Journal of Sport & Exercise Psychology, 2021, 43, 223-233
https://doi.org/10.1123/jsep.2019-0213

Human Kinetics
ORIGINAL RESEARCH

The Development and Psychometric Evaluation of the Exercise Overvaluation Scale

Mary Page Leggett-James,[1] Matthew E. Vanaman,[2] Danielle Lindner,[3] and Robert L. Askew[3]
[1]Florida Atlantic University; [2]City University of New York; [3]Stetson University

While regular exercise is associated with a number of physical and mental health benefits, basing one's self-esteem largely on exercise is likely associated with negative outcomes. In the present studies, the authors developed a novel measure of this construct, something they term "exercise overvaluation." In Study 1, 820 participants completed an online survey measuring self-esteem, exercise attitudes and behaviors, and eating disorder symptoms. Exploratory and confirmatory factor analysis were employed to develop the 14-item Exercise Overvaluation Scale. The results provided evidence of discriminant and convergent validity and internal consistency reliability of scale scores. In Study 2, the Exercise Overvaluation Scale was administered to 134 university athletes, including those who participated in intramural sports, club sports, and collegiate athletics. The results from Study 2 supported the criterion validity and test–retest reliability of scale scores. This scale offers researchers a new tool to help understand the relationships among exercise, self-esteem, and physical and mental health outcomes.

Keywords: athletes, measurement, physical activity, self-esteem

Regular exercise, whether in the form of organized sports or other endeavors, is related to a number of positive outcomes for physical and mental health, including improved cardiovascular health, reduced depression and anxiety, and improved self-esteem (see report by the United States Department of Health and Human Services, 2018, for a review). However, when exercise behavior becomes excessive, when people engage in exercise purely for appearance- or weight-related reasons, or when people feel extreme pressure to perform, some of the observed benefits of exercise may decrease (e.g., DiBartolo et al., 2007; Strelan et al., 2003). It may also be that some of the benefits of exercise are undermined when an individual bases their self-esteem largely on exercise behavior and performance, something we term *exercise overvaluation*. The purpose of this paper is to describe the development and psychometric evaluation of a scale designed to measure exercise overvaluation, with the goal of providing researchers with an instrument that can be used to better elucidate the relationships between exercise and physical and psychological outcomes. Better understanding the relationship between exercise and self-esteem, including the point at which one's investment in exercise becomes problematic, could help researchers identify the kinds of exercise/sports and/or coaching practices that foster a healthy relationship with oneself and one's body. This knowledge could then also be applied to develop training programs and interventions for those participating in sports or coaching sports in order to maximize sports participation's benefits to self-esteem.

Defining Exercise Overvaluation

As previously described, we defined exercise overvaluation as the extent to which one's self-esteem is determined by their exercise behavior and athletic performance. An individual low in exercise overvaluation might be disinterested in exercise altogether, or they might value exercise in addition to other domains (e.g., school, relationships, etc.). In contrast, an individual high in exercise overvaluation bases too much of their self-esteem on exercise and exercise performance and, as a result, may engage in excessive or obligatory exercise or disordered eating, or may experience other negative physical and mental health consequences. Exercise overvaluation is not specifically about the different factors that motivate someone to exercise; instead, it is about the extent to which one's sense of self-esteem is dependent on exercise. Our conceptualization of exercise overvaluation was informed by the available research on self-esteem, body image, and existing theories and measures of motivation for exercise.

