Body Image 9 (2012) 476–487

Contents lists available at SciVerse ScienceDirect

Body Image

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

Development and validation of the Body, Eating, and Exercise Comparison Orientation Measure (BEECOM) among college women☆

Ellen E. Fitzsimmons-Craft, Anna M. Bardone-Cone*, Megan B. Harney

University of North Carolina at Chapel Hill, Department of Psychology, CB#3270-Davie Hall, Chapel Hill, NC 27599, United States

ARTICLE INFO

Article history:
Received 25 February 2012
Received in revised form 20 July 2012
Accepted 20 July 2012

Keywords:
Social comparison
Body dissatisfaction
Eating disorders
Measure development
Factor analysis

ABSTRACT

We constructed and validated a measure of comparison dimensions associated with eating pathology, namely, the Body, Eating, and Exercise Comparison Orientation Measure (BEECOM). Participants were 441 undergraduate women. In Study 1, items were generated and refined via exploratory factor analysis, yielding three interpretable factors (i.e., Body, Eating, and Exercise Comparison Orientation). Confirmatory factor analysis was then used to confirm the three-factor structure of the BEECOM and to investigate the potential presence of a higher-order factor. Given that the lower-order factors loaded strongly onto a higher-order factor, it is appropriate to use a total BEECOM score, in addition to subscale scores. Further, the BEECOM's scores yielded evidence of internal consistency and construct validity in this sample. Study 2 demonstrated two-week test–retest reliability of the BEECOM among college women. Overall, the BEECOM demonstrated good psychometric properties and may be useful for more comprehensively assessing eating disorder-related social comparison behavior.

Social comparison theory holds that humans have a natural drive to assess their progress and standing in life (Festinger, 1954). In order to fulfill this need, they seek out standards against which to compare themselves. Indeed, social comparison behavior is a common social psychological process that has been described as a "core element of human conduct and experience" (Suls, Martin, & Wheeler, 2002, p. 159) that pervades nearly all life domains (Gibbons & Buunk, 1999; Gilbert, Price, & Allan, 1995). When objective standards are not available for comparison purposes, and even oftentimes when they are (Klein, 1997), individuals look to their social environments for available others against which to compare themselves (e.g., Corning, Krumm, & Smitham, 2006). In fact, in many cases, individuals tend to disregard more generic comparison targets, as they are typically viewed as rather dissimilar and irrelevant (e.g., Heinberg & Thompson, 1992; Strahan, Wilson, Cressman, & Buote, 2006). In order to make an accurate comparison, a person very close to one's own abilities and opinions will be chosen (i.e., a peer; Lin & Kulik, 2002).

One very popular instance of social comparison is the assessment of one's own body as compared to friends and other peers. Indeed, there is ample evidence to suggest that women engage in frequent comparisons with peers in order to gain an understanding of their weight/shape status relative to others (e.g., Leahey, Crowther, & Mickelson, 2007; Striegel-Moore, Silberstein, & Rodin, 1986), and research has suggested that social comparison with peers may be one pathway through which internalized pressures for thinness may translate into body dissatisfaction and disordered eating (Dittmar, 2005; Dittmar & Howard, 2004; Fitzsimmons-Craft, Harney, Koehler, Danzi, Riddell, & Bardone-Cone, 2012; Leahey et al., 2007; Wood, 1996). It may be that via social comparison, individuals come to know that there is a discrepancy between their ideal and actual selves. Research has indicated that the tendency to engage in social comparison behavior with peers (i.e., both generally and specific to appearance) is associated with body dissatisfaction and disordered eating (Thompson, Heinberg, & Tantleff, 1991; Trottier, Polivy, & Herman, 2007). This has been confirmed by both experimental (e.g., Cash, Cash, & Butters, 1983; Krones, Stice, Batres, & Orjada, 2005; Lin & Kulik, 2002) and correlational work (e.g., Bamford & Halliwell, 2009; Faith, Leone, & Allison, 1997; Gilbert & Meyer, 2003; Hildebrandt, Shiovitz, Alfano, & Greif, 2008; Morrison, Kalin, & Morrison, 2004; Stormer & Thompson, 1996). Further, Myers and Crowther's (2009) meta-analytic results showed large, significant effect sizes for the relation between social comparison and body dissatisfaction.

Most of the work cited above has used one of a handful of brief instruments that have been developed for assessing social comparison orientation. For instance, the Iowa-Netherlands Comparison Orientation Measure (INCOM) has been validated as a measure of general social comparison orientation (Gibbons & Buunk, 1999). Using this measure, researchers found that differences in the general tendency to socially compare were associated with the

☆ We would like to thank Kerry S. O'Brien, Gregory T. Smith, and J. Kevin Thompson for reviewing the content of the BEECOM.

* Corresponding author. Tel.: +1 919 962 5989; fax: +1 919 962 2537.
E-mail addresses: fitzsimmonscraft@unc.edu (E.E. Fitzsimmons-Craft), bardonecone@unc.edu (A.M. Bardone-Cone), harney@unc.edu (M.B. Harney).

1740-1445/$ – see front matter © 2012 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.bodyim.2012.07.007

presence of eating disorder symptomatology (Corning et al., 2006). Several measures assessing more specific aspects of comparison (i.e., related to the body) have been created, as well. For instance, the Physical Appearance Comparison Scale (PACS) is a very brief (5-item) measure, which assesses the degree to which individuals tend to compare their appearance with others (e.g., "The best way for people to know if they are overweight or underweight is to compare their figure to the figure of others;" Thompson et al., 1991). High scores on this measure have been found to be highly correlated with both body dissatisfaction and eating disturbance (Thompson et al., 1991). Additionally, the Body Image Comparison Scale (BICS) is a 5-item scale, which assesses the frequency with which respondents engage in specific body- and appearance-related social comparison behaviors (e.g., "In social situations, I compare my figure (physique) to the figures (physiques) of others;" Faith et al., 1997). Likewise, Fisher and Thompson's (1998) Body Comparison Scale (BCS) assesses the frequency of comparison for multiple body sites (e.g., how often an individual compares her stomach to the stomachs of other individuals of the same sex). Although the aforementioned measures have been used to evaluate general social comparisons, as well as physical appearance and body comparisons, no measure has yet been developed that evaluates comparisons in other specific domains that may lead to body dissatisfaction and disordered eating, namely eating and exercise comparisons. This represents a marked void in the literature, and a clear need for a measure that more comprehensively assesses the *specific* types of comparisons that may lead to such pathology – that is, body, eating, and exercise comparisons.

To the authors' knowledge, only one prior study has examined any aspect of eating or exercise comparison behavior. In particular, one naturalistic, experimental study found that when women in a recreational facility were exposed to a fit peer exercising near them, they experienced lower body satisfaction, as compared to women who were exposed to an unfit peer or to no peer (Wasilenko, Kulik, & Wanic, 2007), perhaps suggesting that an upward comparison of exercise ability may contribute to feeling dissatisfied with one's own weight and shape. Further, although not directly assessing eating and exercise comparisons, other measures have incorporated the potential importance of such comparison constructs. For instance, one scale that assesses verbal messages received from others that may impact body image and eating disturbance includes items such as, "Don't you think you've eaten enough already?" and "You need to start exercising to lose weight," which suggest that women may indeed care about the food and weight ideals and standards held by other women (Herbozo & Thompson, 2006).

However, to date, no questionnaire has been developed that assesses the degree to which one compares her exercise or eating habits to others, and it may be that the dearth of research examining eating and exercise comparisons is related to the lack of a psychometrically sound questionnaire for measuring such comparison-making. The development of a measure that includes these domains would allow researchers to establish the importance of these comparison constructs and would contribute to closing the gap between eating disorders practice and research that has been the focus of much recent work (e.g., Wilson, Schmidt, Nicholls, & Johnson, 2012). Indeed, the study authors have noted a great abundance of eating-, exercise-, and body-related comparison discussion in therapy sessions with eating disorder clients, as well as in discussions with undergraduate female students in courses on eating disorders and body image. Although likely related to the widely studied notion of appearance-related social comparison, eating and exercise comparisons may focus more on the actions necessary to achieve the appearance-related goal gleaned from the body-related comparison. These behavioral comparisons may act as an "environmental alert" telling a woman she must behave differently in order to achieve her ideal weight/shape. Therefore, examining the role of body comparisons, *as well as* eating and exercise comparisons, will be important in terms of coming to a more complete understanding of the impact that social comparison behavior may have on body dissatisfaction and disordered eating.

The goal of the present project was thus to construct and validate a measure of multiple comparison dimensions (i.e., body, eating, exercise) that are theoretically associated with body dissatisfaction and other eating disorder symptomatology: the Body, Eating, and Exercise Comparison Orientation Measure (BEECOM). This answers the call of Corning et al. (2006) for researchers and clinicians to come to a "more complete understanding of the everyday focus and content of women's social comparisons" (p. 345). Our approach involved several components and followed Clark and Watson's (1995) guidelines for scale development. Study 1 (and work preliminary to Study 1) focused on scale construction and validation and involved item generation and exploratory factor analysis (EFA) to refine the item pool. We also sought to confirm the factor structure of the BEECOM via a confirmatory factor analysis (CFA), examine the possible higher-order factor structure of this social comparison measure, and assess the measure's validity (i.e., construct, incremental) in a female college sample. The two-week test–retest reliability of the measure was examined in Study 2. Of note, this measure was developed with women in mind. Most of the prior work examining social comparison in relation to eating pathology has used female-only samples (e.g., Corning et al., 2006; Leahey et al., 2007; Lin & Kulik, 2002; Thompson et al., 1991), and males have been found to report lower levels of social comparison behavior than women (O'Brien et al., 2009). Thus, although the constructs of body, eating, and exercise comparison and their relations to body dissatisfaction and disordered eating may be applicable for men, prior research and our clinical experience have pointed to the particular relevance of these constructs for women's eating pathology. As such, we chose to focus exclusively on women in the initial development and validation of the BEECOM.

