

The Intuitive Eating Scale–2: Item Refinement and Psychometric Evaluation With College Women and Men

Tracy L. Tylka
Ohio State University

Ashley M. Kroon Van Diest
Texas A&M University

The 21-item Intuitive Eating Scale (IES; Tylka, 2006) measures individuals' tendency to follow their physical hunger and satiety cues when determining when, what, and how much to eat. While its scores have demonstrated reliability and validity with college women, the IES-2 was developed to improve upon the original version. Specifically, we added 17 positively scored items to the original IES items (which were predominantly negatively scored), integrated an additional component of intuitive eating (Body–Food Choice Congruence), and evaluated its psychometric properties with 1,405 women and 1,195 men across three studies. After we deleted 15 items (due to low item–factor loadings, high cross-loadings, and redundant content), the results supported the psychometric properties of the IES-2 with women and men. The final 23-item IES-2 contained 11 original items and 12 added items. Exploratory and second-order confirmatory factor analyses upheld its hypothesized 4-factor structure (its original 3 factors, plus Body–Food Choice Congruence) and a higher order factor. The IES-2 was largely invariant across sex, although negligible differences on 1 factor loading and 2 item intercepts were detected. Demonstrating validity, the IES-2 total scores and most IES-2 subscale scores were (a) positively related to body appreciation, self-esteem, and satisfaction with life; (b) inversely related to eating disorder symptomatology, poor interoceptive awareness, body surveillance, body shame, body mass index, and internalization of media appearance ideals; and (c) negligibly related to social desirability. IES-2 scores also garnered incremental validity by predicting psychological well-being above and beyond eating disorder symptomatology. The IES-2's applications for empirical research and clinical work are discussed.

Keywords: intuitive eating, scale development, body image, assessment, eating disorder symptomatology

Supplemental materials: <http://dx.doi.org/10.1037/a0030893.supp>

Research on eating behavior has been quite lopsided—it is heavily centered on describing and predicting disordered eating and less focused on identifying and promoting adaptive eating attitudes and behaviors (Tylka, 2006). Recently, however, the study of intuitive eating has gained momentum. *Intuitive eating* is described as an adaptive form of eating characterized by a strong connection with internal physiological hunger and satiety cues (Tribole & Resch, 1995). Individuals who eat intuitively are not preoccupied with food or dieting and do not label certain foods as “good” or “bad.” Although taste is important, they often choose foods for the purpose of enhancing their body’s functioning. They are aware of and trust their body’s internal hunger and satiety cues and use these cues to determine when and how much to eat. A 24-week intervention based on intuitive eating principles has been shown to decrease cholesterol, triglycerides, and systolic blood pressure among middle-aged women at a 52-week follow-up (Bacon et al., 2002). This same intervention has also been shown to

decrease eating restraint, disinhibited eating, drive for thinness, bulimic symptomatology, body dissatisfaction, poor interoceptive awareness, and depression at a 24-month follow-up (Bacon, Stern, Van Loan, & Keim, 2005). Given the strength-based and prevention foci of their field and training (Gelso & Fretz, 2001), counseling psychologists are well situated to study intuitive eating. The Intuitive Eating Scale (IES; Tylka, 2006) was developed as an initial instrument to lead this research.

The Original Intuitive Eating Scale

Ten principles of intuitive eating were offered by Tribole and Resch (2003), who coined the intuitive eating construct, and Tylka (2006) clustered these principles into three domains and developed the 21-item IES to measure these domains. Unconditional Permission to Eat (UPE) reflects individuals’ willingness to eat when hungry (i.e., not try to stave off hunger) and refusal to label certain foods as forbidden. People who allow themselves to eat unconditionally are less likely to overindulge in food, engage in binge eating, and experience guilt when eating (Polivy & Herman, 1999). Eating for Physical Rather Than Emotional Reasons (EPR) represents individuals’ patterns of eating; they eat because they are physically hungry rather than to cope with emotional distress, such as anxiety, loneliness, and boredom. Indeed, the higher women’s tendency to eat for physical rather than emotional reasons, the lower their binge eating and food preoccupation (Tylka & Wilcox,

Tracy L. Tylka, Department of Psychology, Ohio State University; Ashley M. Kroon Van Diest, Department of Psychology, Texas A&M University.

Correspondence concerning this article should be addressed to Tracy L. Tylka, Department of Psychology, Ohio State University, 1465 Mt. Vernon Avenue, Marion OH 43302. E-mail: tylka.2@osu.edu

2006). Last, Reliance on Hunger and Satiety Cues (RHSC) reflects individuals' trust in their internal hunger and satiety cues and reliance on these cues to guide their eating behavior. Individuals who do not trust and follow their internal hunger and satiety cues lose their innate ability to regulate their food intake and are more likely to experience dietary restraint, weight gain, and emotional eating (Birch & Fisher, 2000; Birch, Fisher, & Davison, 2003). Each of these three domains represents an IES subscale.

Research has supported the psychometric properties of the IES as well as the adaptive properties of intuitive eating with women. Exploratory and confirmatory factor analyses upheld its three-factor structure, with the three subscales loading on a higher order intuitive eating factor (Tylka, 2006). Internal consistency estimates have ranged from .85 to .88 for the IES total scale score and between .85 and .89 for two of its three subscales (Tylka, 2006). Many studies have supported the construct validity of its scores with women. For example, studies with women between 18 and 65 years of age have consistently found that the IES total and subscale scores are (a) negatively associated with disordered eating symptoms (Shouse & Nilsson, 2011; Tylka, 2006; Tylka & Wilcox, 2006), thin-ideal internalization (Tylka, 2006), body mass index (Augustus-Horvath & Tylka, 2011; Dittmann & Freedman, 2009; Kroon Van Diest & Tylka, 2010), and self-silencing (Shouse & Nilsson, 2011) as well as (b) positively associated with body appreciation (Augustus-Horvath & Tylka, 2011; Avalos & Tylka, 2006; Dittmann & Freedman, 2009; Kroon Van Diest & Tylka, 2010; Oh, Wiseman, Hendrickson, Phillips, & Hayden, 2012). The IES has also been positively associated with measures of psychological well-being, such as optimism, proactive coping, general life satisfaction, and self-esteem (Tylka, 2006; Tylka & Wilcox, 2006) as well as awareness and acceptance of emotions and different emotional states (Shouse & Nilsson, 2011). Furthermore, the IES was unrelated to social desirability among college women, supporting the discriminant validity of its scores (Tylka, 2006).

While the original IES provides a psychometrically sound measure for assessing intuitive eating in women, the measure is limited in three main ways. First, the original IES does not assess a component of intuitive eating articulated by Tribole and Resch (2003, p. 193): that of honoring health or practicing "gentle nutrition." *Gentle nutrition* reflects the tendency for individuals to make food choices that honor their health and body functioning (e.g., choosing foods that promote energy, stamina, and body performance) as well as taste good. Individuals who practice gentle nutrition may ask themselves "How does this food make my body feel? Do I like this feeling?" and "Does this food give me lasting energy? How do I feel after I eat it?" and may use this information to guide their future food choices, but not in a rigid way. This component was likely missed in the original IES because it was not presented alongside the other three aspects of intuitive eating in Tribole and Resch's (1995) original *Intuitive Eating* text, which shaped the development of Tylka's (2006) original IES. Tribole and Resch (1995, p. 233) placed gentle nutrition at the end of their book out of concern that individuals would focus on nutrition before learning the other intuitive eating principles: "Our experience has shown that if a healthy relationship with food is not in place first, it's difficult to pursue a truly healthy diet." Nevertheless,

they emphasized the importance of gentle nutrition to the intuitive eating experience: If intuitive eating is considered adaptive, it should also represent individuals often choosing nutritious foods to help their bodies function well. Thus, we added items to the IES-2 in order to assess this domain of intuitive eating and labeled it Body-Food Choice Congruence (B-FCC), as it measures the extent to which individuals match their food choices with their bodies' needs.

Second, although intuitive eating is characterized in part by resistance to dieting and emotional eating, intuitive eating should not be largely measured by the absence of these attitudes and behaviors. In the original IES, 13 out of 21 items were written to assess the absence of intuitive eating attitudes and behaviors (e.g., "I don't trust myself around fattening foods," "I use food to help me soothe my negative emotions"). In this study, items added to the IES-2 were written to assess the presence of intuitive eating attitudes and behaviors (e.g., "I allow myself to eat what food I desire at the moment") with the intent of replacing some of the original IES items while maintaining its psychometric integrity. Further, on the original IES, the 13 items that were written to assess the absence of intuitive eating have to be reverse scored, which increases the difficulty of calculating total and subscale scores as well as the likelihood of miscalculation if all necessary items are not correctly reverse scored. Thus, all newly-developed items were positively scored.

Third, Cronbach's alpha estimates for the original RHSC subscale have been in the low to mid .70s (Tylka, 2006; Tylka & Wilcox, 2006), which is at the low end of the acceptable limit (i.e., .70) set forth by Nunnally and Bernstein (1994). Therefore, we added and replaced items in an attempt to improve internal consistency reliability estimates for this subscale in the IES-2.

While not a limitation of the original IES's structural integrity, the fact that the sample consisted entirely of women limited the investigation of its psychometric properties (Tylka, 2006). To date, research has not estimated the IES's psychometric properties with men. Hence, in the present study, we examined whether the IES-2 yielded reliable and valid scores with samples of both men and women.

The Present Research

Additional research on intuitive eating is necessary to increase researchers' and practitioners' understanding of ways to promote adaptive eating attitudes and behaviors. This study examined whether the IES-2 could guide this research. The IES-2 was developed to achieve the following goals: (a) to yield four factors to capture the four domains of intuitive eating, with each factor loading on a higher order intuitive eating factor; (b) to demonstrate strong psychometric properties with both women and men; and (c) to include mostly positively scored rather than negatively scored items. We chose to begin this investigation of the IES-2 with college students because (a) the original IES's psychometric properties were explored with college women (thus providing a direct comparison), (b) college students represent the nonclinical population necessary to be able to find and assess individuals who eat adaptively, and (c) large samples needed to perform factor analyses and tests of measurement invariance could be obtained. Three studies are presented to assess the IES-2, each of which examined various dimensions of its psychometric properties.

Study 1

The purpose of Study 1 was to develop items for the IES-2, explore its factor structure, and examine the internal consistency reliability, convergent validity, and test-retest reliability of its scores. Worthington and Whittaker's (2006) recommendations for scale development were followed.

Development of Additional IES Items and Expert Review

Newly developed items for the IES-2 were drafted by Tracy Tylka, who consulted a psychologist and a nutritionist, both of whom had extensive knowledge of intuitive eating. Item content was guided by the themes noted throughout Tribble and Resch's (2003) *Intuitive Eating* text. The item development team also considered the original 21 items when designing the new items. In total, 17 items were developed: two for UPE, seven for EPR, four for RHSC, and four for B-FCC. These 17 items were designed to be positively scored in order to offset the large number of reverse-scored items included in the original IES.

Following item generation, the 17 developed items and the 21 original items were reviewed by two other researchers familiar with intuitive eating and its components to (a) examine the extent to which the items comprehensively assessed the four components of intuitive eating and (b) determine whether the wording of the items needed to be modified for clarity or style. On the basis of their recommendations, five of the developed items were modified. All developed items (Items 22–38), along with the original IES items (Items 1–21), can be found in the supplemental materials online.

Hypotheses

First, we hypothesized that the IES-2 would adhere to a four-factor solution: the three original factors and the newly-developed B-FCC factor. The factor structure was evaluated separately for women and men to ensure that the emerging factors and their items were equivalent. Second, we expected that the items within these dimensions, as well as all IES-2 items, would be internally consistent because they measure intuitive eating as well as specific facets of this construct. Third, we predicted that the original IES and IES-2 total scale and subscale scores would tap into the same construct (i.e., because many new items were created, we wanted to make sure that the original and revised scales/subscales would be similar in content). Fourth, we hypothesized that men would have higher IES-2 total and subscale scores than women given that men experience fewer sociocultural pressures than women to be ultra-thin and to restrict their eating (Fredrickson & Roberts, 1997; Kroon Van Diest & Tylka, 2010). Fifth, we expected that IES-2 total and subscale scores would be stable over a 3-week period.