Most researchers who study self-esteem focus on *global self-esteem*, defined as an individual's overall opinion of themselves or general judgment of self-worth (Crocker & Wolfe, 2001; Rosenberg et al., 1995). *Specific self-esteem* refers to an individual's evaluation of themselves in a given domain (e.g., athletic, academic, etc.), whether positive or negative (Rosenberg et al., 1995). Global self-esteem and specific self-esteem are often only moderately correlated, and global self-esteem is related to psychological well-being, while specific self-esteem predicts behavioral outcomes related to the domain of interest. Rosenberg et al. (1995) argued that the behavioral outcomes of specific self-esteem depend upon the importance of the domain of interest to the individual. According to this argument, if an individual bases more of their self-esteem on exercise, they may engage in exercise behavior to maintain a positive self-view. This theoretical approach, however, does not account for potentially negative behavioral outcomes that could arise from placing too much emphasis on exercise in self-evaluation. In 2009, Holm-Denoma et al. (2009) examined undergraduate athletes, including collegiate varsity athletes, club athletes, independent exercisers, and nonexercisers, for eating disorder symptoms and body image concerns. The findings indicated that women athletes who participated in regular exercise reported higher rates of eating disorder symptoms than nonexercisers.

Leggett-James is with Florida Atlantic University, Boca Raton, FL, USA. Vanaman is with the City University of New York, New York, NY, USA. Lindner and Askew are with Stetson University, DeLand, FL, USA. Lindner (dlindner@stetson.edu) is corresponding author.

Moreover, a meta-analysis of female athletes indicated that athletes report high drive for thinness and more body dissatisfaction than nonathletes (Smolak et al., 2000). Previous studies suggest that collegiate athletes have many protective factors, such as regular exercise, increased global self-esteem, and social support, which reduce the risk of depression symptoms (e.g., Dishman et al., 2006; Morgan, 2000; Williams & Galliher, 2006). However, when athletes derive too much of their self-esteem from exercise, some of these benefits may be reduced. Taken together, these findings suggest that measuring only global self-esteem leads researchers to miss valuable opportunities to better understand the relationships between sport/exercise participation and behavioral outcomes.

In contrast to specific self-esteem, Deci and Ryan (1995) defined *contingent self-esteem* as "feelings about oneself, which are dependent on some standard of interpersonal or intrapsychic expectations" (p. 32). In other words, individuals with high contingent self-esteem are concerned with how they measure up on a specific criterion. When performance in a specific domain is poor, individuals with high contingent self-esteem will go to great measures to avoid feelings of incompetence (Deci & Ryan, 1995). For example, they may distort their memory of their performance to avoid negative feelings. Similarly, Paradise and Kernis (1999) believed that self-esteem was contingent upon meeting the expectations of self and others and argued that contingent self-esteem may be adaptive in some circumstances and problematic in others. To examine the intensity of contingent self-esteem, they developed the Contingent Self-Esteem Scale. Basing self-esteem on domains where one has achieved a level of competence can motivate behavior and have a positive impact on global self-esteem, and having a broad range of domains from which one can draw self-esteem can provide a buffering effect (i.e., when performance or feedback in one area is not positive, positive performance or feedback in other areas may mitigate the effects of negative feedback). However, basing self-esteem on domains where one is not competent, on domains that depend on approval from others, or on domains that involve superficial aspects of the self (e.g., appearance) likely have a negative impact on global self-esteem. While the Contingent Self-Esteem Scale measures the intensity of contingent self-esteem, it does not measure specific domains (e.g., academics, appearance, and exercise) that an individual will describe as important to their self-esteem. Moreover, there are undoubtedly unique characteristics to individual domains, like exercise, and incorporating these unique experiential facets would enhance validity and precision in measurement.

To address this gap, Crocker and Wolfe (2001) introduced the term "contingencies of self-worth" to describe the domains upon which self-esteem is based, and they argued that the nature of these domains, including the types and relative importance, varies by the individual and is shaped through a variety of sources. They noted that "social groups, institutions, and families may have implicit or explicit criteria regarding what makes a person worthwhile or valuable" (p. 613), and people will often internalize the standards of the groups to which they belong. Crocker et al. (2003) developed the Contingencies of Self-Worth Scale to assess these domains upon which people base their self-esteem. The Contingencies of Self-Worth Scale examines the domains of family support, competition, appearance, God's love, school competence, virtue, and approval from others. Absent is a domain related to athletics and exercise, an area upon which those who have learned that exercise or their sports performance is important, perhaps through regular exercise or participation in a sport, may base their self-esteem. While the appearance (e.g., "My sense of self-worth suffers whenever I think I don't look good") and competition (e.g., "My self-worth is influenced by how well I do on competitive tasks") may be related to one's sense of self-worth around exercise or sports participation, none of the Contingencies of Self-Worth Scale's items ask specifically about sports and exercise.