Study 1: Item Generation, Scale Refinement, Identification and Confirmation of the Factor Structure, and Measure Validation

The goal of this study was to generate and evaluate body, eating, and exercise comparison items and to provide information on the reliability, factor structure, construct validity, and incremental validity of the resultant questionnaire in a sample of college women. In examining and confirming the factor structure of the BEECOM, we also sought to test whether the BEECOM factors could be explained by a single higher-order factor. Regarding construct validity, we tested the convergent validity of the BEECOM scores in this sample of college women by examining whether they were positively correlated with general and physical appearance social comparison tendencies. We examined the concurrent validity of the BEECOM scores by examining whether they were positively correlated with eating disorder symptomatology and body dissatisfaction. These four constructs were chosen for examining the convergent and concurrent validity of the BEECOM given that we believed that the BEECOM would be correlated with other measures of social comparison (including a general measure, as well as one specific to appearance) and that prior work has linked social comparison tendencies with body dissatisfaction and disordered eating (e.g., Corning et al., 2006; Thompson et al., 1991; Trottier et al., 2007). We tested the discriminant validity of BEECOM scores by examining whether the scores were less strongly correlated with a measure of sexual self-efficacy than with the aforementioned constructs used in examining convergent and concurrent validity in this sample. That is, we hypothesized that BEECOM total and subscale scores would be more strongly correlated with general social comparison tendencies, physical appearance social comparison

tendencies, eating disorder symptomatology, and body dissatisfaction than with a measure of sexual self-efficacy. A measure is thought to have discriminant validity if it has low correlations with a measure that is supposedly not measuring the same construct (Campbell & Fiske, 1959). Finally, we examined body mass index (BMI) in relation to BEECOM scores. Given that Blowers, Loxton, Grady-Flesser, Occhipinti, and Dawe (2003) found a significant positive correlation between BMI and physical appearance comparison behavior in a sample of preadolescent girls, we hypothesized that BMI would be significantly positively correlated with BEECOM scores.

Preliminary Work: Initial Item Generation and Reduction

Keeping women's experiences in mind, items were generated by the study authors and by undergraduate research assistants based on a thorough review of the relevant literature. This included a review of other measures of social comparison (e.g., Faith et al., 1997; Fisher & Thompson, 1998; Gibbons & Buunk, 1999; Thompson et al., 1991) and descriptions of how social comparison relates to body dissatisfaction and disordered eating (e.g., Fairburn, 2008), with most of this literature focusing on the experiences of women. Given the dearth of research and clinical materials on eating- and exercise-related comparisons, reports of female undergraduate research assistants' experiences with and observations of such behaviors were also incorporated into item generation. Items were constructed using guidelines for proper item format (e.g., simple language, avoidance of double-barreled items), and a self-report questionnaire was then developed using the generated items.

We aimed to develop an item pool that was more comprehensive and broader than the target construct – that is, we erred on the side of overinclusiveness (Clark & Watson, 1995). Sixty items were constructed; 20 targeted body-related social comparisons, 20 targeted eating-related social comparisons, and 20 targeted exercise-related social comparisons. Items used a 7-point response scale, with responses of 1 (*never*), 2 (*almost never*), 3 (*seldom*), 4 (*sometimes*), 5 (*often*), 6 (*almost always*), and 7 (*always*).

These 60 items were administered in hard copy format to 46 undergraduate women who were mostly friends and acquaintances of undergraduate research assistants. These women were asked to respond to the questions as honestly as possible and were told that their responses would be completely anonymous (i.e., no demographic or other identifying information was collected). The goal of this preliminary work was to shorten this long list of items to a shorter, but still overinclusive, list of items, given that we were unable include the full set of 60 BEECOM items in our larger scale development study (given additional study goals and research interests). Of note, the body-, eating-, and exercise-related items were mixed randomly in the questionnaire (rather than grouped together into individual sections) in this and every administration of the BEECOM discussed in this manuscript. The response distributions of individual items and item-total correlations were then examined. Items that were highly skewed were considered for elimination (e.g., if almost everyone answered "never" to a given item, little information could be gained from such an item and it was excluded from further consideration). In particular, as per Kline (2005), if the absolute value of the ratio of the unstandardized skew index over its standard error was greater than three, the item was eliminated. The information regarding item skewness, along with low corrected item-total correlations (items with corrected item-total correlations $\leq$.3 were eliminated and other items with low corrected item-total correlations were considered for elimination; Clark & Watson, 1995; Field, 2005), subjective judgments (e.g., content coverage considerations), as well as a desire to retain an equal number of body-, eating-, and exercise-related items, contributed to our choice of which items to retain and which to eliminate. As per Clark and Watson (1995), "there is no substitute for good theory and careful thought when using these [factor analytic] techniques" (p. 314), and as such, we incorporated analytic results and theory in making decisions regarding item retention versus item elimination. Using these points of consideration as a guide, 28 of the 60 items were retained with 8 targeting body-related social comparisons, 10 targeting eating-related social comparisons, and 10 targeting exercise-related social comparisons. At this stage, we then wrote and added in two additional body-related items that we believed were important but whose content had not been covered in the original set of 60 items (see items 9 and 13 in Table 1), bringing us to a total of 30 items. The purpose of Study 1 was to refine this set of items further, using a large sample, to a final, shorter version of the scale that yielded evidence of reliable and valid scores and an empirically backed and conceptually sound factor structure.

Method

Participants. Participants were 441 women attending a Southeastern university who were recruited through introductory psychology courses. These women ranged in age from 17 to 24 years, with a mean age of 18.71 years (SD = 1.01). Most women (73.2%) identified themselves as Caucasian, 9.1% as Black, 8.0% as Latina, 5.0% as Asian, .2% as Pacific Islander, 4.3% as biracial/biethnic, and .2% as other races/ethnicities. Highest parental education was used as a proxy for socioeconomic status and ranged from 7 to 21 years (M = 17.01 years, SD = 2.67).

So that we could conduct both an EFA and a CFA and as has been done in previous work (e.g., Wheaton, Berman, Franklin, & Abramowitz, 2010), this sample was randomly divided into two groups using the SPSS "Random sample of cases" function. The first sample (Sample 1) was used to conduct the EFA, and the second sample (Sample 2) was used for the CFA, as well as to examine construct and incremental validity. Sample 1 was comprised of 226 women who ranged in age from 17 to 24 years, with a mean age of 18.67 years (SD = .97). Most women (74.3%) identified themselves as Caucasian, 8.4% as Black, 7.1% as Latina, 3.1% as Asian, .4% as Pacific Islander, 6.2% as biracial/biethnic, and .4% as other races/ethnicities. Highest parental education ranged from 8 to 21 years (M = 16.93 years, SD = 2.68). Sample 2 was comprised of 215 women who ranged in age from 17 to 23 years, with a mean age of 18.75 years (SD = 1.04). Most women (72.0%) identified themselves as Caucasian, 9.8% as Black, 8.9% as Latina, 7.0% as Asian, and 2.3% as biracial/biethnic. Highest parental education ranged from 7 to 21 years (M = 17.09 years, SD = 2.66). Results indicated that age, $t(435) = .83$, $p = .407$, and highest parental education, $t(439) = .61$, $p = .543$, did not differ across the two samples; additionally, about three-fourths of both samples identified as Caucasian. Thus, the two samples were similar demographically. Further, we also compared the means of all of the study variables listed below across samples; none of these means were found to significantly differ across groups (all $ps > .05$), providing further support for the similarity of Sample 1 and Sample 2.

Measures

Demographics. Demographic data for age, parents' highest levels of education attained, and race/ethnicity were collected via a set of questionnaires created for this study.

Body, Eating, and Exercise Comparison Orientation Measure (BEECOM). The 30-item preliminary BEECOM measure was used to assess an individual's tendency to engage in social comparison in domains related to the body, eating, and exercise. Items were rated on a 7-point scale ranging from 1 (*never*) to 7 (*always*).

General social comparison. General social comparison behavior was measured using the Iowa-Netherlands Comparison

Table 1
Rotated item factor loadings and maximum likelihood communalities for each BEECOM factor obtained from an exploratory factor analysis of Sample 1.