Method

Participants. Data were analyzed from 878 college students (487 women and 391 men) from a regional campus of a large Midwestern university. Participants ranged in age from 18 to 56 years ($M = 20.40$ years, $SD = 5.19$). They identified as White (77.3%), African American (13.1%), Asian American (4.0%), Latina/Latino (1.3%), Native American (0.7%), or multiracial

(2.7%). Eight participants (0.9%) did not report an ethnic identification. Participants indicated that they were first-year college students (75.7%), sophomores (18.2%), juniors (4.0%), or seniors (2.1%).

A subset of these participants ($n = 219$; 140 women and 79 men) completed the IES-2 three weeks after the first administration to estimate its stability. They ranged in age from 18 to 47 years ($M = 20.33$ years, $SD = 4.61$). They identified as White (80.4%), African American (9.2%), multiracial (6.4%), Asian American (3.2%), Latino/Latina (0.5%), or Native American (0.5%). They were first-year students (75.8%), sophomores (18.3%), juniors (4.1%), or seniors (1.8%).

Measure. The 21 original IES items and the 17 newly developed items were administered. Participants were instructed, "For each item, please check the answer that best characterizes your eating attitudes or behaviors." The item-response scale mirrored the original 5-point scale (Tylka, 2006): 1 = *strongly disagree*, 2 = *disagree*, 3 = *neutral*, 4 = *agree*, 5 = *strongly agree*.

Procedure. All procedures were approved by the university's Institutional Review Board (IRB). Participants were recruited in introductory psychology classes over a 2-year period. The study was posted alongside other studies on the psychology department's research management website. The study was described as an investigation of eating habits among college students. Interested participants could sign up for the study on the website and, when ready to complete the survey, click a link to a webpage that hosted the informed consent sheet. After providing their consent, students were directed to the survey webpage. Students received research credit for their participation.

Several strategies were used to detect and delete duplicate and erroneous data. Participants were deleted if they took the survey more than once ($n = 10$), failed at least one of two validity questions (e.g., "Do not answer this item so we know you are paying attention," $n = 31$), terminated early ($n = 75$), or had significant missing data (five or more IES-2 items, $n = 4$). From the initial data set of 998 participants, 878 remained.

For additional research credit, participants had the option of completing the IES-2 again 3 weeks after they initially completed the first part of the study. Participants were not made aware that they were taking the same survey again or that the purpose of the study was to gauge the stability of the IES-2; they were told that this was a follow-up study on their eating habits. Interested participants signed up for the study via the research management system and were told that they needed to take the follow-up survey exactly 3 weeks later. Nineteen days after they took the first part of the survey, Tracy Tylka sent them an e-mail with the link to the informed consent sheet and the study (Day 19 was chosen to give participants a small time window to receive the e-mail and complete the study). A total of 297 of the original participants completed the second part to this study and were matched (via the student code assigned to them by the research management system) to their initial answers on the IES-2 items. Participants who submitted their surveys more than once ($n = 2$), terminated early or left more than five IES-2 items blank ($n = 20$), were random in their responding (as indicated by their inaccurate responses to the two validity questions, $n = 19$), or completed the survey later than 23 days after the original administration ($n = 37$) were deleted. These screening procedures reduced the test-retest data set to 219

participants, who, on average, completed the second administration 20.57 days after the first administration.

Results and Discussion

Preliminary analyses. A total of 17.0% of the participants had at least one missing data point. For individual IES-2 items, the count for missing data points ranged from 0 to 1.3%. Data were missing completely at random as indicated by Little's Missing Completely at Random (MCAR) analysis, $\chi^2(542) = 586.24$, $p = .092$, suggesting that the amount and pattern of missingness were not problematic. Thus, multiple imputation (MI) via SPSS 19.0 was used to estimate missing data values. Specifically, five imputed data sets were created using available data, and subsequent analyses were conducted on all five data sets, with results pooled.

IES-2 items were examined for normality of distribution. Skewness and kurtosis were evaluated by sex because the IES-2 factor structure was evaluated separately for women and men. Items with skewness values > 3 and/or kurtosis values > 10 may pose problems in regression and structural analyses and therefore should be transformed (Kline, 2005). Skewness and kurtosis values for the IES-2 items were lower than these limits; thus, no item was transformed.

Item means and standard deviations both for the combined sample and for women and men separately are presented in the supplemental materials online. After controlling for experiment-wise error via the Bonferroni correction (i.e., $p = .05/38$ items, or .00131), we found that women and men differed significantly on 22 items, with men reporting higher levels of intuitive eating for each of these items.

Exploring the IES-2's factor structure. Two principal axis factor (PAF) analyses were conducted on the 38 IES-2 items using SPSS 19.0: the first on the sample of women ($n = 487$), and the second on the sample of men ($n = 391$). For each analysis, the number of participants exceeded the recommended 10:1 cases-to-parameter ratio needed to confidently examine a model (Bentler, 1990). In this study, a minimum of 380 participants were required for each analysis because there were 38 item-factor parameters to be estimated. For each PAF analysis, direct oblimin rotation with a delta weight of zero was chosen to allow for moderate relationships between the factors.

Parallel analysis was used to determine the number of factors to extract. Parallel analysis more accurately estimates the number of factors in a data set than do the eigenvalue > 1 criterion, examining the scree plot of the eigenvalues for breaks or discontinuities, and maximum likelihood procedures (Brown, 2006; Fabrigar, Wegener, MacCallum, & Strahan, 1999). The basis of parallel analysis is that the factors underlying a measure should account for more variance than is expected by chance. Therefore, factor analysis is performed on the actual data as well as multiple sets of random data (in this case, 1,000) that have the same dimensions as the actual data set. If the eigenvalue generated from the analysis of the actual data exceeds the corresponding pooled eigenvalue from the analysis of the random data, then that factor is retained.

Items were retained if they had an item-factor loading of at least .50 on a primary factor and cross-loadings less than .30 on other factors (Brown, 2006; Tabachnick & Fidell, 2007). However, the anti-image correlation matrix was also used to determine if the item was highly correlated with other items. Only one item within

each set of highly correlated items was retained. If two items were similar and one was reverse scored and one was positively scored, the positively scored item was retained.

For women's data, the size of the Kaiser-Meyer-Okin (KMO) measure of sampling adequacy ($KMO = .901$) revealed that the IES-2 items had adequate common variance for factor analysis, and the significance of Bartlett's test of sphericity, $\chi^2(703) = 12,486.53$, $p < .001$, suggested that the correlation matrix was factorable (Tabachnick & Fidell, 2007). Results of the parallel analysis suggested that four factors should be retained. A total of 27 items had factor loadings $\geq .50$ on their primary factor and cross-loadings $\leq .30$ on the other three factors. The anti-image correlation matrix revealed that two retained item pairs had high correlations: Items 6 and 24 ($r = -.54$) and Items 22 and 23 ($r = -.71$). Item 24 was retained because it was positively scored, and Item 23 was retained because it had a higher factor loading than Item 22. Additionally, all three positively scored items for UPE did not reach the $\geq .50$ criterion, which is possibly due to the large number of reverse scored items on this factor. Thus, the three positively scored items (Items 4, 28, and 29) and three of the eight reverse scored items (Items 1, 9, and 14) that best represented the UPE construct were retained.

Therefore, of the original 38 items, 23 were retained (eight were deleted for low primary item-factor loadings and/or high cross-loadings, two were deleted for redundant content, and five were deleted for being reverse scored items on the UPE factor). A second factor analysis using PAF with oblimin rotation ($\delta = 0$) on this set of 23 items revealed a four-factor solution, which accounted for 65.06% of the total variance in the items. Item-factor loadings exceeded .50 on their primary factor and were less than .30 on the other factors. Table 1 includes these 23 items and the item-factor loadings.

For men's data, the size of the Kaiser-Meyer-Okin measure of sampling adequacy ($KMO = .903$) and the significance of Bartlett's test of sphericity, $\chi^2(703) = 9293.30$, $p < .001$, indicated that we could proceed with the factor analysis. Similar to results with the women's data, results of the parallel analysis of men's data suggested that four factors should be retained. A total of 25 items had factor loadings $\geq .50$ on their primary factor and cross-loadings $\leq .30$ on other factors (for men, Items 7 and 15 produced clear item-factor loadings, whereas they did not for women). The anti-image correlation matrix revealed that the following retained item combinations had high correlations: Items 4 and 28 ($r = -.44$), Items 5 and 29 ($r = -.43$), Items 6 and 24 ($r = -.55$), and the trio of Items 7, 15, and 37 ($r_s = -.44, -.47$, and $-.39$). Items 28, 29, and 24 were retained because they were positively scored. Items 7 and 15 were deleted and Item 37 was retained to be consistent with the items selected via the analysis of the women's data. Similar to the results with the women's data, Items 4, 28, and 29 (all positively scored) did not load highly on the UPE factor, which also contained eight items that were reverse scored. This factor's three positively scored items and three reverse scored items (Items 1, 9, and 14) were retained, and its five reverse scored items (Items 5, 18, 19, 20, and 21) were deleted.

The remaining 23 items were the same items obtained from the analysis of the women's data. A second factor analysis using PAF with oblimin rotation ($\delta = 0$) on these 23 items yielded a four-factor solution that accounted for 63.92% of the total variance in the items for men. Item-factor loadings exceeded .50 on their

Table 1
Intuitive Eating Scale-2 (IES-2) Item-Factor Loadings for Study 1

Factor and item	Factor 1	Factor 2	Factor 3	Factor 4	Item-total <i>r</i>
Factor 1: Eating for Physical Rather Than Emotional Reasons					
3. I find myself eating when I'm feeling emotional (e.g., anxious, depressed, sad), even when I'm not physically hungry. ^a	.84/.78	.03/.09	.02/.03	.02/.01	.77/.74
10. I find myself eating when I am lonely, even when I'm not physically hungry. ^a	.83/.86	.01/.05	.02/.00	.01/.01	.78/.80
16. I use food to help me soothe my negative emotions. ^a	.90/.82	.04/.05	.08/.05	.01/.04	.83/.76
17. I find myself eating when I am stressed out, even when I'm not physically hungry. ^a	.91/.85	.04/.04	.14/.01	.03/.01	.82/.80
23. I am able to cope with my negative emotions (e.g., anxiety, sadness) without turning to food for comfort.	.62/.50	.03/.12	.17/.11	.07/.13	.69/.59
24. When I am bored, I do NOT eat just for something to do.	.62/.71	.15/.15	.20/.04	.01/.07	.68/.68
25. When I am lonely, I do NOT turn to food for comfort.	.78/.83	.05/.11	.10/.09	.03/.00	.81/.82
26. I find other ways to cope with stress and anxiety than by eating.	.61/.60	.02/.07	.22/.17	.05/.09	.70/.68
Factor 2: Unconditional Permission to Eat					
1. I try to avoid certain foods high in fat, carbohydrates, or calories. ^a	.02/.09	.68/.66	.07/.02	-.15/-.21	.63/.66
4. If I am craving a certain food, I allow myself to have it.	.12/.19	.59/.69	.06/.05	.00/-.08	.51/.53
9. I get mad at myself for eating something unhealthy. ^a	.22/.28	.60/.51	.05/.00	.04/-.09	.52/.56
14. I have forbidden foods that I don't allow myself to eat. ^a	.06/.12	.71/.64	.12/.05	-.06/-.06	.62/.62
28. I allow myself to eat what food I desire at the moment.	-.12/-.02	.68/.74	.09/.09	-.04/-.07	.60/.65
29. I do NOT follow eating rules or dieting plans that dictate what, when, and/or how much to eat.	-.01/-.02	.63/.56	.11/.13	-.02/-.03	.57/.55
Factor 3: Reliance on Hunger and Satiety Cues					
11. I trust my body to tell me when to eat.	.01/.01	.04/.01	.73/.78	.06/.00	.70/.73
12. I trust my body to tell me what to eat.	.02/.02	.06/.06	.61/.68	.06/.01	.60/.62
13. I trust my body to tell me how much to eat.	.07/.04	.02/.00	.86/.85	.08/.01	.76/.79
35. I rely on my hunger signals to tell me when to eat.	.08/.00	.08/.10	.60/.63	.04/.02	.62/.63
36. I rely on my fullness (satiety) signals to tell me when to stop eating.	.08/.08	.04/.06	.66/.69	.06/.00	.65/.70
37. I trust my body to tell me when to stop eating.	.04/.08	.03/.00	.85/.84	.05/.02	.77/.80
Factor 4: Body-Food Choice Congruence					
31. Most of the time, I desire to eat nutritious foods.	.02/.06	-.11/-.19	.01/.05	.65/.63	.66/.65
32. I mostly eat foods that make my body perform efficiently (well).	-.02/.03	.01/.01	.06/.02	.98/.87	.83/.79
33. I mostly eat foods that give my body energy and stamina.	.01/.03	.07/.11	-.02/-.04	.88/.89	.77/.74

Note. Values to the left of the diagonal are item-factor loadings derived from women's data ($n = 487$), and values to the right of the diagonal are item-factor loadings derived from men's data ($n = 391$). Numbers corresponding to the items are from the original list of 38 items in Study 1 (see online supplemental materials Table 1) and do not correspond to the numbered items in the Appendix.

