheimer, S., Petzschner, F. H., Pollatos, O., Rhudy, J. L., Schramm, L. P., Simmons,

W. K., Stein, M. B., … Interoception Summit 2016 participants. (2018). Interoception and mental health: A roadmap. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*, *3*(6), 501–513. https://doi.org/10.1016/j.bpsc.2017.12.004

Klastrup, C., Frølich, J., Winkler, L. A. D., & Støving, R. K. (2020). Hunger and satiety perception in patients with severe anorexia nervosa. *Eating and Weight Disorders*, *25*(5), 1347–1355. https://doi.org/10.1007/s40519-019-00769-7

Lorenzo-Seva, U., & ten Berge, J. M. F. (2006). Tucker's congruence coefficient as a meaningful index of factor similarity. *Methodology: European Journal of Research Methods for the Behavioral and Social Sciences*, *2*(2), 57–64. https://doi.org/10.1027/1614-2241.2.2.57

Martin, E., Dourish, C. T., Rotshtein, P., Spetter, M. S., & Higgs, S. (2019). Interoception and disordered eating: A systematic review. *Neuroscience and Biobehavioral Reviews*, *107*, 166–191. https://doi.org/10.1016/j.neubiorev.2019.08.020

Mayr, A., Gefeller, O., Prokosch, H. U., Pirkl, A., Fröhlich, A., & de Zwaan, M. (2012). Web-based data collection yielded an additional response bias—But had no direct effect on outcome scales. *Journal of Clinical Epidemiology*, *65*(9), 970–977. https://doi.org/10.1016/j.jclinepi.2012.03.005

Meule, A., Richard, A., Schnepper, R., Reichenberger, J., Georgii, C., Naab, S., Voderholzer, U., & Blechert, J. (2021). Emotion regulation and emotional eating in anorexia nervosa and bulimia nervosa. *Eating Disorders: The Journal of Treatment & Prevention*, *29*(2), 175–191. https://doi.org/10.1080/10640266.2019.1642036

Miller, L. C., Murphy, R., & Buss, A. H. (1981). Consciousness of body: Private and public. *Journal of Personality and Social Psychology*, *41*(2), 397–406. https://doi.org/10.1037/0022-3514.41.2.397

Monteleone, A. M., & Cascino, G. (2021). A systematic review of network analysis studies in eating disorders: Is time to broaden the core psychopathology to non specific symptoms. *European Eating Disorders Review*, *29*(4), 531–547. https://doi.org/10.1002/erv.2834

Murphy, J., Brewer, R., Plans, D., Khalsa, S. S., Catmur, C., & Bird, G. (2020). Testing the independence of self-reported interoceptive accuracy and attention. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, *73*(1), 115–133. https://doi.org/10.1177/1747021819879826

Murphy, J., Catmur, C., & Bird, G. (2019). Classifying individual differences in interoception: Implications for the measurement of interoceptive awareness. *Psychonomic Bulletin & Review*, *26*(5), 1467–1471. https://doi.org/10.3758/s13423-019-01632-7

Nagl, M., Hilbert, A., de Zwaan, M., Braehler, E., & Kersting, A. (2016). The German version of the Dutch eating behavior questionnaire: Psychometric properties, measurement invariance, and population-based norms. *PLOS ONE*, *11*(9), Article e0162510. https://doi.org/10.1371/journal.pone.0162510

Paul, T., & Thiel, A. (2005). *Eating Disorder Inventory-2 (EDI-2): Deutsche Version*. Hogrefe.

Paulus, C. (2009). *Der Saarbrücker Persönlichkeitsfragebogen SPF (IRI) zur Messung von Empathie: Psychometrische Evaluation Der Deutschen Version Des Interpersonal Reactivity Index* [The Saarbrueck Personality Questionnaire on empathy: Psychometric evaluation of the German

version of the Interpersonal Reactivity Index]. https://doi.org/10.23668/psycharchives.9249

Poovey, K., Ahlich, E., Attaway, S., & Rancourt, D. (2022). General versus hunger/satiety-specific interoceptive sensibility in predicting disordered eating. *Appetite*, *171*, Article 105930. https://doi.org/10.1016/j.appet.2022.105930

Pudel, V., & Westenhöfer, J. (1989). *Fragebogen zum Essverhalten (FEV)-Handanweisung* [Three-Factor Eating Questionnaire (TFEQ)-Manual]. Verlag für Psychologie Dr. CJ Hogrefe.

Quadflieg, N., & Fichter, M. M. (2019). Long-term outcome of inpatients with bulimia nervosa-Results from the Christina Barz Study. *International Journal of Eating Disorders*, *52*(7), 834–845. https://doi.org/10.1002/eat.23084

Richard, A., Meule, A., Georgii, C., Voderholzer, U., Cuntz, U., Wilhelm, F. H., & Blechert, J. (2019). Associations between interoceptive sensitivity, intuitive eating, and body mass index in patients with anorexia nervosa and normal-weight controls. *European Eating Disorders Review*, *27*(5), 571–577. https://doi.org/10.1002/erv.2676

Rindermann, H. (2009). *Emotionale-Kompetenz-Fragebogen: EKF; Einschätzung emotionaler Kompetenzen und emotionaler Intelligenz aus Selbst- und Fremdsicht* [Emotional Competency Questionnaire: Assessment of emotional competencies and emotional intelligence from the perspective of self and others]. Hogrefe.

Schaefer, L. M., Smith, K. E., Anderson, L. M., Cao, L., Crosby, R. D., Engel, S. G., Crow, S. J., Peterson, C. B., & Wonderlich, S. A. (2020). The role of affect in the maintenance of binge-eating disorder: Evidence from an ecological momentary assessment study. *Journal of Abnormal Psychology*, *129*(4), 387–396. https://doi.org/10.1037/abn0000517

Suksasilp, C., & Garfinkel, S. N. (2022). Towards a comprehensive assessment of interoception in a multi-dimensional framework. *Biological Psychology*, *168*, Article 108262. https://doi.org/10.1016/j.biopsycho.2022.108262

Tabachnik, B. G., & Fidell, L. S. (2007). *Using multivariate statistics* (5th ed.). Allyn and Bacon.

Tucker, L. R. (1951). *A method for synthesis of factor analysis studies*. Personnal Research Section Report No. 984. Educational Testing Service.

van Dyck, Z., Herbert, B. M., Happ, C., Kleveman, G. V., & Vögele, C. (2016). German version of the intuitive eating scale: Psychometric evaluation and application to an eating disordered population. *Appetite*, *105*, 798–807. https://doi.org/10.1016/j.appet.2016.07.019

van Dyck, Z., Schulz, A., Blechert, J., Herbert, B. M., Lutz, A. P. C., & Vögele, C. (2021). Gastric interoception and gastric myoelectrical activity in bulimia nervosa and binge-eating disorder. *International Journal of Eating Disorders*, *54*(7), 1106–1115. https://doi.org/10.1002/eat.23291

Westwood, H., Kerr-Gaffney, J., Stahl, D., & Tchanturia, K. (2017). Alexithymia in eating disorders: Systematic review and meta-analyses of studies using the Toronto Alexithymia Scale. *Journal of Psychosomatic Research*, *99*, 66–81. https://doi.org/10.1016/j.jpsychores.2017.06.007