Item	Factor loadings			Communalities
	1	2	3	
Factor 1: Body Comparison Orientation				
2. I pay attention to whether or not I am as thin as, or thinner, than my peers.	**.70**	.30	−.07	*.76*
4. In social situations, I think about how my figure "matches up" to the figures of those around me.	**.80**	.10	.02	*.77*
9. I notice how I compare with my peers in terms of specific parts of the body (e.g., stomach, hips, breasts, etc.).	**.78**	.10	−.02	*.69*
12. I compare my body shape to that of my peers.	**.88**	.18	−.13	*.84*
13. When I see a peer who is wearing revealing clothing, I have thoughts of how my own body compares.	**.83**	−.07	.07	*.70*
17. I pay attention to whether or not I am as toned as my peers.	**.64**	.11	.20	*.73*
Body Comparison Orientation Items that were Deleted				
a. I use my peers as a standard of comparison for what my body should look like.	.45	.49	−.07	*.62*
b. I am aware of what clothing size my peers wear in comparison to myself.	.38	.12	.23	*.41*
c. When I see peers, I find myself wondering if my body is as romantically appealing or sexy to potential romantic partners as theirs.	.80	−.15	.15	*.65*
d. I like to know how much my peers weigh so that I can compare my own weight.	.38	.01	.46	*.58*
Factor 2: Eating Comparison Orientation				
1. I look at the amount of food my peers leave on their plate in comparison to me when they are finished eating.	.04	**.65**	.11	*.58*
3. During meals, I compare what I am eating to what others are eating.	.21	**.80**	−.11	*.75*
7. I find myself thinking about how my food choices compare with the food choices of my peers.	.05	**.81**	.07	*.81*
8. I am quick to notice how healthy (or unhealthy) my peers' food choices are compared to my own food choices.	.06	**.75**	.03	*.66*
11. When I go to the dining hall or out to eat, I pay attention to how much I am eating compared to other people.	.10	**.73**	.15	*.81*
16. I pay attention to how much junk food my peers eat compared to me.	.07	**.73**	.13	*.76*
Eating Comparison Orientation Items that were Deleted				
e. Before I order in a restaurant/cafeteria, I find out what my friends are ordering to make sure I order something as healthy as they do.	−.12	.58	.36	*.64*
f. I compare how often I eat late night food with often my peers eat it.	.04	.55	.26	*.61*
g. I pay attention to how many times throughout the day my peers eat in comparison to me.	.07	.50	.38	*.74*
h. I evaluate how often my peers snack compared to how often I snack.	.07	.39	.46	*.70*
Factor 3: Exercise Comparison Orientation				
5. When I am exercising (e.g., at the gym, running outdoors), I pay attention to the length of time that those around me work out.	.13	.13	**.61**	*.63*
6. I pay close attention when I hear peers talking about exercise (in order to determine if I am exercising as much as they are).	.01	.14	**.78**	*.80*
10. When working out around other people, I think about how many calories I am burning in comparison to my peers.	.11	.13	**.61**	*.60*
14. I like to know how often my friends are working out so I can figure out if the number of times I work out "matches up".	−.04	.13	**.83**	*.82*
15. When I exercise (e.g., at the gym, running outdoors), I pay attention to the intensity level of the workouts of those around me.	.24	−.10	**.70**	*.63*
18. When I work out, I evaluate how hard my workout was compared to how hard my friends say they worked out.	.04	.07	**.76**	*.70*
Exercise Comparison Orientation Items that were Deleted				
i. I am quick to notice other people my age exercising in public (e.g., jogging on the sidewalk, biking).	.52	.06	.15	*.44*
j. I use the amount of time my friends exercise as a standard for the amount of time that I should exercise.	−.08	.33	.53	*.55*
k. I pay attention to how often my peers work out.	.17	.23	.55	*.72*
l. When working out around other people, I compare my performance to that of others.	.39	−.03	.44	*.51*

Note. $n = 226$. Factor 1, Body Comparison Orientation; Factor 2, Eating Comparison Orientation; Factor 3, Exercise Comparison Orientation. The item numbers listed are those that are used in the final 18-item version of the BEECOM.

Orientation Measure (INCOM; Gibbons & Buunk, 1999). This scale consists of 11 items that are rated on a 5-point scale ranging from 1 (*I disagree strongly*) to 5 (*I agree strongly*), with higher scores indicating a greater tendency to engage in social comparisons across life domains; responses to the 11 items are summed to create a total INCOM score. An example item is, "I always pay a lot of attention to how I do things compared with how others do things." Evidence of construct validity in American adolescent and college samples is suggested by the measure's significant relationships with neuroticism, self-monitoring, public and private self-consciousness, and social anxiety in these samples (Gibbons & Buunk, 1999). Gibbons and Buunk (1999) found that estimates of internal consistency ranged from .78 to .85 in American college students; in the present study, alpha was .83 in Sample 2.

Physical appearance social comparison. Appearance-related social comparison tendencies were assessed using the Physical Appearance Comparison Scale (PACS; Thompson et al., 1991). This scale assesses an individual's tendency to compare her own appearance to the appearance of others and consists of five items that are rated on a 5-point scale ranging from 1 (*never*) to 5 (*always*); responses to the five items are summed to create a total PACS score. Construct validity among college women is demonstrated by the strong correlations between the PACS and measures of body dissatisfaction and eating disturbance (Thompson et al., 1991).

Thompson et al. (1991) found adequate internal consistency (coefficient alpha = .78) in a sample of college women; in the present study, alpha was .74 in Sample 2.

Eating disorder symptomatology. The Eating Attitudes Test-26 (EAT-26; Garner, Olmsted, Bohr, & Garfinkel, 1982) was used to assess general eating disorder symptoms. The EAT-26 is one of the most widely used standardized measures of eating disorder attitudes and behaviors (Garner, 2002) and is comprised of 26 items that are rated on a 6-point scale ranging from 1 (*never*) to 6 (*always*), with higher scores reflecting greater eating pathology. An example item is, "Find myself preoccupied with food." Items endorsed as 1, 2, or 3 are scored as "0," while items marked as 4, 5, or 6, are scored as "1," "2," or "3," respectively; the 26 items are summed to create a total EAT-26 score. Studies have found the measure to be effective as a screening measure, with a cutoff score of 20 indicating a probable eating disorder (King, 1989, 1991), and good internal consistency (alpha = .83–.90) and two-week test–retest reliability (r = .84) have been demonstrated in samples of young women (Carter & Moss, 1984; Garner et al., 1982). In the present study, alpha was .90 in Sample 2.

Body dissatisfaction. Body dissatisfaction experienced over the past 28 days was assessed via the Weight Concern and Shape Concern subscales of the Eating Disorder Examination-Questionnaire (EDE-Q; Fairburn & Beglin, 2008), which is one of the most commonly used measures of disordered eating attitudes and behaviors in clinical and community populations (Anderson & Williamson, 2002). These two subscales focus on weight and shape dissatisfaction and the degree to which one's self-worth and acceptance of oneself are defined by weight or shape, and were combined since previous work has indicated that these two subscales load onto one underlying factor (Peterson et al., 2007). In particular, the 12 items, rated on a 7-point scale ranging from 0 to 6 (with items either rated on a *no days* to *everyday* or *not at all* to *markedly* scale), were averaged to reflect body dissatisfaction. An example item is, "How dissatisfied have you been with your weight?" The Weight Concern and Shape Concern subscales have yielded evidence of internal consistency (alphas of .89–.93; Luce & Crowther, 1999) and convergent validity (Fairburn & Beglin, 1994; Grilo, Masheb, & Wilson, 2001) among samples of college women and community and patient groups. In the current study, alpha was .95 in Sample 2.

Sexual self-efficacy. Sexual self-efficacy was assessed using the Refusal subscale of the Sexual Self-Efficacy Scale (SSES; Soet, Dudley, & Dilorio, 1999). The original item wording was slightly modified from "can" to "can/could" so that participants who had not engaged in sexual intercourse could respond to items. This subscale is comprised of four items that are scored on a 10-point scale ranging from 1 (*not at all sure I can do it*) to 10 (*completely sure I can do it*) and that assess an individual's perceived ability to say "no" to a potential sexual partner for various reasons (e.g., "I can/could always say no to sex with someone who is pressuring me to have sex"). Items are summed to create a total Refusal subscale score. The Refusal subscale of the SSES has exhibited internally consistent scores in a sample of college women (α = 74; Soet et al., 1999), and in the current study, alpha was .86 in Sample 2.

Body mass index (BMI). Participants reported on their current weight and height, and we used this information to compute BMI. There is evidence that individuals are generally accurate with their self-reported weights (Shapiro & Anderson, 2003).

Procedure. All participants completed the same online survey in a private setting of their choosing (e.g., their home) as part of a study presented as a study of peers and body image. A link to the survey and consent form was emailed to the participants, followed up by a call from a research assistant to highlight aspects of the consent form and answer any questions about the study. After participants provided electronic consent, they were directed to the questionnaires, which were presented in a fixed order and took 45–60 min to complete; they received course credit for completing the survey.

Of note, 499 undergraduate women originally participated in this study; however, of these individuals, 58 "failed" at least one of the study's three validity checks and were excluded. In particular, the validity checks were made up of three items placed throughout the survey that asked the participants to choose a specific response choice (e.g., "Please choose 'Slightly Agree'"); not responding appropriately to one or more of the three validity check items suggests possible random or inattentive responding. Thus, 11.6% (58 out of 499) of the original study participants were dropped from analyses because of possibly invalid reporting, leaving us with a final sample size of 441. Of these 441 participants who were included in the study, some had missing data. Missing data ranged from a low of .23% for EDE-Q Weight Concern/Shape Concern and SSES Refusal subscale scores to a high of 11.79% for the EAT-26. No individual item had more than 2.5% of values missing, and Little's Missing Completely at Random analysis was non-significant, $\chi^2(8040) = 8022.46$, $p = .553$, indicating that the data for all of the study items were missing completely at random. Overall, this information suggests that the amount and pattern of missingness should not be problematic. This study was reviewed and approved by the university's Institutional Review Board.

Results and Discussion

EFA on Sample 1. Following the recommendations of Clark and Watson (1995), items were retained and submitted to EFA based on item distribution, average correlation with the other items, and item-total correlation. After examining skewness, kurtosis, average correlations with the other items, and item-total correlations, all items were submitted to EFA. That is, no items were dropped due to low average correlations (r < .40) with the other items (Clark & Watson, 1995; McCarthy, Pederson, Thompsen, & Leuty, 2006), low corrected-item total correlations ($\leq$.3; Clark & Watson, 1995; Field, 2005), or very high kurtosis or skewness. Standardized kurtosis values greater than 10 may suggest a problem (Kline, 2005); no items exhibited kurtosis values greater than 10. As aforementioned, an absolute value of the ratio of the unstandardized skew index over its standard error greater than three may be problematic (Kline, 2005); two items exhibited values greater than three. Given that tests of variances and covariances (e.g., factor analyses) are affected by kurtosis more so than skew (DeCarlo, 1997; Jobson, 1991; Mardia, Kent, & Bibby, 1979) and the fact that in large samples (i.e., greater than 200), skewness cutoffs should not be applied (because of the problem of small standard errors; Field, 2005), these two items were retained – their kurtosis values were in the acceptable range and in examining their actual distributions (which Field (2005) recommends should be done in large samples), although skewed, their distributions appeared acceptable; further, they exhibited strong average correlations with the other items (i.e., rs of .56 and .61) and corrected item-total correlations (i.e., rs of .70 and .80).