^a Reverse scored.

primary factor and were less than .30 on the other factors. The item-factor loadings are presented in Table 1.

The following findings are presented for women's and men's data, respectively. Factor 1 accounted for 31.17% and 31.75% (eigenvalues = 7.17 and 7.94) of the variance after rotation and was labeled Eating for Physical Rather Than Emotional Reasons (EPR). Eight items loaded on this factor, and its item-factor loadings ranged from .61 to .91 for women and from .50 to .86 for men. Factor 2, labeled Unconditional Permission to Eat (UPE), accounted for 16.97% and 15.69% (eigenvalues = 3.90 and 3.92) of the variance after rotation. Six items loaded on this factor, and item-factor loadings ranged from .59 to .71 for women and from .51 to .74 for men. Factor 3, Reliance on Hunger and Satiety Cues (RHSC), included six items that accounted for 11.50% and 10.74% (eigenvalues = 2.65 and 2.68) of the variance after rotation; its item-factor loadings ranged from .60 to .86 for women and from .63 to .85 for men. Last, Factor 4, Body-Food Choice Congruence (B-FCC), included three items that accounted for 5.43% and 5.74% (eigenvalues = 1.25 and 1.30) of the variance after rotation,

and its factor loadings ranged from .65 to .98 for women and from .63 to .89 for men. According to the factor correlation matrix for women's and men's data, respectively, EPR correlated .05 and .28 with UPE, .35 and .37 with RHSC, and .36 and .26 with B-FCC. UPE correlated .25 and .35 with RHSC and -.30 and -.28 with B-FCC. RHSC and B-FCC were correlated .40 for women and .16 for men.

Despite being analyzed separately, women's and men's data both converged on a four-factor solution for the IES-2 items. The same items loaded on each factor for women and men, yielding a 23-item measure (see the Appendix) that can be administered to both women and men.

Internal consistency reliability. For women and men, respectively, Cronbach's coefficient alphas were .87 and .89 for the total 23-item IES-2, .93 and .92 for EPR, .81 and .82 for UPE, .88 and .89 for RHSC, and .87 and .85 for B-FCC. Item-total correlations are presented in Table 1. Consequently, IES-2 total and subscale scores were internally consistent for women and men.

Convergent validity. The IES-2 total and subscale scores should assess the same construct as the original IES total and subscale scores. Therefore Pearson r correlation coefficients between the original and revised scale/subscales should be .80 or higher, the criterion for multicollinearity or substantial construct overlap (Kline, 2005; Studenmund, 2006). Correlations were .87 for women and .91 for men between the original IES and the IES-2 total scores, .95 for women and .94 for men between the original and revised EPR subscales, .90 for both women and men between the original and revised UPE subscales, and .86 for women and .90 for men between the original and revised RHSC subscales. Thus, the original IES and the IES-2 scales and subscales have considerable construct overlap.

IES-2 total scale and subscale means. Means for the IES-2 total scale and subscales are included in Table 2. To determine whether women's and men's means were significantly different, the Bonferroni correction was used (i.e., $.05/5 = .01$). Additionally, Cohen's d was used to estimate the degree of difference between women's and men's average values, whereby a value of .30 represents a small effect size, .50 indicates a medium effect size, and .80 reveals a large effect size (Cohen, 1988). Therefore, when compared to women, men demonstrated higher IES-2 total scores and EPR subscale scores, yielding medium to large effects, and higher UPE and RHSC subscale scores, yielding small effects. Women and men had similar scores on B-FCC.

Test-retest reliability. Intraclass correlation coefficients were used to estimate the stability of the IES-2's scores using data from the subsample of 219 participants who completed the IES-2 at both administrations 3 weeks apart. The relationship between the first and second administrations was .88 among women and .92 among men for the IES-2 total score, .81 among women and .84 among men for EPR, .86 among women and .89 among men for UPE, .80 among women and .90 among men for RHSC, and .77 among women and .75 among men for B-FCC. As such, IES-2 total scores and subscale scores were stable across a 3-week period.

Study 2

In Study 2, a confirmatory factor analysis (CFA) of the 23-item IES-2 was conducted in order to determine whether its four-factor structure identified in Study 1 would be confirmed within a different sample of participants. Furthermore, this CFA would allow

us to determine whether the four first-order latent factors load on a higher order intuitive eating factor. We hypothesized that (a) the IES-2 items would load on their respective first-order latent factors, (b) the first-order latent factors would be related and load on a second-order intuitive eating factor, and (c) the overall model would provide an acceptable fit to the data. If supported, these findings would uphold the use of an IES-2 total score as well as subscale scores. This model was also tested for measurement invariance to ensure that the IES-2 assesses the same construct for women and men so meaningful group comparisons can be made. Thus, our model was hypothesized to be invariant across sex.

Also, IES-2 total and subscale scores were expected to demonstrate construct validity via their relationships to several eating- and body-related variables and psychological well-being indices for both women and men. Because intuitive eating is conceptualized as adaptive eating (Tribole & Resch, 2003; Tylka, 2006; Tylka & Wilcox, 2006) and eating disorder symptomatology is conceptualized as maladaptive eating (Garner, Olmsted, Bohr, & Garfinkel, 1982; Tribole & Resch, 2003), IES-2 total and subscale scores were predicted to be negatively related to eating disorder symptomatology. Further, because those who respect and appreciate their bodies are more aware of their bodily needs, including their hunger and satiety signals, and eat according to those signals (Avalos & Tylka, 2006), IES-2 scores were hypothesized to be positively related to body appreciation. Additionally, those who listen to their inner experience to guide their behavior are more likely to resist internalizing societal ideals (Rogers, 1961), which include media appearance ideals, and evaluating their bodies based on these unrealistic standards (Augustus-Horvath & Tylka, 2011). Therefore, IES-2 scores were hypothesized to be inversely related to internalization of media ideals, body surveillance, and body shame. Because intuitive eating requires awareness of internal body experiences, IES-2 scores were predicted to be inversely related to poor interoceptive awareness, or lowered emotional and physiological awareness. Given that emotional, restrained, and situational (i.e., tendency to overeat whenever food cues are present) eating has been found to be related positively to body mass index (BMI) to a small degree (Koenders & van Strien, 2011), intuitive eating should be inversely related to BMI to a small degree. Therefore, we hypothesized that IES-2 scores would be inversely related to BMI. Each of these relationships has been demonstrated using the original 21-item IES among women

Table 2

A Comparison of IES-2 Total and Subscale Mean (SD) Scores for Women and Men Across Studies 1–3

IES-2 total score and subscale	Study 1				Study 2				Study 3			
	Women	Men	t	d	Women	Men	t	d	Women	Men	t	d
IES-2 total score	3.37 (0.51)	3.69 (0.54)	-9.00*	-.61	3.38 (0.51)	3.67 (0.53)	-9.56*	-.55	3.38 (0.48)	3.67 (0.52)	-6.58*	-.58
EPR	3.19 (0.90)	3.77 (0.85)	-9.58*	-.65	3.17 (0.90)	3.69 (0.87)	-9.94*	-.57	3.18 (0.85)	3.72 (0.84)	-7.54*	-.66
UPE	3.46 (0.76)	3.69 (0.80)	-4.33*	-.29	3.48 (0.76)	3.65 (0.78)	-3.77*	-.22	3.50 (0.70)	3.70 (0.80)	-3.16*	-.28
RHSC	3.52 (0.70)	3.72 (0.71)	-4.09*	-.28	3.53 (0.61)	3.68 (0.65)	-4.21*	-.24	3.57 (0.66)	3.71 (0.66)	-2.29	-.20
B-FCC	3.34 (0.80)	3.40 (0.80)	-0.99	-.07	3.35 (0.79)	3.48 (0.77)	-2.84*	-.16	3.29 (0.80)	3.33 (0.79)	-0.48	-.04

Note. Study 1 $N = 878$ ($n = 487$ women, $n = 391$ men); Study 2 $N = 1,200$ ($n = 680$ women, $n = 520$ men); Study 3 $N = 522$ (238 women, 284 men). A Cohen's $d \approx .30$ represents a small effect size, $\approx .50$ indicates a medium effect size, and $\approx .80$ reveals a large effect size. IES-2 = Intuitive Eating Scale-2, EPR = Eating for Physical Rather Than Emotional Reasons, UPE = Unconditional Permission to Eat, RHSC = Reliance on Hunger and Satiety Cues, B-FCC = Body-Food Choice Congruence.

* $p < .01$.

(Tylka, 2006), and many of these links have been supported with men (Kroon Van Diest & Tylka, 2010); therefore, these relationships were expected to hold for the IES-2 as well.

IES-2 total and subscale scores were also expected to be related to indices of psychological well-being, further supporting their construct validity. Intuitive eating is conceptualized as adaptive to psychological health and has been found to be linked to many indices of psychological strength (Tylka & Wilcox, 2006); hence, IES-2 scores were expected to be positively related to self-esteem, positive affect, and life satisfaction and inversely related to negative affect. IES-2 scores were also hypothesized to predict these well-being indices above and beyond their associations with eating disorder symptomatology. These results would support the distinctiveness of the IES-2 from low levels of eating disorder symptomatology, which would uphold the incremental validity of its scores. Indeed, the incremental validity of the original IES has been supported in this manner (Tylka & Wilcox, 2006).

Method

Participants. Study 2 included data from 1,200 college students (680 women and 520 men) from the same regional campus of the large Midwestern university used in Study 1. Participants ranged in age from 18 to 53 years ($M = 20.45$ years, $SD = 5.06$). They identified as White (81.7%), African American (5.5%), Asian American (3.5%), Latina/Latino (1.8%), Native American (0.1%), or multiracial (4.5%); 2.9% endorsed "other." They categorized themselves as first-year college students (75.0%), sophomores (16.3%), juniors (4.3%), or seniors (4.5%). Participants' self-reported height and weight were used to calculate BMI (kg/m^2). Women's average BMI was 24.02 ($SD = 5.68$, range = 15.98 to 56.25), and men's average BMI was 25.38 ($SD = 5.48$, range = 16.50 to 59.06).

Measures.

Intuitive eating. The 23-item IES-2 (see Appendix) described in Study 1 was used in Study 2. In Study 2, Cronbach's alphas among women and men, respectively, were .88 and .89 for the IES-2 total score, .93 and .92 for EPR, .81 and .82 for UPE, .85 and .89 for RHSC, and .86 and .83 for B-FCC.

Eating disorder symptomatology. The Eating Attitudes Test-26 (EAT-26; Garner et al., 1982) assessed participants' levels of disordered eating attitudes and behaviors. Its 26 items (e.g., "I am terrified about being overweight") are rated along a 6-point scale ranging from 1 (*never*) to 6 (*always*). The continuous scoring method was followed because it utilizes the full range of responses, which reduces skewness in the distribution of scores due to the relatively low base rate of clinical eating disorders (e.g., Mazzeo, 1999; Tylka & Subich, 2004). Therefore, the values assigned to each item were averaged to arrive at the total score. Higher scores reflect greater eating disorder symptomatology. When scored continuously, EAT-26 scores have been estimated to be internally consistent (e.g., $\alpha = .91$ for college women, Tylka, 2006; $\alpha = .91$ for college men, Tylka, 2011) and stable across a 3-week period ($r = .86$ for college women; Mazzeo, 1999) as well as strongly related to an eating disorder diagnostic instrument ($r = .66$ for college women; Tylka & Subich, 2004). In the present study, Cronbach's alphas for the EAT-26 items were .91 among both women and men.

Poor interoceptive awareness. The 10-item Interoceptive Awareness subscale of the Eating Disorder Inventory-2 (EDI-2-IA; Garner, 1991) was used to measure participants' disconnection to their internal body states, such as emotions, hunger, and satiety. Items are rated along a scale ranging from 1 (*never true of me*) to 6 (*always true of me*). Continuous scoring was used to prevent range restriction; other researchers (Tylka, 2006; Tylka & Subich, 2004) have used this scoring method with college students. Scores are averaged, and higher scores indicate poorer interoceptive awareness. Among college students, scores on the EDI-2-IA subscale have been estimated to be internally consistent ($\alpha = .86$; Anestis, Holm-Denoma, Gordon, Schmidt, & Joiner, 2008), stable over a 3-week period ($r = .85$; Wear & Pratz, 1987), and related to alexithymia ($r = .77$; Tylka & Subich, 2004). For the present study, Cronbach's alphas for EDI-2-IA items were .87 among women and .84 among men.

Body appreciation. The Body Appreciation Scale (BAS; Avalos, Tylka, & Wood-Barcalow, 2005) is a 13-item instrument designed to assess individuals' favorable opinions and treatment of their bodies despite their perceived appearance flaws. Its items (e.g., "Despite its imperfections, I still like my body") are rated along a 5-point scale that ranges from 1 (*never*) to 5 (*always*). Item responses are averaged, and higher scores reflect greater body appreciation. Among college women, BAS scores have been estimated to be internally consistent ($\alpha = .91$), stable over a 3-week period ($r = .90$), and inversely related to body shame ($r = -.73$; Avalos et al., 2005). Among college men, internal consistency estimates ($\alpha = .85$) and correlations with body esteem ($r = .44$) have supported its psychometric properties (Swami, Stieger, Hübner, & Voracek, 2008). For the present study, Cronbach's alphas for BAS items were .93 and .92 among women and men, respectively.

Body surveillance and body shame. Two subscales of the Objectified Body Consciousness Scale (OBCS; McKinley & Hyde, 1996) measured participants' tendency to habitually monitor their appearance (Body Surveillance; eight items) and to feel shame due to their bodies not meeting societal appearance standards (Body Shame; eight items). Body Surveillance items (e.g., "During the day, I think about how I look many times") and Body Shame items (e.g., "I feel like I must be a bad person when I don't look as good as I could") are rated along a 7-point scale ranging from 1 (*strongly disagree*) to 7 (*strongly agree*). Subscale items are averaged, and higher scores reflect greater body surveillance and body shame, respectively. Among college women, Body Surveillance ($\alpha = .89$) has been significantly related to public self-consciousness ($r = .73$), and Body Shame ($\alpha = .75$) has been inversely related to self-esteem ($r = -.51$; McKinley & Hyde, 1996). Among college men, Body Surveillance ($\alpha = .84$) and Body Shame ($\alpha = .77$) have been positively correlated with internalization of societal appearance standards ($r = .49$ for Body Surveillance, $r = .38$ for Body Shame; Parent & Moradi, 2011). In the current study, Cronbach's alphas were .87 for Body Surveillance items (among both women and men) and .86 and .85 (among women and men, respectively) for Body Shame items.

Internalization of media appearance ideals. Women were given the 8-item Internalization subscale of the Sociocultural Attitudes Toward Appearance Questionnaire-Revised (SATAQ-I-R; Heinberg, Thompson, & Stormer, 1995) to assess the extent to which they have internalized the thin media ideal as their personal

standard (e.g., “Women who appear in TV shows and movies project the type of appearance that I see as my goal”). Men were given the 11-item male version of the SATAQ-I-R (Thompson, Heinberg, Altabe, & Tantleff-Dunn, 1999), which measures men’s internalization of the muscular and fit media ideal as their personal standard (e.g., “I would like my body to look like the men who appear in TV shows and magazines”). On both versions, items are rated along a 5-point scale ranging from 1 (*completely disagree*) to 5 (*completely agree*) and averaged. Higher scores reflect greater internalization of media appearance ideals. Among college women, the SATAQ-I-R yielded internally consistent score estimates ($\alpha = .90$; Tylka, 2006) and was strongly related to perceptions of ideal body type ($r_s = .62-.64$; Tylka & Subich, 2004). Among college men, the SATAQ-I-R demonstrated internally consistent score estimates ($\alpha = .87$) and was strongly related to appearance comparison ($r = .57$; Tylka, Bergeron, & Schwartz, 2005). In the present study, Cronbach’s alphas for SATAQ-I-R items were .92 among women and .91 among men.

Self-esteem. Participants’ level of self-esteem was measured by the 10-item Rosenberg Self-Esteem Scale (RSES; Rosenberg, 1965). Its items (e.g., “I feel that I have a number of good qualities”) are rated on a 4-point scale ranging from 1 (*strongly disagree*) to 4 (*strongly agree*). Item responses are averaged, and higher scores reflect greater self-esteem. The estimated internal consistency of its scores has been upheld for women ($\alpha = .90$; Tylka, 2006) and men ($\alpha = .86$; Tylka et al., 2005). Supporting its construct validity with college students, the RSES was found to be strongly related to proactive coping ($r = .63$) and optimism ($r = .73$) among college women (Tylka, 2006) as well as proactive coping ($r = .72$) and hardiness ($r = .64$) among college men (Bergeron & Tylka, 2007). In the present study, Cronbach’s alphas for RSES items were .91 among women and .90 among men.

Positive and negative affect. The Positive and Negative Affect Schedule–Expanded (PANAS-X; Watson & Clark, 1994; Watson, Clark, & Tellegen, 1988) contains two 10-item subscales that were used to gauge participants’ levels of positive affect (PA; e.g., “enthusiastic,” “determined”) and negative affect (NA; e.g., “irritable,” “nervous”). Participants were asked to rate the degree they experienced each emotion in general along a 5-point scale ranging from 1 (*very slightly or not at all*) to 5 (*extremely*). Subscale items are averaged, with higher scores reflecting greater PA and NA. Among college students, PA ($\alpha = .87$) and NA ($\alpha = .85$) have been found to yield stable scores over a 2-month period ($r = .70$ for PA, $r = .71$ for NA; Watson & Clark, 1994) and construct validity via their correlations with depressive symptoms ($r = -.36$ for PA, $r = .58$ for NA) and anxiety ($r = -.35$ for PA, $r = .51$ for NA; Watson et al., 1988). In the present study, Cronbach’s alphas for PA items were .85 for women and .84 for men, and Cronbach’s alphas for NA items were .86 for women and .83 for men.

Life satisfaction. The 5-item Satisfaction With Life Scale (SWLS; Diener, Emmons, Larsen, & Griffin, 1985) was used to assess participants’ global life satisfaction. Its items (e.g., “In most ways my life is close to ideal”) are rated along a 7-point scale ranging from 1 (*strongly disagree*) to 7 (*strongly agree*). Item responses are averaged, and higher scores indicate stronger life satisfaction. Among college students, the SWLS ($\alpha = .87$) has yielded stable scores over a 2-month period ($r = .82$) and shown strong relationships to positive affect ($r = .50$) and self-esteem

($r = .54$; Diener et al., 1985). In the present study, Cronbach’s alphas for SWLS items were .91 among women and .87 among men.

Procedure. After receiving IRB approval, participants were recruited in introductory and upper-level psychology classes over a 3-year period, which commenced 3 months after the data collection from Study 1 ended. This study was posted alongside others on the psychology department’s research management website. Students interested in participating signed up on this website and, from there, clicked a link to a webpage that hosted the informed consent sheet. After providing their consent, they were directed to the survey webpage. This study was described as an investigation of eating habits, body attitudes, and well-being among college students. The measures were counterbalanced to control for order effects.

Twelve students who took this study more than once, 47 participants who failed at least one of the five embedded validity questions, 85 participants who exited the survey before completion, and 13 participants who had significant missing data (not completing at least 80% of each measure) were deleted. These screening procedures reduced the initial data set of 1,357 participants to 1,200.

Results and Discussion

Preliminary analyses. A total of 15.82% of participants had at least one missing data point. The count for missing data points ranged from 0 to 1.0%, and Little’s MCAR analysis indicated that these data points were missing completely at random, $\chi^2(376) = 391.42, p = .281$. MI was used to estimate missing data values by creating five imputed data sets using available data. Subsequent analyses were conducted on all five data sets, with results pooled.

All skewness and kurtosis values for the 23 IES-2 items and all other study measures were below the critical limits (Kline, 2005); thus, no item, scale, or subscale was transformed. IES-2 item means and standard deviations are presented in the online supplemental materials, both for the combined sample and for women and men separately. After using the Bonferroni correction (i.e., $p = .05/23$ items, or .00217), we found that women and men differed significantly on 16 items; in each case, the men’s mean item score was higher. Mean total scale and subscale scores also were calculated for all measures and are presented in Table 2. Men scored higher than women on the IES-2 total scale and on each subscale, with moderate effect sizes emerging between women and men for the IES-2 total score and for EPR, and very small to small effect sizes emerging between men and women for UPE, RHSC, and B-FCC. Again, men reported higher intuitive eating scores than women, although for three subscales the difference was negligible.

Confirming the IES-2’s higher order factor structure. Mplus Version 6.12 (Muthén & Muthén, 1998–2011), with maximum likelihood estimation and the covariance matrix as input, was used to conduct the second-order CFA. Adequacy of model fit was determined via consensus among three indices recommended by Hu and Bentler (1999): the comparative fit index (CFI), the standardized root-mean square residual (SRMR), and the root mean square error of approximation (RMSEA). Specifically, CFI values around .95 and higher, SRMR values around .08 or lower,

and RMSEA values around .06 and lower indicate a relatively good fit of the model to the data.

Each IES-2 item was specified to load only on its latent first-order factor. Relationships between the four hypothesized latent factors were estimated, and these four latent factors were specified to load on a second-order intuitive eating factor. Given that there were similarly worded phrases used between certain IES-2 items (i.e., four items that begin with “I trust my body to tell me . . . ,” three items that begin with “I find myself eating when I am . . . ,” two items that begin with “I mostly eat foods that . . . ,” and two items that end with “. . . turning to food for comfort”), these items were expected to share method variance. Therefore, correlated errors between these similarly phrased items were estimated. This decision was based on recommendations from scholars (e.g., Brown, 2006; Kline, 2005) who have argued that it is both acceptable and preferable to correlate the errors between certain items that share method effects, such as similar phrasings within items.

This model provided a good fit to the data for the combined sample, CFI = .96, SRMR = .06, RMSEA = .05, 90% CI [.050, .057], $\chi^2(206, N = 1200) = 908.31, p < .001$. This model also provided a good fit to the data for women, CFI = .96, SRMR = .06, RMSEA = .05, 90% CI [.045, .055], $\chi^2(206, N = 680) = 558.90, p < .001$, and an acceptable to good fit to the data for men, CFI = .94, SRMR = .07, RMSEA = .06, 90% CI [.057, .068], $\chi^2(206, N = 520) = 627.38, p < .001$. Table 3 includes the

item-factor loadings as well as the first-order factor loadings on the second-order intuitive eating factor for the overall and sex-specific models. The factor structure obtained in Study 1, then, was confirmed in Study 2 for the overall sample as well as for women and men separately.