The 30 items were submitted to EFA with maximum likelihood factoring and oblique Crawford–Ferguson direct quartimax rotation using Comprehensive Exploratory Factor Analysis (CEFA) Version 3.04 (Browne, Cudeck, Tateneni, & Mels, 2010). Oblique rotation was used because we believed the factors reflecting dimensions of social comparison behavior would be correlated. In general, oblique rotation represents a more realistic approach in the search for underlying factors in that it allows rotated factors to be correlated (Fabrigar, Wegener, MacCallum, & Strahan, 1999). The significance of Bartlett's test of sphericity, $\chi^2(435) = 5766.56$, p < .001, and the size of the Kaiser–Meyer–Olkin measure of sampling adequacy (KMO = .96) suggested that the 30 BEECOM items had adequate common variance and that factor analysis was

appropriate for these data (Field, 2005; Worthington & Whittaker, 2006). Further, the root-mean-square error of approximation (RMSEA) for this EFA was .087 (90% confidence interval: .080–.093), which indicates fit in the reasonable to mediocre range (Browne & Cudeck, 1993).

The number of factors to retain was determined by parallel analysis (Horn, 1965). This technique has been found to be more accurate in determining the number of factors than other methods, such as retaining factors with eigenvalues greater than or equal to one (Guttman, 1954; Kaiser, 1960) or examining the scree plot of eigenvalues for breaks (Cattell, 1966; Fabrigar et al., 1999). Such methods are subjective, ambiguous, and tend to overestimate the number of factors (Hayton, Allen, & Scarpello, 2004). Parallel analysis involves extracting eigenvalues from random data sets that parallel the actual data set with regard to the number of cases and variables. The eigenvalues derived from the actual data are then compared to the criterion eigenvalues (i.e., 95th percentile) derived from the random data (O'Connor, 2000). Factors in the actual data are only retained if their eigenvalues are greater than the eigenvalues from the random data (Hayton et al., 2004; O'Connor, 2000).

To determine the number of factors to interpret, we generated 1000 random data sets and compared the eigenvalues from the actual data to the criterion eigenvalues (i.e., 95th percentile) from the random data. The first three factors from the actual data had eigenvalues greater than the criterion eigenvalues generated from the random data (i.e., 16.76 [actual data] compared to 1.09 [random data] for the first factor, 2.10 [actual data] compared to .94 [random data] for the second factor, and 1.20 [actual data] compared to .84 [random data] for the third factor). The remaining factors derived from the actual data had eigenvalues that were lower than the corresponding criterion eigenvalues generated from the random data. As such, we retained three factors, which was also consistent with our conceptualization. These three factors explained 63.98% of the common variance.

Further details of the results of this EFA, including rotated item factor loadings and maximum likelihood communalities, for the 30 BEECOM are provided in Table 1. Based on low factor loadings (<.40; Floyd & Widaman, 1995; Ford, MacCallum, & Tait, 1986), loading on more than one factor (>.30 on second factor; Bosworth, Espelage, & Simon, 1999; Cicero, Kerns, & McCarthy, 2010; Costello & Osborne, 2005; McCarthy et al., 2006; Worthington & Whittaker, 2006), how well items represented the target construct (including content coverage considerations), and our goal of creating a relatively short measure with an equal number of items per subscale, 12 items were deleted. For example, although the body comparison item (Table 1, item c), "When I see peers, I find myself wondering if my body is as romantically appealing or sexy to potential romantic partners as theirs" exhibited a strong loading on the body comparison factor and low loadings on the other two factors, in the end, we decided that this item did not capture an essential piece of the construct of "body comparison orientation." In particular, we decided that we did not want to bring in the notion of considering whether one's own body is as romantically appealing/sexy as a peer's into the body comparison construct and that this aspect of body comparison was likely not a central component of the construct. Although the eating comparison item (item f), "I compare how often I eat late night food with how often my peers eat it," exhibited a relatively strong loading on the eating comparison factor and lower loadings on the other two factors, we believed this item was not as central to the construct of "eating comparison orientation" as retained items and would be less relevant to a broader population (i.e., non-college students). With regard to the exercise comparison items, although one of the items (item k; "I pay attention to how often my peers work out") exhibited relatively strong loadings on the exercise comparison factor and low loadings on the other factors, we believed that the content of this item was already covered in a different item that we chose to retain (i.e., item 14, which also had a stronger loading on the exercise factor than item k).

In the end, most deleted items were excluded because of low factor loadings or loading onto more than one factor: three body comparison items (items a, b, and d), three eating comparison items (items e, g, and h), and three exercise comparison items (items i, j, and l). One body comparison item (item c), one eating comparison item (item f), and one exercise item (item k) were deleted based on subjective reasons and the desire to create a parsimonious scale with an equal number of items per subscale.

The final 18-item scale is depicted in Table 1. On the basis of the content of the items comprising each factor, as expected, the BEECOM is composed of three factors that represent the hypothesized dimensions of social comparison that may play a role in body dissatisfaction and disordered eating. The first factor, labeled Body Comparison Orientation (six items), captures an individual's propensity to engage in body-related comparisons. Previous research has indicated the importance of assessing the impact of body-related comparisons on eating disorder psychopathology (e.g., Leahey et al., 2007; Thompson et al., 1991). The second factor, labeled Eating Comparison Orientation (six items), assesses an individual's propensity to make a variety of eating-related comparisons (e.g., regarding amount and food choices). Finally, the third factor, labeled Exercise Comparison Orientation (six items), assesses an individual's propensity to engage in exercise-related social comparisons (e.g., regarding length of time, calories burned, and intensity level). We additionally found that internal consistency was high for the full scale ($\alpha = .96$), as well as for the subscales (Body: $\alpha = .94$; Eating: $\alpha = .94$; Exercise: $\alpha = .93$) in this sample, and as anticipated, correlations between the subscales were strong (Body & Eating: $r = .70$; Body & Exercise: $r = .68$; Eating & Exercise: $r = .76$; all $ps < .001$).

CFA on Sample 2. CFA with the second random sample that was generated (Sample 2) was used to confirm the three-factor structure of the BEECOM obtained via the EFA on Sample 1. We also examined whether the three BEECOM factors could be explained by a single higher-order factor. Of note, with only three lower-order factors, the higher order factor model (where one higher-order factor influences the three lower-order factors) is equivalent in terms of model fit to the correlated CFA model with only lower-order factors because the higher-order model is just identified (MacCallum, personal communication). Thus, we were not able to empirically evaluate whether one of these models statistically provided a better fit than the other; however, if the BEECOM factors load strongly onto a higher-order factor, this suggests that it is appropriate to compute a total BEECOM score that represents eating disorder-related social comparison orientation. If the lower order factors do not load strongly onto the higher-order factor, then the BEECOM should be viewed as multidimensional and items should not be computed as a total score (Rubio, Berg-Weger, & Tebb, 2001).

CFA was conducted using Mplus Version 5.21 (Muthén & Muthén, 2007). Goodness-of-fit was evaluated using the RMSEA, the standardized root-mean-square residual (SRMR), the comparative fit index (CFI), and the Tucker–Lewis Index (TLI). Good model fit was defined by the following criteria: RMSEA values of about .08 or below (Browne & Cudeck, 1993), SRMR values less than about .05 (Byrne, 1998; Hu & Bentler, 1999), CFI values of about .95 or above (Bentler, 1990; Hu & Bentler, 1999), and TLI values above about .90 (Hu & Bentler, 1999). Multiple fit indices were used together because they provide a more conservative and reliable approach to the evaluation of model fit than the examination of a single index of fit.

The normalized estimate for Mardia's (1970) test of multivariate kurtosis was 13.24 ($p < .001$) in Sample 2. Bentler and Wu (2002) have suggested that values greater than three are indicative of non-normally distributed data that may lead to χ^2 and standard error

Table 2
Standardized maximum likelihood factor loadings with robust standard errors for the Body, Eating, and Exercise Comparison Orientation Measure obtained from a confirmatory factor analysis of Sample 2.

Item	Factor loadings			Higher-order factor
	1	2	3	
Factor 1: Body Comparison Orientation				.82
2. I pay attention to whether or not I am as thin as, or thinner, than my peers.	.89			
4. In social situations, I think about how my figure "matches up" to the figures of those around me.	.88			
9. I notice how I compare with my peers in terms of specific parts of the body (e.g., stomach, hips, breasts, etc.).	.84			
12. I compare my body shape to that of my peers.	.91			
13. When I see a peer who is wearing revealing clothing, I have thoughts of how my own body compares.	.77			
17. I pay attention to whether or not I am as toned as my peers.	.74			
Factor 2: Eating Comparison Orientation				.97
1. I look at the amount of food my peers leave on their plate in comparison to me when they are finished eating.		.82		
3. During meals, I compare what I am eating to what others are eating.		.90		
7. I find myself thinking about how my food choices compare with the food choices of my peers.		.89		
8. I am quick to notice how healthy (or unhealthy) my peers' food choices are compared to my own food choices.		.87		
11. When I go to the dining hall or out to eat, I pay attention to how much I am eating compared to other people.		.89		
16. I pay attention to how much junk food my peers eat compared to me.		.87		
Factor 3: Exercise Comparison Orientation				.87
5. When I am exercising (e.g., at the gym, running outdoors), I pay attention to the length of time that those around me work out.			.80	
6. I pay close attention when I hear peers talking about exercise (in order to determine if I am exercising as much as they are).			.85	
10. When working out around other people, I think about how many calories I am burning in comparison to my peers.			.81	
14. I like to know how often my friends are working out so I can figure out if the number of times I work out "matches up".			.86	
15. When I exercise (e.g., at the gym, running outdoors), I pay attention to the intensity level of the workouts of those around me.			.84	
18. When I work out, I evaluate how hard my workout was compared to how hard my friends say they worked out.			.92	

Note. $n = 215$. Factor 1, Body Comparison Orientation; Factor 2, Eating Comparison Orientation; Factor 3, Exercise Comparison Orientation. Items are rated on a 1–7 scale with the following anchors: *never*, *almost never*, *seldom*, *sometimes*, *often*, *almost always*, and *always*, and item responses are summed to create subscale and total scores. Items are preceded by the following set of instructions:
"Please rate each of the following items regarding how often you compare yourself to your **same-sex peers** in terms of appearance, exercise, and eating. Remember, there are no right or wrong answers, so please be as honest as possible.
Regarding the items that refer to comparisons you might make when you are **exercising** (e.g., running outside, playing an organized sport, using a cardio machine at a gym): **If you are not currently exercising, think back to times when you have exercised (e.g., participated in gym class, played an organized sport, walked or ran outside) and answer accordingly.**"

biases. Thus, because the data were multivariate kurtotic (i.e., the multivariate distribution of the observed variables had both tails and peaks that differed from those characteristic of a multivariate normal distribution; DeCarlo, 1997; Raykov & Marcoulides, 2006) and because multivariate kurtosis can be problematic in structural equation modeling analyses, maximum likelihood parameter estimates with standard errors and a χ^2 test statistic that are robust to non-normality were used for these analyses. The RMSEA (.078; 90% confidence interval: .067–.090), SRMR (.041), CFI (.946), and TLI (.937) all approximated good fit according to the aforementioned criteria. As previously noted, because the lower-order model and the higher-model were statistically equivalent, the higher-order model had the same fit as the lower-order model (RMSEA: .078; SRMR: .041; CFI: .946; TLI: .937). See Table 2 for standardized factor loadings. Given that the lower-order factors loaded strongly onto a higher-order factor (loadings of .82–.97), we concluded that it is appropriate to use a total BEECOM score as an indicator of an overarching eating disorder-related social comparison orientation. That is, in addition to BEECOM subscale scores being available for the researcher interested in a specific aspect of eating disorder-related comparison, the total score is also appropriate to use.

Reliabilities and mean scores observed in Sample 2. As seen in Table 3, the BEECOM (the subscales and total) exhibited excellent internal consistency in Sample 2 ($\geq$.93). The mean total BEECOM score was 67.68 ($SD = 23.84$), just below the scale's mid-point of 72 (BEECOM total possible range: 18–126). According to the Flesch reading ease formula, the final 18-item version of the BEECOM was at a grade reading level of 4.4 (Flesch, 1948).

Examining construct and incremental validity of the BEECOM in Sample 2. As seen in Table 3, the BEECOM total score and BEECOM subscale scores (which were significantly, highly correlated) were significantly positively correlated with measures of general social comparison orientation and physical appearance social comparison behavior (*r*s from .42 to .76), providing evidence of convergent validity in this sample. Of note and as expected, the PACS (i.e., a measure of physical appearance comparison behavior) was significantly more highly correlated with the BEECOM Body Comparison Orientation subscale ($r = .76$) than the Eating Comparison Orientation ($r = .57$; $Z = 5.42$, $p < .001$) or Exercise Comparison Orientation ($r = .53$; $Z = 6.08$, $p < .001$) subscales (Meng, Rosenthal, & Rubin, 1992). The BEECOM total and subscale scores were also significantly positively correlated with measures of eating disorder symptomatology and body dissatisfaction (*r*s from .60 to .75), providing evidence of concurrent validity in this sample. As hypothesized, the BEECOM total and subscale scores were significantly more highly correlated with measures of general social comparison orientation, physical appearance social comparison behavior, eating disorder symptomatology, and body dissatisfaction than with

Table 3
Correlations among and means and standard deviations of the measured variables in Sample 2.

	1	2	3	4	5	6	7	8	9	10	
1. BEECOM – Body Comparison Orientation	**.93**										$M = 26.81$, $SD = 7.91$
2. BEECOM – Eating Comparison Orientation	.76***	**.95**									$M = 22.00$, $SD = 9.26$
3. BEECOM – Exercise Comparison Orientation	.70***	.80***	**.94**								$M = 19.07$, $SD = 8.85$
4. BEECOM – Total	.89***	.93***	.92***	**.97**							$M = 67.68$, $SD = 23.84$
5. INCOM	.58***	.53***	.42***	.55***	**.83**						$M = 39.07$, $SD = 6.60$
6. PACS	.76***	.57***	.53***	.67***	.57***	**.74**					$M = 16.68$, $SD = 3.58$
7. EAT-26	.64***	.66***	.60***	.70***	.34***	.54***	**.90**				$M = 9.23$, $SD = 9.86$
8. Weight Concern/Shape Concern	.74***	.68***	.61***	.75***	.40***	.57***	.74***	**.95**			$M = 2.47$, $SD = 1.62$
9. SSES Refusal subscale	−.29***	−.18*	−.21**	−.26***	−.04	−.26***	−.16*	−.24***	**.86**		$M = 35.16$, $SD = 6.81$
10. BMI	.20**	.17*	.14*	.19**	.06	.10	.08	.44***	−.05	–	$M = 22.57$, $SD = 4.06$

Note. $n = 215$. BEECOM = Body, Eating, and Exercise Comparison Orientation Measure; INCOM = Iowa-Netherlands Comparison Orientation Measure; PACS = Physical Appearance Comparison Scale. EAT-26 = Eating Attitudes Test-26. Weight Concern/Shape Concern = combined subscales from the Eating Disorder Examination-Questionnaire. SSES = Sexual Self-Efficacy Scale. BMI = body mass index. Variables are continuous, with higher values reflecting higher levels of the construct. Possible ranges for the study variables are as follows: BEECOM total (18–126), BEECOM subscales (6–42), INCOM (11–55), PACS (5–25), EAT-26 (0–78), EDE-Q: Weight Concern/Shape Concern (0–6), SSES Refusal subscale (4–40). Reliabilities of measures are reported on the diagonal.

* $p < .05$.
** $p < .01$.
*** $p < .001$.

a measure of sexual self-efficacy (*r*s from −.18 to −.29; *Z* statistics for all comparisons of correlations, $p < .001$, Meng et al., 1992), providing support for the discriminant validity of the BEECOM in the current sample of college women. Finally, the BEECOM total and subscale scores were significantly positively correlated with BMI.

Further, we evaluated whether the BEECOM subscales predicted variance in eating disorder symptomatology and body dissatisfaction above and beyond that of existing comparison measures (i.e., physical appearance comparison behavior – as measured by the PACS, general appearance comparison orientation – as measured by the INCOM) using hierarchical multiple regression in order to examine the potential incremental validity of this new measure. In order to provide an even more stringent test of the explanatory power of the BEECOM, we also controlled for BMI, given that it is a significant predictor of disordered eating and body image disturbance (Stice, 2002). Thus, for eating disorder symptomatology and body dissatisfaction as separate dependent variables, we entered BMI at Step 1, PACS and INCOM scores at Step 2, and the BEECOM subscales at Step 3. A statistically significant R^2 at Step 3 would indicate incremental validity evidence for the BEECOM subscales as a set. We also examined the significance of the βs for each of the BEECOM subscales in order to determine which ones significantly accounted for unique variance in eating disorder symptomatology and body dissatisfaction above and beyond all of the other variables in the model.

As seen in Table 4, after considering the contribution of BMI and existing social comparison measures (i.e., PACS, INCOM), as a set, the BEECOM subscales predicted unique variance in both eating disorder symptomatology (ΔR^2 at Step 3 = .20) and body dissatisfaction (ΔR^2 at Step 3 = .18). That is, the BEECOM subscales accounted for between 18% and 20% of additional variance in college women's eating disorder symptomatology and body dissatisfaction, above and beyond BMI and other comparison measures. In examining the significance of the βs for each of the BEECOM subscales, it appears that the BEECOM Body and Eating Comparison Orientation subscales accounted for unique variance in eating disorder symptomatology and body dissatisfaction, while the BEECOM Exercise Comparison Orientation subscale did not.

Overall, Study 1 resulted in a set of items generated from both theory and empirical data. The results from the EFA indicated the presence of three interpretable factors that were consistent with our original conceptualization of the types of social comparison that play a role in body dissatisfaction and disordered eating. We confirmed the factor structure of the BEECOM that was obtained via an EFA in Sample 1 using a CFA in Sample 2. The fit indices suggested good model fit, and the BEECOM was found to be composed of three highly correlated factors that represent related aspects of social comparison behavior. Results also suggested that because the lower-order factors (i.e., Body Comparison Orientation, Eating Comparison Orientation, and Exercise Comparison Orientation) loaded highly onto the higher-order factor, the BEECOM has a single higher-order factor (i.e., eating disorder-related social comparison orientation). This suggests that not only is it meaningful to compute and use the BEECOM subscale scores to assess targeted interests, but that it is also meaningful and useful to compute scores across all items. This total score represents an individual's eating disorder-related social comparison orientation in a more comprehensive manner than has been done to date; in other words, this total score represents her propensity to engage in social comparison behavior in domains that have implications for body image disturbance and eating disorder symptomatology.

Study 1 also provided support for the internal consistency, convergent validity, concurrent validity, and discriminant validity of the BEECOM total and subscale scores in this sample of college women. There was some evidence for the incremental validity of the BEECOM subscales in this sample as well, particularly for the Body and Eating Comparison Orientation subscales. Although, the BEECOM Exercise Comparison Orientation subscale did not significantly predict eating disorder symptomatology and body dissatisfaction above and beyond the effects of BMI, physical appearance and general social comparison tendencies, and the BEECOM Body and Eating Comparison Orientation subscales, future research may wish to examine if there are certain, specific eating disorder symptoms that the BEECOM Exercise subscale may predict (e.g., compulsive exercise behavior). Additionally, examining the Exercise Comparison Orientation subscale in a sample engaging in regular exercise merits consideration.