Tests of measurement invariance. Measurement invariance, determining whether the IES-2 was invariant among women and men, was tested at three different levels: (a) configural invariance (i.e., whether similar factors are measured in women and men), (b) factor loading invariance (i.e., whether the magnitude of factor loadings is the same across women and men), and (c) intercept invariance (i.e., whether the intercept of the regression relating each item to its factor is the same across women and men; Chen, 2007). Configural invariance is determined by CFI, SRMR, and RMSEA model fit indices. A chi-square difference (i.e., $\Delta\chi^2$) test allows a statistical comparison between nested models. However, this test is almost always large and statistically significant with complex models and large samples (as in the present study) and therefore is an impractical and unrealistic criterion on which to base evidence of invariance (e.g., Byrne & Stewart, 2006; Chen, Sousa, & West, 2005). For this reason, Chen (2007) recommended also exploring practical model fit changes: If $\Delta\text{CFI} \leq .010$, $\Delta\text{RMSEA} \leq .015$, and $\text{SRMR} \leq .030$ for tests of factor loading invariance, and $\Delta\text{CFI} \leq .010$, $\Delta\text{RMSEA} \leq .015$, and $\text{SRMR} \leq$

Table 3
Standardized Item and Factor Loadings From Confirmatory Factor Analyses of the Overall Sample, Women, and Men From Study 2

Factor and IES-2 item (see Appendix)	First-order			Second-order		
	Overall	Women	Men	Overall	Women	Men
Eating for Physical Rather Than Emotional Reasons (EPR)				.55	.55	.58
Item 2	.81	.79	.79			
Item 5	.82	.80	.82			
Item 10	.88	.87	.87			
Item 11	.87	.85	.88			
Item 12	.71	.75	.61			
Item 13	.69	.67	.68			
Item 14	.79	.79	.77			
Item 15	.72	.69	.73			
Unconditional Permission to Eat (UPE)				.66	.62	.63
Item 1	.72	.72	.72			
Item 3	.50	.48	.53			
Item 4	.66	.65	.66			
Item 9	.64	.64	.65			
Item 16	.69	.70	.65			
Item 17	.69	.71	.66			
Reliance on Hunger and Satiety Cues (RHSC)				.89	.88	.84
Item 6	.66	.64	.67			
Item 7	.59	.50	.66			
Item 8	.65	.61	.70			
Item 21	.69	.66	.70			
Item 22	.67	.67	.69			
Item 23	.77	.78	.79			
Body-Food Choice Congruence (B-FCC)				.74	.69	.71
Item 18	.72	.75	.71			
Item 19	.87	.86	.86			
Item 20	.79	.81	.74			

Note. $N = 1,200$; $n = 680$ women and $n = 520$ men. For the model based on the data for the overall sample (and for women and men, respectively), factor interrelationships were calculated. EPR was related to UPE, $r = .16$ ($r_s = .08$ for women and $.21$ for men); to RHSC, $r = .49$ ($r_s = .46$ for women and $.49$ for men); and to B-FCC, $r = .28$ ($r_s = .25$ for women and $.27$ for men). UPE was related to RHSC, $r = .34$ ($r_s = .24$ for women and $.42$ for men), and to B-FCC, $r = -.35$ ($r_s = -.30$ for women and $-.45$ for men). RHSC was related to B-FCC, $r = .30$ ($r_s = .31$ for women and $.27$ for men).

.010 for tests of intercept invariance, then measurement invariance is evidenced.

The configural invariance model fit the data well, CFI = .950, RMSEA = .056, SRMR = .064, $\chi^2(412, N = 1,200) = 1,186.278$, $p < .001$, suggesting that the IES-2 has four distinct dimensions that each load on a higher order factor for both college women and men. When testing for factor loading invariance, we constrained all factor loadings equally across women and men. This model also provided a good fit to the data, CFI = .949, RMSEA = .055, SRMR = .066, $\chi^2(431, N = 1,200) = 1,223.541$, $p < .001$, but differed significantly from the configural model, $\Delta\chi^2(19) = 37.26$, $p = .0074$. However, the fit indices changed very little ($\Delta CFI = .001$, $\Delta RMSEA = .001$, and $\Delta SRMR = .002$) and were well within Chen's (2007) recommendations for factor loading invariance, indicating that the factor score estimates were largely invariant and similar between women and men.

Nevertheless, we explored which factor parameters were statistically different for women and men. Following recommendations by Byrne and Stewart (2006), we performed additional item-level analyses to better understand the statistically significant chi-square difference and identify which items were more equivalent indicators of intuitive eating across sex. After applying a Bonferroni adjustment to control for Type I error ($p = .05/27 = .00185$), we found that only one factor parameter (i.e., for Item 12 in the Appendix) was statistically significant/noninvariant, $\Delta\chi^2(1) = 10.41$, $p = .00130$ (path for women = .747, path for men = .610); however, the fit indices were nearly identical ($\Delta CFI = .001$, $\Delta RMSEA = .000$, and $\Delta SRMR = .000$), suggesting that the overall impact of this deviation was trivial.

Last, intercept invariance was evaluated, and all item and first-order factor intercepts were constrained equally across women and men. This model fit the data well, CFI = .946, RMSEA = .056,

SRMR = .064, $\chi^2(450, N = 1,200) = 1,290.436$, $p < .001$, but differed significantly from the factor loading invariant model, $\Delta\chi^2(19) = 66.90$, $p < .001$. Again, though, the fit indices changed very little and were well within Chen's (2007) recommendations for intercept invariance ($\Delta CFI = .003$, $\Delta RMSEA = .001$, and $\Delta SRMR = .002$), indicating that the intercepts were largely invariant for women and men. An exploration of the 27 intercepts revealed that only two item intercepts were noninvariant: the Item 2 intercept, $\Delta\chi^2(1) = 10.36$, $p = .0013$ (intercept for women = 2.36, intercept for men = 2.57), and the Item 4 intercept, $\Delta\chi^2(1) = 17.69$, $p < .0001$ (intercept for women = 2.39, intercept for men = 2.71); Items 2 and 4 can be found in the Appendix. The fit indices for these two item intercepts (Item 2 intercept: $\Delta CFI = .001$, $\Delta RMSEA = .000$, and $\Delta SRMR = .001$; Item 4 intercept: $\Delta CFI = .001$, $\Delta RMSEA = .001$, and $\Delta SRMR = .001$) suggested that the overall impact of these different intercepts was negligible. For cases such as this, Karcher and Sass (2010) recommend that all scale items be retained rather than deleting the noninvariant items. As such, all 23 IES-2 items were retained.

Construct validity. The IES-2 total scale and subscales were expected to correlate (a) in a negative direction with eating disorder symptomatology, body surveillance, body shame, internalization of media appearance ideals, poor interoceptive awareness, and negative affect and (b) in a positive direction with body appreciation, self-esteem, positive affect, and satisfaction with life. Further, Cohen's (1992) criteria (i.e., $r_s \geq .50$ indicate a large/strong effect size; r_s around .30, a medium/moderate effect size; and r_s around .10, a slight/negligible effect size) were used to determine the strength of the IES-2's associations with the study variables. These correlations are presented in Table 4.

IES-2 total scores. Our hypotheses that IES-2 total scores would be related to the various eating- and body-related measures

Table 4
Means (SDs) and Intercorrelations of the Measured Variables in Study 2

Variable	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1. IES-2 total scale	—	.84*	.53*	.75*	.29*	-.59*	.53*	-.32*	-.60*	-.37*	-.43*	.41*	.30*	-.35*	.33*	-.31*
2. IES-2 EPR subscale	.82*	—	.17*	.43*	.26*	-.39*	.42*	-.25*	-.46*	-.32*	-.36*	.37*	.27*	-.36*	.33*	-.25*
3. IES-2 UPE subscale	.45*	.05	—	.35*	-.34*	-.64*	.21*	-.27*	-.50*	-.25*	-.22*	.11	-.05	-.09	.06	-.17*
4. IES-2 RHSC subscale	.68*	.34*	.24*	—	.24*	-.43*	.45*	-.21*	-.34*	-.25*	-.27*	.30*	.26*	-.24*	.19*	-.24*
5. IES-2 B-FCC subscale	.34*	.23*	-.23*	.23*	—	.08	.31*	-.04	.00	-.04	-.11	.25*	.40*	-.20*	.25*	-.15*
6. Disordered eating	-.50*	-.20*	-.68*	-.37*	.05	—	-.47*	.42*	.72*	.46*	.23*	-.28*	-.08	.30*	-.20*	.33*
7. Body appreciation	.52*	.32*	.33*	.40*	.29*	-.49*	—	-.49*	-.63*	-.53*	-.35*	.59*	.49*	-.40*	.39*	-.32*
8. Body surveillance	-.36*	-.25*	-.26*	-.24*	-.19*	.42*	-.50*	—	.56*	.63*	.08	-.32*	-.20*	.33*	-.18*	.04
9. Body shame	-.56*	-.35*	-.52*	-.39*	-.08	.67*	-.66*	.49*	—	.59*	.27*	-.46*	-.23*	.48*	-.26*	.26*
10. Internalization	-.37*	-.30*	-.30*	-.22*	-.13*	.51*	-.56*	.59*	.58*	—	.26*	-.36*	-.25*	.35*	-.18*	.07
11. PIA	-.35*	-.28*	-.17*	-.24*	-.13*	.29*	-.31*	.18*	.35*	.27*	—	-.37*	-.27*	.33*	-.17*	.12
12. Self-esteem	.41*	.26*	.27*	.30*	.22*	-.38*	.63*	-.27*	-.53*	-.35*	-.33*	—	.51*	-.55*	.53*	-.09
13. Positive affect	.26*	.18*	.00	.22*	.37*	-.10	.45*	-.20*	-.26*	-.22*	-.19*	.57*	—	-.26*	.40*	-.05
14. Negative affect	-.36*	-.31*	-.15*	-.21*	-.19*	.36*	-.43*	.30*	.45*	.39*	.31*	-.52*	-.27*	—	-.43*	.01
15. Life satisfaction	.33*	.28*	.12	.21*	.18*	-.24*	.49*	-.14*	-.36*	-.28*	-.18*	.66*	.56*	-.48*	—	-.23*
16. BMI	-.27*	-.22*	-.14*	-.17*	-.06	.13*	-.33*	.07	.31*	.05	.05	-.07	-.03	.02	-.22*	—
M, women	3.38	3.17	3.48	3.53	3.35	2.38	3.55	4.74	3.49	3.16	2.95	3.13	3.62	2.38	4.70	24.02
SD, women	0.51	0.90	0.76	0.61	0.79	0.69	0.78	1.21	1.32	1.11	0.90	0.54	0.61	0.69	1.42	5.67
M, men	3.67	3.69	3.65	3.68	3.48	1.98	3.82	4.04	2.96	3.09	3.30	3.23	3.72	2.12	4.73	25.38
SD, men	0.53	0.87	0.78	0.65	0.77	0.70	0.70	1.22	1.17	0.87	1.05	0.52	0.57	0.58	1.33	5.48
Possible range	1-5	1-5	1-5	1-5	1-5	1-6	1-5	1-7	1-7	1-5	1-6	1-4	1-5	1-5	1-7	

Note. $N = 1,200$ ($n = 680$ women, $n = 520$ men). Values for women are presented below the diagonal, whereas values for men are presented above the diagonal. IES-2 = Intuitive Eating Scale-2, EPR = Eating for Physical Rather Than Emotional Reasons, UPE = Unconditional Permission to Eat, RHSC = Reliance on Hunger and Satiety Cues, B-FCC = Body-Food Choice Congruence, PIA = poor interoceptive awareness, BMI = body mass index. * $p < .001$.

and psychological well-being indices were supported. Specifically, the IES-2 total score had strong negative associations with eating disorder symptomatology and body shame and strong positive associations with body appreciation. Further, the IES-2 total score had moderate to strong negative links to internalization of media appearance ideals and poor interoceptive awareness and moderate to strong positive links to self-esteem. Moderately sized positive links were observed from the IES-2 total score to positive affect and life satisfaction, and moderately sized negative links were noted from the IES-2 total score to body surveillance, negative affect, and BMI. All abovementioned trends were noted for both women and men.