Content validity. Beyond the empirical data presented, we also asked three context experts (in the areas of social comparison, body image and disordered eating, and eating disorder-related measure development) to review the final 18-item version of the measure and to indicate whether they believed the set of items comprehensively represented the constructs it was designed to measure. All three reported that the BEECOM comprehensively assessed the constructs of body, eating, and exercise social comparison orientation; for example, they noted that the constructs were covered "quite well" and that the BEECOM "does a good job." Thus, both the psychometric properties of the BEECOM and reviews from content experts indicate that the BEECOM is a sound measure of what it purports to measure – that is, body, eating, and exercise comparison orientation.

Table 4
Incremental variance in eating disorder symptomatology and body dissatisfaction accounted for by BEECOM subscale scores in Sample 2.

	Cumulative R^2	Adjusted R^2	ΔR^2	ΔF (*df*s)	β	*t* (*df*s)
Dependent variable: eating disorder symptomatology (EAT-26), overall $F(6,146) = 25.18$, $p < .001$						
Step 1	.01	.00	.01	0.97 (1,151)		
BMI					.08	0.98 (1,151)
Step 2	.31	.29	.30	32.12 (2,149)***		
BMI					.05	0.73 (3,149)
PACS					.46	5.65 (3,149)***
INCOM					.14	1.66 (3,149)
Step 3	.51	.49	.20	20.08 (3,146)***		
BMI					−.03	−0.46 (6,146)
PACS					.12	1.32 (6,146)
INCOM					−.05	−0.73 (6,146)
BEECOM Body					.23	1.98 (6,146), $p = .05$
BEECOM Eating					.34	3.26 (6,146)**
BEECOM Exercise					.15	1.50 (6,146)
Dependent variable: body dissatisfaction (Weight Concern/Shape Concern), overall $F(6,165) = 50.83$, $p < .001$						
Step 1	.18	.17	.18	36.20 (1,170)***		
BMI					.42	6.02 (1,170)***
Step 2	.47	.46	.29	46.32 (2,168)***		
BMI					.38	6.76 (3,168)***
PACS					.43	6.26 (3,168)***
INCOM					.17	2.47 (3,168)*
Step 3	.65	.64	.18	28.25 (3,165)***		
BMI					.29	6.11 (6,165)***
PACS					.04	0.50 (6,165)
INCOM					−.02	−0.25 (6,165)
BEECOM Body					.44	4.70 (6,165)***
BEECOM Eating					.23	2.62 (6,165)*
BEECOM Exercise					.08	0.98 (6,165)

Note. $n = 215$. EAT-26 = Eating Attitudes Test-26; Weight Concern/Shape Concern = combined subscales from the Eating Disorder Examination-Questionnaire; BMI = body mass index; PACS = Physical Appearance Comparison Scale; INCOM = Iowa-Netherlands Comparison Orientation Measure; BEECOM = Body, Eating, and Exercise Comparison Orientation Measure.

* $p < .05$.
** $p < .01$.
*** $p < .001$.

Study 2: Two-Week Test–Retest Reliability of the BEECOM

The primary goal of Study 2 was to examine the test–retest reliability of the BEECOM over the course of two weeks. Given that theory has generally conceptualized one's overall tendency to make social comparisons as a trait (e.g., Tiggemann & McGill, 2004) and that past research has indicated the stability of social comparison measures (e.g., Gibbons & Buunk, 1999; Thompson et al., 1991), we hypothesized that the BEECOM total and subscale scores would exhibit at least a reasonable level of temporal stability among a sample of college women.

Method

Participants. All 441 women who took part in the first part of the study were contacted and asked to provide additional data about two weeks later. As part of the informed consent process for the first part of the study, all participants were informed that they would be re-contacted in two weeks and asked to respond to a small subset of the original study questionnaires that was expected to take 10–15 min. They were informed that in exchange for their participation in this follow-up portion of the study, they would be entered into a drawing for one of four $100 cash prizes. Of the original sample of 441, 362 (82%) completed a small subset of the original study questionnaires online about two weeks later. These women ranged in age from 17 to 24 years, with a mean age of 18.72 years ($SD = 1.01$). Most women (73.4%) identified themselves as Caucasian, 8.9% as Black, 7.8% as Latina, 5.3% as Asian, .3% as Pacific Islander, 4.2% as biracial/biethnic, and .3% as other races/ethnicities. Highest parental education ranged from 7 to 21 years ($M = 16.98$, $SD = 2.64$). Individuals who provided data two weeks later were compared with individuals who did not provide such data using *t*-tests; these groups were not significantly different from each other on demographic variables (i.e., age, highest parental education, BMI), BEECOM total or subscale scores, body dissatisfaction, or eating disorder symptoms. Additionally, about three-fourths of both samples identified as Caucasian. Thus, those that completed the test–retest portion of the study appear to be representative of those who began the study on these study variables, which minimizes concerns regarding attrition.

Of note, the actual length of time between administrations ranged from 9 to 41 days, with an average of 14 days between administrations and the vast majority (85%) completing the retest within 12–16 days of the initial test. In addition to running test–retest analyses on the full set of 362 participants who completed the test–retest portion of the study, we also ran these analyses including only those whose length of time between administrations was between 12 and 16 days. That is, we reran analyses deleting the extreme outliers (i.e., anyone with lower than 12 and higher than 16 days between administrations; $n = 47$) and individuals for whom the time between administrations was unknown ($n = 6$). In total, there were 53 such cases (15% of the test–retest sample).

Measures

Body, Eating, and Exercise Comparison Orientation Measure (BEECOM). Although all 30 BEECOM items were administered to study participants, the final 18-item version of the BEECOM was used in all Study 2 analyses. Items were rated on a 7-point scale ranging from 1 (*never*) to 7 (*always*).

Results and Discussion

Test–retest reliability for the BEECOM total and subscale scores was determined by means of calculating correlation coefficients between scores at the first and second administrations of the BEECOM. Test–retest reliability was high for the total scale ($r = .90$), as well as for the subscales (Body Comparison Orientation: $r = .85$; Eating Comparison Orientation: $r = .88$; Exercise Comparison Orientation: $r = .84$) in the full sample of test–retest completers ($N = 362$). All correlations were significant at the $p < .001$ level. Results of the analyses on those participants for whom we knew the length of time between administrations was between 12 and 16 days ($n = 309$) were nearly identical to that of the full sample; that is, test–retest reliability was high for the total scale ($r = .89$), as well as for the subscales (Body Comparison Orientation: $r = .85$; Eating Comparison Orientation: $r = .87$; Exercise Comparison Orientation: $r = .84$). All correlations were significant at the $p < .001$ level.

Overall, results indicated that two-week test–retest reliabilities for the BEECOM total and subscale scores were high. This is in line with research supporting the existence of a trait-like tendency to engage in social comparison behavior (both generally and specific to appearance; e.g., Corning et al., 2006; Thompson et al., 1991). Results of the current study suggest that the tendencies to engage in body-, eating-, and exercise-related social comparison behavior exhibit a trait-like quality as well. However, past work has also suggested that social comparison behavior (e.g., comparing one's body to another's) may be triggered by the presentation of certain targets and/or by certain individual motivational factors (Tiggemann & McGill, 2004); given the dearth of research on eating and exercise comparisons, future research may wish to examine whether these processes can be triggered as well, and if so, the moderating factors that influence their activation.

General Discussion

The goal of this study was to develop and explore the psychometric properties of a more comprehensive measure of the types of social comparison that likely play a role in the development and maintenance of body dissatisfaction and disordered eating. Results indicated that the BEECOM total and subscale scores were reliable and valid for measuring such a construct in this sample of college women. To date, research has focused only on the relations between general- and appearance-related social comparison tendencies and eating disorder symptomatology (e.g., Corning et al., 2006; Thompson et al., 1991) and has yet to examine how comparisons in other specific domains may be associated with such pathology. The BEECOM fills this marked void by more comprehensively assessing the specific types of social comparisons theoretically associated with eating disorder symptomatology – namely, body, eating, and exercise comparisons.

Study 1 indicated that the BEECOM is comprised of three correlated dimensions of social comparison hypothesized to play a role in eating pathology. The first factor, Body Comparison Orientation, assesses an individual's propensity to engage in body-related social comparisons and is similar to what is assessed by other appearance-related social comparison measures (i.e., the PACS, Thompson et al., 1991; the BICS, Faith et al., 1997; the BCS, Fisher & Thompson, 1998). The second factor, Eating Comparison Orientation, assesses an individual's tendency to engage in comparisons related to their own eating behavior, such as those regarding the amount or types of food she eats as compared to a same-sex peer. Finally, the third factor, Exercise Comparison Orientation, assesses an individual's tendency to engage in exercise-related comparisons; these may involve an individual comparing the amount or intensity of her exercise to that of a same sex peer's. This factor structure was confirmed via a CFA on a second sample. Furthermore, results indicated the presence of a higher-order factor that is a global representation of the three lower-order factors. Thus, results suggest that not only is it appropriate to compute and use the BEECOM subscale scores as indicators of body, eating, and exercise comparison behavior, but that the BEECOM total score is also appropriate as a comprehensive measure of eating disorder-related social comparison orientation. Results of Study 1 also provided evidence for the convergent, concurrent, and discriminant validity of the BEECOM total and subscale scores in college women. Incremental validity of the BEECOM, particularly of the Body and Eating Comparison Orientation subscales, was demonstrated, as well. Finally, in Study 2, the two-week test–retest reliability of the BEECOM total and subscale scores provided evidence of the trait-like nature of these constructs in college women.

The current research has several strengths. Given that, to date, only measures of general- and appearance-related social comparisons exist, one strength is the development of a measure that more comprehensively assesses the types of social comparison that may be associated with body dissatisfaction and eating disorder symptoms. In particular, to the authors' knowledge, this is the first study to specifically acknowledge the potential roles that eating and exercise comparisons may play in the development and maintenance of eating pathology. Interestingly, some research has suggested that general measures of social comparison may be too general and that appearance-related measures of social comparison may be too narrow to adequately predict body dissatisfaction (Fitzsimmons-Craft et al., 2012), further supporting the need for this new measure. Another strength is that multiple important psychometric aspects of the BEECOM were examined, including information on internal consistency, convergent validity, concurrent validity, discriminant validity, incremental validity, and two-week test–retest reliability. Lastly, the large sample size is a strength and permitted the application of both EFA and CFA in the development and validation of the BEECOM.

An additional strength of the BEECOM is its flexibility. Researchers are able to use the total score if they are interested in comprehensively capturing the types of social comparison that may be associated with disordered eating; however, if researchers have more targeted interests in capturing a certain type of social comparison behavior, using the BEECOM subscales scores affords that possibility. It may be that moderators act differently in relations between BEECOM subscale scores and measures of eating pathology, thus compelling the use of the subscales. For example, it may be that motivations for exercise would be an important moderator to consider in the Exercise Comparison Orientation-eating pathology relation but not in the Eating Comparison Orientation-eating pathology relation. As another example of the flexibility of the BEECOM, researchers can examine whether there are individuals who focus their comparisons on one domain more so than the others and whether there are implications for this. For example, might individuals who endorse eating-related social comparisons more so than other comparisons benefit more from interventions targeted at increasing awareness of these specific tendencies?

One limitation of this study is that the generalizability of these findings is limited to similar samples, namely undergraduate women. Future research should examine the BEECOM and its psychometric properties in other groups, including women of different ages as well as men. While it is believed that the BEECOM would generally be applicable to men, although perhaps less closely tied with eating pathology than it is for women, it is likely that some items would need to be modified to better capture their experiences. In particular, items from the Body Comparison Orientation subscale would require modification (e.g., items that focus on thinness may need to instead focus on muscularity, item that discusses "revealing clothing"). Further, since our sample was

primarily Caucasian, work remains to determine the psychometrics of the BEECOM in samples of diverse races/ethnicities. Future research may also want to examine the relation between scores on the BEECOM and a measure of social desirability to ensure that this measure is not substantially related to participants' response style. Additional limitations include the fixed order of the questionnaires, as order effects could not be controlled for, and the largely cross-sectional design (with the exception of the test–retest portion) which prevented obtaining information about predictive power. Future research should examine if the BEECOM total and/or subscale scores prospectively predict the development of or increases in body dissatisfaction and eating pathology. Finally, the BEECOM does not differentiate between tendencies toward upward and downward comparisons; however, research has indicated that more than 80% of comparisons made by women in their natural environment are in the upward direction (Leahey et al., 2007). Future research may still want to explicitly examine if the body-, eating-, and exercise-related comparisons captured by the BEECOM are most representative of upward or downward comparisons.

If future research bears out a causal role of body, eating, and exercise comparison behavior in the development and maintenance of body dissatisfaction and/or eating pathology, interventions should focus on mitigating such comparisons or their effects. For example, these comparisons could be targeted by having individuals identify and monitor the specific body, eating, and exercise comparisons that they make. Then, clinicians could work with individuals on understanding the function that such comparisons serve, as well as their precursors/triggers and consequences. It may be that the act of simply tracking and becoming more aware of such comparison behaviors will lead an individual to make fewer of them, demonstrating the phenomenon of reactivity (Campbell & Stanley, 1963). Indeed, in an ecological momentary assessment (EMA) study of appearance-related social comparison behavior, Leahey, Crowther, and Ciesla (2011) found that participants reported that recording their appearance-related comparisons increased their awareness of how often they engaged in such behaviors and that the number of comparisons made during the first two days of the EMA study period was significantly greater than the number of comparisons made during the last two days of the study period. Thus, making individuals more aware of their comparison behavior and helping them identify triggers and consequences of such behavior may be an effective intervention in decreasing body, eating, and exercise comparisons and their negative effects. Indeed, Fairburn (2008) suggests that addressing comparison-making in cognitive-behavioral therapy for eating disorders may be an important step in addressing weight/shape concerns and other forms of eating pathology.

In conclusion, this research aimed to improve our understanding and assessment of the types of social comparisons that play a role in eating disorder psychopathology. Fairburn (2008) has noted the importance of addressing appearance social comparison behavior in eating disorders treatment, and the current study suggests that not only is it likely important to address body-related comparisons but that eating- and exercise-related comparisons may be important intervention targets, as well. This study presents the BEECOM as a tool for assessing body-, eating-, and exercise-related social comparisons that has demonstrated good psychometric properties and that will likely be useful for more comprehensively assessing the types of social comparison behaviors that have an influence on eating disorder-related psychopathology.

References

Anderson, D. A., & Williamson, D. A. (2002). Outcome measurement in eating disorders. In W. W. IsHak, T. Burt & L. I. Sederer (Eds.), *Outcome measurement in psychiatry: A critical review* (pp. 289–301). Washington, DC: American Psychiatric Press.

Bamford, B., & Halliwell, E. (2009). Investigating the role of attachment in social comparison theories of eating disorders within a non-clinical female population. *European Eating Disorders Review*, *17*, 371–379.

Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin*, *107*, 238–246.

Bentler, P. M., & Wu, E. J. C. (2002). *EQS for Windows user's guide*. Encino, CA: Multivariate Software, Inc.

Blowers, L. C., Loxton, N. J., Grady-Flesser, M., Occhipinti, S., & Dawe, S. (2003). The relationship between sociocultural pressure to be thin and body dissatisfaction in preadolescent girls. *Eating Behaviors*, *4*, 229–244.

Bosworth, K., Espelage, D. L., & Simon, T. R. (1999). Factors associated with bullying behavior in middle school students. *Journal of Early Adolescence*, *19*, 341–362.

Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 136–192). Newbury Park, CA: Sage.

Browne, M. W., Cudeck, R., Tateneni, K., & Mels, G. (2010). *CEFA: Comprehensive exploratory factor analysis, Version 3.04 (Computer software and manual)*.

Byrne, B. M. (1998). *Structural equation modeling with LISREL, PRELIS, and SIMPLIS: Basic concepts, applications, and programming*. Mahwah. NJ: Erlbaum.

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait–multimethod matrix. *Psychological Bulletin*, *56*, 81–105.

Campbell, D. T., & Stanley, J. C. (1963). *Experimental and quasi-experimental designs for research*. Boston: Houghton Mifflin.

Carter, P. I., & Moss, R. A. (1984). Screening for anorexia and bulimia nervosa in a college population: Problems and limitations. *Addictive Behaviors*, *9*, 417–419.

Cash, T. F., Cash, D. W., & Butters, J. W. (1983). Mirror, mirror, on the wall. . .?: Contrast effects and self-evaluations of physical attractiveness. *Personality and Social Psychology Bulletin*, *9*, 351–358.

Cattell, R. B. (1966). The scree test for the number of factors. *Multivariate Behavioral Research*, *1*, 245–276.

Cicero, D. C., Kerns, J. G., & McCarthy, D. M. (2010). The Aberrant Salience Inventory: A new measure of psychosis proneness. *Psychological Assessment*, *22*, 688–701.

Clark, L. A., & Watson, D. (1995). Constructing validity: Basic issues in objective scale development. *Psychological Assessment*, *7*, 309–319.

Corning, A. F., Krumm, A. J., & Smitham, L. A. (2006). Differential social comparison processes in women with and without eating disorder symptoms. *Journal of Counseling Psychology*, *53*, 338–349.

Costello, A. B., & Osborne, J. W. (2005). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical Assessment, Research & Evaluation*, *10*, 1–9.

DeCarlo, L. T. (1997). On the meaning and use of kurtosis. *Psychological Methods*, *2*, 292–307.

Dittmar, H. (2005). Vulnerability factors and processes linking sociocultural pressures and body dissatisfaction. *Journal of Social and Clinical Psychology*, *24*, 1081–1087.

Dittmar, H., & Howard, S. (2004). Thin ideal internalization and social comparison tendency as moderators of media models impact on women's body focused anxiety. *Journal of Social and Clinical Psychology*, *23*, 768–791.

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.

Fairburn, C. G. (2008). *Cognitive behavior therapy and eating disorders*. New York: Guilford.

Fairburn, C. G., & Beglin, S. J. (1994). Assessment of eating disorders: Interview or self-report questionnaire? *International Journal of Eating Disorders*, *16*, 363–370.

Fairburn, C. G., & Beglin, S. (2008). Eating Disorder Examination Questionnaire (EDE-Q 6.0). In C. G. Fairburn (Ed.), *Cognitive behavior therapy and eating disorders* (pp. 309–313). New York: Guilford.

Faith, M. S., Leone, M. A., & Allison, D. B. (1997). The effects of self-generated comparison targets, BMI, and social comparison tendencies on body image appraisal. *Eating Disorders*, *5*, 128–140.

Festinger, L. (1954). A theory of social comparison processes. *Human Relations*, *7*, 117–140.

Field, A. (2005). *Discovering statistics using SPSS* (2nd ed.). London: Sage.

Fisher, E., & Thompson, J. K. (1998). *Social comparison and body image: An investigation of body comparison processes using multidimensional scaling*. Unpublished manuscript, University of South Florida.

Fitzsimmons-Craft, E. E., Harney, M. B., Koehler, L. G., Danzi, L. E., Riddell, M. K., & Bardone-Cone, A. M. (2012). Explaining the relation between thin ideal internalization and body dissatisfaction among college women: The roles of social comparison and body surveillance. *Body Image*, *9*, 43–49.