Eating for Physical Rather Than Emotional Reasons (EPR) subscale scores. As expected, EPR was associated with the disordered eating, body-related, and well-being measures. Associations between EPR and eating disorder symptomatology were negative and slight to moderate for women and moderate to strong for men. Relations between EPR and body shame were negative and moderate to strong in size for women and men. Links from EPR to body appreciation and self-esteem were positive and moderate in strength for women and moderate to strong for men. A moderately sized positive link was noted from EPR to life satisfaction, and moderately sized negative links were demonstrated from EPR to body surveillance, internalization of media appearance ideals, poor interoceptive awareness, and negative affect; these trends were noted for both women and men. Last, EPR was moderately and positively related to positive affect and moderately and negatively related to BMI for women and men.

Unconditional Permission to Eat (UPE) subscale scores. Consistent with our hypotheses, UPE was related to the eating- and body-related measures; however, its associations with the well-being measures were less stable. For women and men, (a) links from UPE to eating disorder symptomatology and body shame were negative and strong in size, (b) the link from UPE to body appreciation was positive and moderate in size, (c) links from UPE to body surveillance and internalization of media appearance ideals were negative and moderate in size, and (d) links from UPE to poor interoceptive awareness and BMI were negative and slight to moderate in size. For women, UPE was positively associated with self-esteem to a moderate degree, positively associated with life satisfaction to a negligible degree, and inversely associated with negative affect to a slight degree; yet, UPE was not associated with positive affect. For men, UPE was associated with self-esteem to a negligible degree but not associated with positive affect, negative affect, or life satisfaction.

Reliance on Hunger and Satiety Cues (RHSC) subscale scores. As predicted, RHSC was related to all eating- and body-related and psychological well-being measures; the findings presented below apply to both women and men. RHSC was moderately to strongly related in a negative direction to eating disorder symptomatology and body shame and in a positive direction to body appreciation. Slight to moderate negative associations were noted between RHSC and body surveillance, internalization of media ideals, poor interoceptive awareness, and BMI. Findings revealed moderate positive associations between RHSC and self-esteem, slight to moderate positive associations from RHSC to positive affect and life satisfaction, and a slight to moderate inverse association between RHSC and negative affect.

Body-Food Choice Congruence (B-FCC) subscale scores. B-FCC was related to certain body-related measures and all psychological well-being indices, offering partial support for our hypothesis. Specifically, for men and women, B-FCC was moderately associated with body appreciation in a positive direction, moderately to strongly associated with positive affect in a positive direction, slightly to moderately associated with self-esteem and life satisfaction in a positive direction, and slightly to moderately associated with negative affect in an inverse direction. For women only, B-FCC was slightly to moderately related in a negative direction to body surveillance, poor interoceptive awareness, and internalization of media ideals. For men only, B-FCC was slightly related in an inverse direction with BMI. B-FCC was not associated with eating disorder symptomatology and body shame for women or men, with body surveillance, internalization of media ideals, and poor interoceptive awareness for men, or with BMI for women.

Incremental validity. Finally, we determined whether the IES-2 predicted psychological well-being (i.e., self-esteem, positive affect, negative affect, and life satisfaction) above and beyond the variance accounted for by eating disorder symptomatology. The EAT-26 total score (eating disorder symptomatology) was entered at Step 1 and the IES-2 total score was entered at Step 2 in the prediction of each of the four criteria, yielding four hierarchical multiple regression equations. The p level was adjusted to .013 (.05/4) to control for Type I error. A statistically significant increment in R^2 at Step 2 would indicate incremental validity evidence for the IES-2. Per Cohen (1992), R^2 values of .02 indicate a small effect size; of .15, a medium effect size; and of .35 and above, a large effect size.

These findings, presented in Table 5, support the incremental validity of the IES-2. For both women and men, IES-2 total scores predicted unique variance in self-esteem, positive affect, negative affect, and life satisfaction. All R^2 values at Step 2 were small to medium in effect size. Interestingly, eating disorder symptomatology no longer predicted self-esteem and negative affect for men and positive affect and life satisfaction for women at Step 2 when the IES-2 total score was entered into the equations, indicating that eating disorder symptomatology did not predict these dimensions of psychological functioning beyond its overlap with intuitive eating. Therefore, IES-2 total scores are distinct from low levels of eating disorder symptomatology, lending support to its incremental validity.

IES-2 subscales were entered at Step 2 of the hierarchical regressions (in lieu of the IES-2 total score) to determine which subscales uniquely contributed to the four indices of psychological functioning above and beyond the variance accounted for by eating disorder symptomatology at Step 1. EPR, RHSC, and B-FCC each predicted unique variance in self-esteem for both women ($\Delta R^2 = .087$) and men ($\Delta R^2 = .139$). RHSC and B-FCC each predicted unique variance in positive affect for women ($\Delta R^2 = .153$), and these two subscales as well as EPR each predicted unique variance in positive affect for men ($\Delta R^2 = .222$). EPR and B-FCC each predicted unique variance in negative affect for both women ($\Delta R^2 = .084$) and men ($\Delta R^2 = .108$). Last, EPR predicted unique variance in life satisfaction for women ($\Delta R^2 = .060$), whereas EPR and B-FCC predicted unique variance in life satisfaction for men ($\Delta R^2 = .119$). A table of the cumulative and incremental R^2 values with their companion F values, as well

Table 5
Incremental Variance in Psychological Functioning Accounted for by IES-2 Total Scores in Study 2

	Cum. R^2	ΔR^2	ΔF	β	t
Criterion: RSES (self-esteem)					
Women's overall $F(2, 671) = 87.05^*$					
Men's overall $F(2, 459) = 47.38^*$					
Step 1	.143/.063	.143/.063	111.54*/32.18*		
EAT-26 (disordered eating)				-.38/-.26	-10.56*/-5.67*
Step 2	.207/.171	.064/.106	53.75*/58.55*		
EAT-26 (disordered eating)				-.23/-.02	-5.91*/-0.30
IES-2 total scores				.29/.40	7.33*/7.65*
Criterion: PANAS-PA (positive affect)					
Women's overall $F(2, 671) = 23.99^*$					
Men's overall $F(2, 459) = 35.05^*$					
Step 1	.010/.002	.010/.002	6.82*/32.18*		
EAT-26 (disordered eating)				-.10/-.04	-2.61*/-0.87
Step 2	.067/.132	.057/.131	40.75*/69.23*		
EAT-26 (disordered eating)				.03/.23	0.83/4.19*
IES-2 total scores				.27/.45	6.38*/8.32*
Criterion: PANAS-NA (negative affect)					
Women's overall $F(2, 671) = 69.68^*$					
Men's overall $F(3, 459) = 19.39^*$					
Step 1	.124/.085	.124/.085	95.30*/42.77*		
EAT-26 (disordered eating)				.35/.29	9.76*/6.54*
Step 2	.172/.148	.048/.063	38.71*/33.99*		
EAT-26 (disordered eating)				.23/-.02	5.63*/1.99
IES-2 total scores				-.25/.40	-6.22*/-5.83*
Criterion: SWLS (life satisfaction)					
Women's overall $F(2, 671) = 13.07^*$					
Men's overall $F(2, 459) = 7.29^*$					
Step 1	.064/.015	.064/.015	2.01/32.18*		
EAT-26 (disordered eating)				-.25/-.12	-3.77*/-1.42
Step 2	.113/.102	.049/.086	11.24*/12.39*		
EAT-26 (disordered eating)				-.10/-.08	-1.21/-0.81
IES-2 total scores				.27/.36	3.35*/3.52*

Note. $N = 1,200$; $n = 680$ women, $n = 520$ men. F and Step 2 values that are to the left of the diagonal were derived from the analysis of women's data. F and Step 2 values that are to the right of the diagonal were derived from the analysis of men's data. Regressions using IES-2 subscales in lieu of total scores are available from the first author. Cum. = Cumulative. EAT-26 = Eating Attitudes Test-26, IES-2 = Intuitive Eating Scale-2, RSES = Rosenberg Self-Esteem Scale, PANAS = Positive and Negative Affect Scale (PA = positive affect, NA = negative affect), SWLS = Satisfaction With Life Scale.

* $p < .013$ (i.e., .05/4).

as standardized betas and their companion t values, from analyzing the subscales at Step 2 can be obtained from Tracy Tylka. Overall, these findings indicate that three of the four IES-2 subscales are distinct from low levels of eating disorder symptomatology, upholding their incremental validity, whereas the remaining IES-2 subscale, UPE, overlapped substantially with low levels of eating disorder symptomatology.

Study 3

Because intuitive eating is associated with psychological well-being, women and men may want to project the impression onto others that they eat intuitively. However, to support its discriminant validity, the IES-2 should not overlap substantially with individuals' response style. Consequently, we hypothesized that the IES-2 total and subscale scores would not be more than negligibly related to socially desirable responding.

Method

Participants. Data were analyzed data from 522 college students (238 women and 284 men) attending the same regional campus that was used in the previous two studies. Participants ranged in age from 18 to 56 years ($M = 20.29$ years, $SD = 4.82$). They identified as White (78.4%), African American (5.4%), Asian American (4.8%), Latina/Latino (1.0%), Native American (0.4%), or multiracial (6.3%); 3.6% endorsed "other" and one participant did not provide an ethnic identification. They were first-year college students (77.8%), sophomores (17.0%), juniors (3.6%), or seniors (1.5%).

Measures.

Intuitive eating. The 23-item IES-2, described in Study 1, was used in Study 3. In this study, Cronbach's alphas among women and men, respectively, were .85 and .88 for the IES-2 total score, .92 and .92 for EPR, .77 and .82 for UPE, .85 and .87 for RHSC, and .87 and .84 for B-FCC.

Social desirability. The 33-item Marlowe-Crowne Social Desirability Scale (MCSDS; Crowne & Marlowe, 1960) represents socially approved opinions or behaviors to which most people cannot truthfully claim to adhere at all times (e.g., “I never hesitate to go out of my way to help someone in trouble”). Each item is scored using a dichotomous *true/false* response scale. Participants receive a point when they respond *true* to a socially desirable item and *false* to a socially undesirable item. Points are summed, and higher scores indicate greater socially desirable responding. Among college students, estimates of its internal consistency reliability (i.e., Kuder-Richardson reliability index, or $KR20$, = .77–.88) and 2-week test–retest reliability ($r = .89$) are adequate, and it is consistently related to other measures of social desirability ($r_s = .74$ –.84; Blake, Valdiserri, Neuendorf, & Nemeth, 2006). For the current study, Kuder-Richardson reliability indexes of .77 and .75 were obtained for women and men, respectively.

Procedure. After receiving IRB approval, we recruited participants from introductory and upper-level psychology classes. Data were collected over a 9-month period, which began 3 months after data collection ceased for Study 2. After interested students signed up for this study on the psychology department’s research management website and provided their consent, they were taken to a webpage that hosted the survey. This study was described as an investigation of eating habits and personality, and the two measures were counterbalanced to control for order effects.

Those who did not complete the survey ($n = 43$), failed at least one of the three validity questions ($n = 37$), had excessive amounts of missing data (not completing at least 80% of each measure, $n = 13$), or participated in Study 2 ($n = 7$) were deleted. These screening procedures reduced the data set from 622 to 522 participants.

Results and Discussion

Preliminary analyses. A total of 13.83% of participants had at least one missing data point. When inspecting individual IES-2 and MCSDS items, the count for missing data points ranged from 0 to 0.8%, and the amount and pattern of missingness were completely random per Little’s MCAR analysis, $\chi^2(2725) = 2824.09$, $p = .091$. Thus, MI was used to estimate missing data values by creating five data sets using available data, and subsequent analyses on the pooled results were conducted. All scale and subscale scores had acceptable levels of skewness and kurtosis for women and men. IES-2 total scale and subscale means and standard deviations are shown in Table 2. The Bonferroni adjustment was used ($.05/5 = .01$). Women and men differed in their IES-2 total scores (medium effect size) and scores on the EPR (medium-to-large effect size) and UPE (small effect size) subscales. For the MCSDS, means were 16.40 ($SD = 5.17$) for women and 15.59 ($SD = 5.05$) for men.

Discriminant validity. As hypothesized, IES-2 scores were either unrelated or negligibly related (i.e., r_s around .10; Cohen, 1988) to social desirability for women and men. Specifically, social desirability correlated (r_s) .06 ($p = .396$) among women and .10 ($p = .096$) among men with the IES-2 total score; .06 ($p = .351$) among women and .13 ($p = .033$) among men with EPR; $-.02$ ($p = .823$) among women and $-.09$ ($p = .117$) among men with UPE; .01 ($p = .921$) among women and .13 ($p = .035$) among men with RHSC; and .11 ($p = .103$) among women and .12 ($p =$

.055) among men with B-FCC. Therefore, supporting the IES-2’s discriminant validity, social desirability did not overlap substantially (i.e., $\leq 1.7\%$) with participants’ response style.

General Discussion

In this collection of studies, we report on the development and psychometric evaluation of the IES-2 with three samples. Overall, the IES-2s factor structure and the reliability (i.e., estimated internal consistency and 3-week test–retest) and validity (i.e., construct, discriminant, and incremental) of its scores were upheld for undergraduate women and men. The final 23-item IES-2 included 11 of Tylka’s (2006) original items and 12 newly developed items that, as hypothesized, loaded on four first-order factors (EPR, UPE, RHSC, and B-FCC). These first-order factors, in turn, loaded on a higher order intuitive eating factor. Also, the IES-2 largely demonstrated measurement invariance at the scale, factor loading, and intercept levels. According to the practical perspective (Chen, 2007), the statistically significant sex differences noted in the factor loading of Item 12 and in the intercepts of Items 2 and 4 were indeed ignorable (i.e., due to the large sample and the complexity of the model, the sensitivity of the chi-square difference tests was heightened). This finding indicates that women and men can be compared on the IES-2. Because an investigation of the original IES with men has not been published to date, the present study’s findings extend the applicability of the IES-2 to men and suggest that the IES-2 can be used and interpreted confidently among and between young adult women and men.

Importantly, the IES-2 improved upon the three limitations noted for the original IES. B-FCC, a dimension of intuitive eating identified by Tribble and Resch (2003), was found to be a distinct factor within the IES-2. Items assessing this factor were not included in Tylka’s (2006) original IES. Furthermore, the IES-2 contains fewer reverse-scored items ($n = 7$) and more positively scored items ($n = 16$) than its original version ($n_s = 13$ and 8, respectively). This change in item distribution facilitates the computation of subscale and total scores and contributes to construct integrity (i.e., the presence of intuitive eating is assessed more so than the absence of dieting and emotional eating). Yet, the original IES and IES-2 total and subscale scores were highly correlated (overlap ranging between 74% and 90%), indicating that they tap into the same construct. Thus, researchers can confidently use the IES-2 as a measure of intuitive eating. Also, for the IES-2, the RHSC subscale consistently yielded internal consistency reliability estimates of .85 to .89 for college women and men, an improvement from the low to mid .70s reported for its original version (Tylka, 2006), even though the number of items remained the same. Overall, the IES-2 contains two more items than its original version, which will not substantially increase the time commitment involved for its administration to research participants or clients. These features make the IES-2 a more attractive option than the original IES.

To understand the adaptive properties of intuitive eating and its components, it is important to note how IES-2 scores are linked with other eating, body-related, and well-being measures. For both women and men, higher levels of overall intuitive eating, EPR, and RHSC corresponded to lower eating disorder symptomatology, lower negative and constrictive body attitudes, greater interoceptive awareness, higher body appreciation, and higher psychologi-

cal well-being. UPE was similarly linked to most of these variables for women and men, with the exception being psychological well-being. In particular, unlike the other dimensions of intuitive eating, UPE was not related to life satisfaction, positive affect, or negative affect and did not account for unique variance in psychological well-being above and beyond eating disorder symptomatology. Conversely, B-FCC was positively related to psychological well-being, even after controlling for eating disorder symptomatology. However, this subscale was only associated with certain body-related variables for women (i.e., body appreciation, body surveillance, internalization of media ideals, and poor interoceptive awareness) and only one for men (i.e., body appreciation); it was not associated with eating disorder symptomatology or body shame for either women or men.

Curiously, UPE was inversely related to B-FCC to a small to moderate degree for both women and men. Although both subscales loaded positively on the higher order intuitive eating factor, their negative interrelationship is understandable. Specifically, those who consistently allow themselves to eat a wide variety of foods and foods they are craving (UPE) may not always limit themselves by choosing foods that will give their body energy and stamina (B-FCC). In their *Intuitive Eating* text, Tribole and Resch (2003) suggested that individuals who eat intuitively balance their eating between these different constructs. If they are experiencing an intense craving for a certain food, they will allow themselves to have it without judgment or guilt, but if they are left to choose between a healthy and an unhealthy food in the absence of a craving, they tend to choose the healthy food because it will nourish their body. The positive links from body appreciation to UPE and B-FCC uncovered in this study support that both components of intuitive eating are linked to listening to and appreciating the body.

It is also noteworthy that IES-2 total and subscale scores were either negatively related or unrelated to BMI among women and men. Individuals often report being reluctant to eat intuitively and instead following a dieting plan because they fear they will gain weight if they listen to their bodies (Tribole & Resch, 2003). While it cannot be inferred from these correlational data that people who transition to intuitive eating lose weight, the data do suggest that listening to the body's hunger and satiety signals and eating in accordance with these signals is linked, at least slightly, to lower BMI.

Whereas the factor structure and relationships with the eating and body-related measures and indices of psychological well-being were similar between women and men, sex differences were uncovered in the levels of the IES-2 total scale and subscale scores. A moderate degree of difference was noted between women and men for IES-2 total scores and EPR subscale scores, with men consistently reporting higher intuitive eating scores than women. However, sex differences were less pronounced between the other subscales, with very small to small degrees of difference documented between women and men. In Western cultures, men's bodies are scrutinized less than women's bodies, and women are encouraged to not trust their body's internal signals but instead to try to diet to achieve the thin ideal body frame promulgated for women in the media (Fredrickson & Roberts, 1997). Men do report pressure to be lean (Tylka, 2011), yet these pressures may not be as severe as the pressures women receive to be thin (Thompson et al., 1999). Given that appearance-related pressures can disrupt

awareness of internal cues (Fredrickson & Roberts, 1997), gender differences in cultural pressures may be one potential explanation for the sex differences in intuitive eating revealed in this study.

Limitations and Directions for Future Research

Limitations of the present study and the IES-2 should be noted. First, participants were mostly White college students, all from the same geographic location. This relative homogeneity of the demographic samples limits the generalizability of the results from the current evaluation of the IES-2 to different populations. Thus, future studies should examine whether the psychometric properties of the IES-2 are upheld in samples that are more diverse in terms of ethnicity, cultural background, age, socioeconomic status (SES), and sexual orientation. Items may need to be reworded for children, preadolescents, and adolescents, in particular. Due to food scarcity concerns being more prevalent in low-income households (Melchior et al., 2009), individuals within these households are likely to eat when food is available rather than have the privilege to use their hunger and satiety cues to guide their eating; therefore, the intuitive eating construct may not be appropriate or may operate differently for lower income individuals. Future research should explore the IES-2's connection to income and food security.

Second, participants also self-selected to participate in these studies, which may have led to biases in the sample such that only those interested in and curious about eating habits participated in this study. Further, in each study, many participants with incomplete data or who did not respond correctly to the validity questions were deleted. Therefore, our sample may represent the intuitive eating habits of only conscientious students.

Third, data from the current study are exclusively self-report, which rely entirely on individuals accurately portraying their current levels of functioning, and correlational. Because no direct observational data of participants' eating attitudes and behaviors were collected, we are unable to determine if participants were accurately reporting their eating behaviors. Likewise, while participants may have been reporting their true perceptions of their eating attitudes and behaviors, their perceptions may not be an accurate portrayal of their actual behaviors. Stice, Fisher, and Lowe (2004) found that measures designed to assess dietary restraint were either unrelated or negligibly related to short-term dietary restriction as measured in a controlled lab. Future studies should examine how closely scores on the IES-2 are related to observations of actual eating behaviors in laboratory and naturalistic settings. Using longitudinal designs to investigate the development and maintenance of intuitive eating, along with its risk and protective factors, also is necessary.

Fourth, although the B-FCC subscale was positively related to several indices of psychological well-being, additional research needs to examine its psychometric properties. One important avenue for research is to differentiate this subscale from measures of orthorexia nervosa, a maniacal, inflexible, and unhealthy fixation with what is considered healthy eating (Donini, Marsili, Graziani, Imbriale, & Cannella, 2005). To uphold the discriminant validity of the B-FCC, it should not be strongly related to a diagnostic instrument for orthorexia nervosa (i.e., the ORTO 15; Donini et al., 2005) if it indeed assesses true "gentle nutrition" and not a rigid reliance and focus on health foods.

Further, the IES-2 could prove to be a useful clinical tool, and its evaluation in counseling and psychoeducational settings is needed. It could be administered in college counseling centers to students who have a goal to adopt more positive attitudes and behaviors related to food and eating. If the psychometric properties of the IES-2 are supported in community samples, it could be helpful for clinicians who run intuitive eating or anti-dieting workshops. These clinicians could use the IES-2 at various intervals to gauge participants' progress and rate of change throughout the workshop. Also, if the IES-2 yields reliable and valid scores among those with eating disorders, it could be incorporated within eating disorder treatment settings. Given that treatment goals for individuals with eating disorders include an increase in adaptive eating alongside a decrease in maladaptive eating, clinicians could administer the IES-2 at various intervals throughout therapy. Increases in IES-2 scores, then, would serve as a measure of therapeutic progress. Further, because the IES-2 predicts psychological well-being above and beyond eating disorder symptoms, it may be better suited than measures of eating disorder symptomatology when assessing clients' risk for relapse. The IES-2 is easy to administer and score and takes only a few minutes to complete, which are attractive features that further support its inclusion within appropriate clinical settings.