Flesch, R. (1948). A new readability yardstick. *Journal of Applied Psychology*, *32*, 221–233.

Floyd, F. J., & Widaman, K. F. (1995). Factor analysis in the development and refinement of clinical assessment instruments. *Psychological Assessment*, *7*, 286–299.

Ford, J. K., MacCallum, R. C., & Tait, M. (1986). The application of exploratory factor analysis in applied psychology: A critical review and analysis. *Personnel Psychology*, *39*, 291–314.

Garner, D. M. (2002). Measurement of eating disorder psychopathology. In C. G. Fairburn & K. D. Brownell (Eds.), *Eating disorders and obesity: A comprehensive handbook* (pp. 141–146). New York, NY: Guilford Press.

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.

Gibbons, F. X., & Buunk, B. P. (1999). Individual differences in social comparison: Development of a scale of social comparison orientation. *Journal of Personality and Social Psychology*, *76*, 129–142.
Gilbert, N., & Meyer, C. (2003). Social anxiety and social comparison: Differential links with restrictive and bulimic attitudes among nonclinical women. *Eating Behaviors*, *4*, 257–264.
Gilbert, P., Price, J., & Allan, S. (1995). Social comparison, social attractiveness and evolution: How might they be related? *New Ideas in Psychology*, *13*, 149–165.
Grilo, C. M., Masheb, R. M., & Wilson, G. T. (2001). A comparison of different methods for assessing the features of eating disorders in patients with binge eating disorder. *Journal of Consulting and Clinical Psychology*, *69*, 317–322.
Guttman, L. (1954). Some necessary and sufficient conditions for common factor analysis. *Psychometrika*, *19*, 149–161.
Hayton, J. C., Allen, D. G., & Scarpello, V. (2004). Factor retention decisions in exploratory factor analysis: A tutorial on parallel analysis. *Organizational Research Methods*, *7*, 191–205.
Heinberg, L. J., & Thompson, J. K. (1992). The effects of figure size feedback (positive vs. negative) and target comparison group (particularistic vs. universalistic) on body image disturbance. *International Journal of Eating Disorders*, *12*, 441–448.
Herbozo, S., & Thompson, J. (2006). Development and validation of the verbal commentary on physical appearance scale: Considering both positive and negative commentary. *Body Image*, *3*, 335–344.
Hildebrandt, T., Shiovitz, R., Alfano, L., & Greif, R. (2008). Defining body deception and its role in peer based social comparison theories of body dissatisfaction. *Body Image*, *5*, 299–306.
Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika*, *32*, 179–185.
Hu, L.-T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes on covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, *6*, 1–55.
Jobson, J. D. (1991). *Applied multivariate data analysis. Regression and experimental design* New York: Springer-Verlag.
Kaiser, H. F. (1960). The application of electronic computers to factor analysis. *Educational and Psychological Measurement*, *20*, 141–151.
King, M. B. (1989). Eating disorders in general practice population: Prevalence, characteristics and follow-up at 12 to 18 months. *Psychological Medicine, Monograph Supplement*, *14*, 1–34.
King, M. B. (1991). The natural history of eating pathology in attenders to primary medical care. *International Journal of Eating Disorders*, *10*, 379–387.
Klein, W. M. (1997). Objective standards are not enough: Affective, self-evaluative, and behavioral responses to social cognition information. *Journal of Personality and Social Psychology*, *72*, 763–774.
Kline, R. B. (2005). *Principles and practice of structural equation modeling* (2nd ed.). New York: Guilford.
Krones, P. G., Stice, E., Batres, C., & Orjada, K. (2005). In vivo social comparison to a thin-ideal peer promotes body dissatisfaction: A randomized experiment. *International Journal of Eating Disorders*, *38*, 134–142.
Leahey, T. M., Crowther, J. H., & Ciesla, J. A. (2011). An ecological momentary assessment of the effects of weight and shape social comparisons on women with eating pathology, high body dissatisfaction, and low body dissatisfaction. *Behavior Therapy*, *42*, 197–210.
Leahey, T. M., Crowther, J. H., & Mickelson, K. D. (2007). The frequency, nature, and effects of naturally occurring appearance-focused social comparisons. *Behavior Therapy*, *38*, 132–143.
Lin, L. F., & Kulik, J. A. (2002). Social comparison and women's body satisfaction. *Basic and Applied Social Psychology*, *24*, 115–123.
Luce, K. H., & Crowther, J. H. (1999). The reliability of the Eating Disorder Examination – Self-report Questionnaire Version (EDE-Q). *International Journal of Eating Disorders*, *25*, 349–351.
Mardia, K. V. (1970). Measures of multivariate skewness and kurtosis with applications. *Biometrika*, *57*, 519–530.
Mardia, K. V., Kent, J. T., & Bibby, J. M. (1979). *Multivariate analysis*. New York: Academic Press.
McCarthy, D. M., Pedersen, S. L., Thompsen, D. M., & Leuty, M. E. (2006). Development of a measure of drinking and driving expectancies for youth. *Psychological Assessment*, *18*, 155–164.
Meng, X., Rosenthal, R., & Rubin, D. B. (1992). Comparing correlated correlation coefficients. *Psychological Bulletin*, *111*, 172–175.
Morrison, T. G., Kalin, R., & Morrison, M. A. (2004). Body-image evaluation and body-image investment among adolescents: A test of sociocultural and social comparison theories. *Adolescence*, *39*, 571–592.
Muthén, L. K., & Muthén, B. O. (2007). *Mplus user's guide* (5th ed.). Los Angeles: Muthén & Muthén.
Myers, T. A., & Crowther, J. H. (2009). Social comparison as a predictor of body dissatisfaction: A meta-analytic review. *Journal of Abnormal Psychology*, *118*, 683–698.
O'Brien, K. S., Caputi, P., Minto, R., Peoples, G., Hooper, C., Kell, S., et al. (2009). Upward and downward physical appearance comparisons: Development of scales and examination of predictive qualities. *Body Image*, *6*, 201–206.
O'Connor, B. P. (2000). SPSS and SAS programs for determining the number of components using parallel analysis and Velicer's MAP test. *Behavior Research Methods, Instruments, & Computers*, *32*, 396–402.
Peterson, C. B., Crosby, R. D., Wonderlich, S. A., Joiner, T., Crow, S. J., Mitchell, J. E., et al. (2007). Psychometric properties of the Eating Disorder Examination-Questionnaire: Factor structure and internal consistency. *International Journal of Eating Disorders*, *40*, 386–389.
Raykov, T., & Marcoulides, G. A. (2006). *A first course in structural equation modeling* (2nd ed.). Mahwah, NJ: Erlbaum.
Rubio, D. M., Berg-Wegner, M., & Tebb, S. S. (2001). Using structural equation modeling to test for multidimensionality. *Structural Equation Modeling*, *8*, 613–626.
Shapiro, J. R., & Anderson, D. A. (2003). The effects of restraint, gender, and body mass index on accuracy of self-reported weight. *International Journal of Eating Disorders*, *34*, 177–180.
Soet, J. E., Dudley, W. N., & Dilorio, C. (1999). The effects of ethnicity and perceived power on women's sexual behavior. *Psychology of Women Quarterly*, *23*, 707–723.
Stice, E. (2002). Risk and maintenance factors for eating pathology: A meta-analytic review. *Psychological Bulletin*, *128*, 825–848.
Stormer, S. M., & Thompson, J. K. (1996). Explanations of body image disturbance: A test of maturational status, negative verbal commentary, social comparison, and sociocultural hypotheses. *International Journal of Eating Disorders*, *19*, 193–202.
Strahan, E. J., Wilson, A. E., Cressman, K. E., & Buote, V. M. (2006). Comparing to perfection: How cultural norms for appearance affect social comparisons and self-image. *Body Image*, *3*, 211–227.
Striegel-Moore, R. H., Silberstein, L. R., & Rodin, J. (1986). Toward an understanding of risk factors for bulimia. *American Psychologist*, *41*, 246–263.
Suls, J., Martin, R., & Wheeler, L. (2002). Social comparison: Why, with whom, and with what effect? *Current Directions in Psychological Science*, *11*, 159–163.
Thompson, J. K., Heinberg, L. J., & Tantleff, S. (1991). The Physical Appearance Comparison Scale (PACS). *Behavior Therapist*, *14*, 174.
Tiggemann, M., & McGill, B. (2004). The role of social comparison in the effect of magazine advertisements on women's mood and body dissatisfaction. *Journal of Social and Clinical Psychology*, *23*, 23–44.
Trottier, K., Polivy, J., & Herman, C. P. (2007). Effects of exposure to thin and overweight peers: Evidence of social comparison in restrained and unrestrained eaters. *Journal of Social and Clinical Psychology*, *26*, 155–172.
Wasilenko, K. A., Kulik, J. A., & Wanic, R. A. (2007). Effects of social comparisons with peers on women's body satisfaction and exercise behavior. *International Journal of Eating Disorders*, *40*, 740–745.
Wheaton, M. G., Berman, N. C., Franklin, J. C., & Abramowitz, J. S. (2010). Health anxiety: Latent structure and associations with anxiety-related psychological processes in a student sample. *Journal of Psychopathology and Behavioral Assessment*, *32*, 565–574.
Wilson, G. T., Schmidt, U., Nicholls, D. E., & Johnson, C. (2012, May). The good, the bad and the ugly of integrating clinical practice and research. In V. Mountford & T. J. Raney (Eds.), *Plenary conducted at the meeting of the Academy for Eating Disorders* Austin, TX.
Wood, J. V. (1996). What is social comparison and how should we study it? *Personality and Social Psychology Bulletin*, *22*, 520–538.
Worthington, R. L., & Whittaker, T. A. (2006). Scale development research: A content analysis and recommendations for best practices. *The Counseling Psychologist*, *34*, 806–838.