References

- Anestis, M. D., Holm-Denoma, J. M., Gordon, K. H., Schmidt, N. B., & Joiner, T. E. (2008). The role of anxiety sensitivity in eating pathology. *Cognitive Therapy and Research, 32*, 370–385. doi:10.1007/s10608-006-9085-y
- Augustus-Horvath, C. L., & Tylka, T. L. (2011). The acceptance model of intuitive eating: A comparison of women in emerging adulthood, early adulthood, and middle adulthood. *Journal of Counseling Psychology, 58*, 110–125. doi:10.1037/a0022129
- Avalos, L. C., & Tylka, T. L. (2006). Exploring a model of intuitive eating with college women. *Journal of Counseling Psychology, 53*, 486–497. doi:10.1037/0022-0167.53.4.486
- Avalos, L. C., Tylka, T. L., & Wood-Barcalow, N. (2005). The Body Appreciation Scale: Development and psychometric evaluation. *Body Image, 2*, 285–297. doi:10.1016/j.bodyim.2005.06.002
- Bacon, L., Keim, N. L., Van Loan, M. D., Derricote, M., Gale, B., Kazaks, A., & Stern, J. S. (2002). Evaluating a 'non-diet' wellness intervention for improvement of metabolic fitness, psychological well-being and eating and activity behaviors. *International Journal of Obesity, 26*, 854–865. doi:10.1038/sj.ijo.0802012
- Bacon, L., Stern, J. S., Van Loan, M. D., & Keim, N. (2005). Size acceptance and intuitive eating improve health for obese, female chronic dieters. *Journal of the American Dietetic Association, 105*, 929–936. doi:10.1016/j.jada.2005.03.011
- Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin, 107*, 238–246. doi:10.1037/0033-2909.107.2.238
- Bergeron, D., & Tylka, T. L. (2007). Support for the uniqueness of body dissatisfaction from drive for muscularity among men. *Body Image, 4*, 288–295. doi:10.1016/j.bodyim.2007.05.002
- Birch, L. L., & Fisher, J. O. (2000). Mothers' child-feeding practices influence daughters' eating and weight. *American Journal of Clinical Nutrition, 71*, 1054–1061.
- Birch, L. L., Fisher, J. O., & Davison, K. K. (2003). Learning to overeat: Maternal use of restrictive feeding practices promotes girls' eating in the absence of hunger. *American Journal of Clinical Nutrition, 78*, 215–220.
- Blake, B. F., Valdiserri, J., Neuendorf, K. A., & Nemeth, J. (2006). Validity of the SDS-17 measure of social desirability in the American context. *Personality and Individual Differences, 40*, 1625–1636. doi:10.1016/j.paid.2005.12.007
- Brown, T. A. (2006). *Confirmatory factor analysis for applied research*. New York, NY: Guilford Press.
- Byrne, B. M., & Stewart, S. M. (2006). The MACS approach to testing for multigroup invariance of a second-order factor structure: A walk through the process. *Structural Equation Modeling, 13*, 287–321. doi:10.1207/s15328007sem1302_7
- Chen, F. F. (2007). Sensitivity of goodness of fit indices to lack of measurement invariance. *Structural Equation Modeling, 14*, 464–504. doi:10.1080/10705510701301834
- Chen, F. F., Sousa, K. H., & West, S. G. (2005). Testing measurement invariance of second-order factor models. *Structural Equation Modeling, 12*, 471–492. doi:10.1207/s15328007sem1203_7
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). New York, NY: Erlbaum.
- Cohen, J. (1992). A power primer. *Psychological Bulletin, 112*, 155–159. doi:10.1037/0033-2909.112.1.155
- Crowne, D. P., & Marlowe, D. (1960). A new scale of social desirability independent of psychopathology. *Journal of Consulting Psychology, 24*, 349–354. doi:10.1037/h0047358
- Diener, E., Emmons, R. A., Larsen, R. J., & Griffin, S. (1985). The Satisfaction With Life Scale. *Journal of Personality Assessment, 49*, 71–75. doi:10.1207/s15327752jpa4901_13
- Dittmann, K. A., & Freedman, M. R. (2009). Body awareness, eating attitudes, and spiritual beliefs of women practicing yoga. *Eating Disorders: The Journal of Treatment & Prevention, 17*, 273–292. doi:10.1080/10640260902991111
- Donini, L. M., Marsili, D., Graziani, M. P., Imbriale, M., & Cannella, C. (2005). Orthorexia nervosa: Validation of a diagnosis questionnaire. *Eating and Weight Disorders, 10*, e28–e32.
- Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods, 4*, 272–299. doi:10.1037/1082-989X.4.3.272
- Fredrickson, B. L., & Roberts, T.-A. (1997). Objectification theory: Toward understanding women's lived experiences and mental health risks. *Psychology of Women Quarterly, 21*, 173–206. doi:10.1111/j.1471-6402.1997.tb00108.x
- Garner, D. M. (1991). *Eating Disorder Inventory-2 manual*. Odessa, FL: Psychological Assessment Resources.
- Garner, D. M., Olmsted, M. P., Bohr, Y., & Garfinkel, P. E. (1982). The Eating Attitudes Test: Psychometric features and clinical correlates. *Psychological Medicine, 12*, 871–878. doi:10.1017/S0033291700049163
- Gelso, C., & Fretz, B. (2001). *Counseling psychology* (2nd ed.). Fort Worth, TX: Harcourt.
- Heinberg, L. J., Thompson, J. K., & Stormer, S. (1995). Development and validation of the Sociocultural Attitudes Towards Appearance Questionnaire. *International Journal of Eating Disorders, 17*, 81–89. doi:10.1002/1098-108X(199501)17:1<81::AID-EAT2260170111>3.0.CO;2-Y
- Hu, L., & Bentler, P. (1999). Cutoff criteria for fit indices in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*, 1–55. doi:10.1080/10705519909540118
- Karcher, M. J., & Sass, D. (2010). A multicultural assessment of adolescent connectedness: Testing measurement invariance across gender and ethnicity. *Journal of Counseling Psychology, 57*, 274–289. doi:10.1037/a0019357
- Kline, R. B. (2005). *Principles and practices of structural equation modeling* (2nd ed.). New York, NY: Guilford Press.
- Koenders, P. G., & van Strien, T. (2011). Emotional eating, rather than lifestyle behavior, drives weight gain in a prospective study in 1,562 employees. *Journal of Occupational and Environmental Medicine, 53*, 1287–1293. doi:10.1097/JOM.0b013e31823078a2

- Kroon Van Diest, A. M., & Tylka, T. L. (2010). The Caregiver Eating Messages Scale: Development and psychometric investigation. *Body Image, 7*, 317–326. doi:10.1016/j.bodyim.2010.06.002
- Mazzeo, S. E. (1999). Modification of an existing measure of body image preoccupation and its relationship to disordered eating in female college students. *Journal of Counseling Psychology, 46*, 42–50. doi:10.1037/0022-0167.46.1.42
- McKinley, N. M., & Hyde, J. S. (1996). The Objectified Body Consciousness Scale: Development and validation. *Psychology of Women Quarterly, 20*, 181–215. doi:10.1111/j.1471-6402.1996.tb00467.x
- Melchior, M., Caspi, A., Howard, L. M., Ambler, A. P., Bolton, H., Mountain, N., & Moffitt, T. E. (2009). Mental health context of food insecurity: A representative cohort of families with young children. *Pediatrics, 124*, e564–e572. doi:10.1542/peds.2009-0583
- Muthén, L. K., & Muthén, B. O. (1998–2011). *Mplus user's guide* (6th ed.). Los Angeles, CA: Author.
- Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). New York, NY: McGraw-Hill.
- Oh, K. H., Wiseman, M. C., Hendrickson, J., Phillips, J. C., & Hayden, E. W. (2012). Testing the acceptance model of intuitive eating with college women athletes. *Psychology of Women Quarterly, 36*, 88–98. doi:10.1177/0361684311433282
- Parent, M. C., & Moradi, B. (2011). His biceps become him: A test of objectification theory's application to drive for muscularity and propensity for steroid use in college men. *Journal of Counseling Psychology, 58*, 246–256. doi:10.1037/a0021398
- Polivy, J., & Herman, C. P. (1999). Distress and eating: Why do dieters overeat? *International Journal of Eating Disorders, 26*, 153–164. doi:10.1002/(SICI)1098-108X(199909)26:2<153::AID-EAT4>3.0.CO;2-R
- Rogers, C. R. (1961). *On becoming a person*. Boston, MA: Houghton Mifflin.
- Rosenberg, M. (1965). *Society and the adolescent self-image*. Princeton, NJ: Princeton University Press.
- Shouse, S. H., & Nilsson, J. (2011). Self-silencing, emotional awareness, and eating behaviors in college women. *Psychology of Women Quarterly, 35*, 451–457. doi:10.1177/0361684310388785
- Stice, E., Fisher, M., & Lowe, M. R. (2004). Are dietary restraint scales valid measures of acute dietary restriction? Unobtrusive observational data suggest not. *Psychological Assessment, 16*, 51–59. doi:10.1037/1040-3590.16.1.51
- Studenmund, A. H. (2006). *Using econometrics: A practical guide*. New York, NY: Pearson.
- Swami, V., Stieger, S., Haubner, T., & Voracek, M. (2008). German translation and psychometric evaluation of the Body Appreciation Scale. *Body Image, 5*, 122–127. doi:10.1016/j.bodyim.2007.10.002
- Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics* (5th ed.). Boston, MA: Allyn & Bacon.
- Thompson, J. K., Heinberg, L. J., Altabe, M., & Tantleff-Dunn, S. (1999). *Exacting beauty: Theory, assessment, and treatment of body image disturbance*. Washington, DC: American Psychological Association. doi:10.1037/10312-000
- Tribole, E., & Resch, E. (1995). *Intuitive eating*. New York, NY: St. Martin's Griffin.
- Tribole, E., & Resch, E. (2003). *Intuitive eating* (2nd ed.). New York, NY: St. Martin's Griffin.
- Tylka, T. L. (2006). Development and psychometric evaluation of a measure of intuitive eating. *Journal of Counseling Psychology, 53*, 226–240. doi:10.1037/0022-0167.53.2.226
- Tylka, T. L. (2011). Refinement of the tripartite influence model for men: Dual body image pathways to body change behaviors. *Body Image, 8*, 199–207. doi:10.1016/j.bodyim.2011.04.008
- Tylka, T. L., Bergeron, D., & Schwartz, J. P. (2005). Development and psychometric evaluation of the Male Body Attitudes Scale (MBAS). *Body Image, 2*, 161–175. doi:10.1016/j.bodyim.2005.03.001
- Tylka, T. L., & Subich, L. M. (2004). Examining a multidimensional model of eating disorder symptomatology among college women. *Journal of Counseling Psychology, 51*, 314–328. doi:10.1037/0022-0167.51.3.314
- Tylka, T. L., & Wilcox, J. A. (2006). Are intuitive eating and eating disorder symptomatology opposite poles of the same construct? *Journal of Counseling Psychology, 53*, 474–485. doi:10.1037/0022-0167.53.4.474
- Watson, D., & Clark, L. A. (1994). *The PANAS-X: Manual for the Positive and Negative Affect Schedule—Expanded Form*. Unpublished manuscript, University of Iowa.
- Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and evaluation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology, 54*, 1063–1070. doi:10.1037/0022-3514.54.6.1063
- Wear, R. W., & Pratz, O. (1987). Test–retest reliability for the Eating Disorder Inventory. *International Journal of Eating Disorders, 6*, 767–769. doi:10.1002/1098-108X(198711)6:6<767::AID-EAT2260060611>3.0.CO;2-V
- Worthington, R. L., & Whittaker, T. A. (2006). Scale development research: A content analysis and recommendations for best practices. *The Counseling Psychologist, 34*, 806–838. doi:10.1177/0011000006288127

Appendix

Intuitive Eating Scale-2 (IES-2; Final Version)

Note that permission to use this measure is not required. However, we do request that you notify us via e-mail (tylka.2@osu.edu) if you use the Intuitive Eating Scale-2 in your research. Please seek permission if any item is modified.

For each item, the following response scale should be used: 1 = *strongly disagree*, 2 = *disagree*, 3 = *neutral*, 4 = *agree*, 5 = *strongly agree*.

Directions for Participants

For each item, please circle the answer that best characterizes your attitudes or behaviors. [Note to experimenter: use “check” in lieu of “circle” if survey is online.]

1. I try to avoid certain foods high in fat, carbohydrates, or calories.
2. I find myself eating when I'm feeling emotional (e.g., anxious, depressed, sad), even when I'm not physically hungry.
3. If I am craving a certain food, I allow myself to have it.
4. I get mad at myself for eating something unhealthy.
5. I find myself eating when I am lonely, even when I'm not physically hungry.
6. I trust my body to tell me when to eat.
7. I trust my body to tell me what to eat.
8. I trust my body to tell me how much to eat.
9. I have forbidden foods that I don't allow myself to eat.
10. I use food to help me soothe my negative emotions.
11. I find myself eating when I am stressed out, even when I'm not physically hungry.
12. I am able to cope with my negative emotions (e.g., anxiety, sadness) without turning to food for comfort.
13. When I am bored, I do NOT eat just for something to do.
14. When I am lonely, I do NOT turn to food for comfort.
15. I find other ways to cope with stress and anxiety than by eating.
16. I allow myself to eat what food I desire at the moment.
17. I do NOT follow eating rules or dieting plans that dictate what, when, and/or how much to eat.
18. Most of the time, I desire to eat nutritious foods.
19. I mostly eat foods that make my body perform efficiently (well).
20. I mostly eat foods that give my body energy and stamina.
21. I rely on my hunger signals to tell me when to eat.
22. I rely on my fullness (satiety) signals to tell me when to stop eating.
23. I trust my body to tell me when to stop eating.

Scoring Procedure

1. Reverse score Items 1, 2, 4, 5, 9, 10, and 11.
2. Total IES-2 scale score: Add together all items and divide by 23 to create an average score.
3. Unconditional Permission to Eat subscale: Add together Items 1, 3, 4, 9, 16, and 17; divide by 6 to create an average score.
4. Eating for Physical Rather Than Emotional Reasons subscale: Add together Items 2, 5, 10, 11, 12, 13, 14, and 15; divide by 8 to create an average score.
5. Reliance on Hunger and Satiety Cues subscale: Add together Items 6, 7, 8, 21, 22, and 23; divide by 6 to create an average score.
6. Body-Food Choice Congruence subscale: Add together Items 18, 19, and 20; divide by 3 to create an average score.

Received May 13, 2012

Revision received October 8, 2012

Accepted October 11, 2012 ■