Kernis (2003) suggested that individuals with fragile self-esteem enhance their self-esteem by participating in activities they feel competent in. Their competency in the activity promotes positive self-feelings and replenishes their self-esteem. Crocker and Wolfe (2001) suggested that basing self-esteem on several distinct domains may provide a buffering effect, but that problems may arise when an individual is overly invested in one particular domain, especially a domain in which the contingency is difficult to satisfy or the standard is unrealistic. The research on the *overvaluation of weight and shape* provides an excellent example of this phenomenon. According to the transdiagnostic model of eating disorders (Fairburn, 2008), overvaluation of weight and shape is the core psychopathology behind all eating disorders, and it is defined as "the judging one's self-worth largely, or even exclusively, in terms of shape and weight and the ability to control them" (p. 11). In other words, Fairburn suggested that basing one's self-esteem on the single domain of weight, shape, and control leads to disordered eating behavior. High levels of overvaluation of weight and shape or investment in appearance are found in samples of eating disorder patients, and in nonclinical samples, overvaluation of weight and shape is correlated with disordered eating behavior (e.g., Hrabosky et al., 2009; Woodward et al., 2014). It may be that exercise overvaluation functions similarly to overvaluation of weight and shape, such that those who place undue emphasis on exercise performance or behavior in their self-evaluation experience mental health consequences. These consequences may be specific to exercise, or, given the relationship between body image, exercise, and self-esteem, may also be related to body image disturbance and disordered eating. A scale designed to measure exercise overvaluation would help researchers and clinicians determine the point at which investment in an exercise becomes problematic, increasing the risk of the aforementioned consequences.

By definition, someone high in exercise overvaluation derives their self-esteem predominately from engaging in exercise. Other theorists have studied similar concepts. For example, self-determination theory (SDT; Deci & Ryan, 1985) argues that people differ in the extent to which they are motivated intrinsically, by internal factors, and extrinsically, by external factors. While on the surface intrinsic motivation seems the more desirable or adaptive option, SDT argues that some forms of extrinsic motivation are also useful, especially those in which the external forces come from an entity with which the individual has a high level of commitment or identification. However, while SDT helps to explain why people might be motivated to engage in exercise, exercise overvaluation is unique in that it assesses the degree to which one's exercise behavior contributes to their self-esteem. A person high in exercise overvaluation might be more motivated to exercise, specifically with the goal of maintaining a positive self-view. A person with a low-to-moderate level of exercise overvaluation might also be motivated to exercise, but perhaps in a less driven or obsessive way, given that, by definition, someone low in exercise overvaluation would not base their self-esteem largely on exercise.

Beyond SDT, several researchers have attempted to measure attitudes toward and/or motivations for exercise. For instance, the Obligatory Exercise Questionnaire (OEQ; Pasman & Thompson,

1988) was developed to quantify behaviors and attitudes related to patterns of excessive and/or overly rigid exercise. There are conceptual similarities between obligatory exercise as measured by the OEQ and exercise overvaluation; however, they are conceptually distinct in that the OEQ focuses primarily on behavioral and/or emotional symptoms of obligatory exercise (e.g., "Sometimes, I feel a need to exercise twice in one day, even though I may feel a little tired," "When I don't exercise I feel guilty"), while exercise overvaluation focuses more on the role of exercise in determining self-esteem. While we certainly expect that someone driven to engage in exercise in a pattern that would result in high OEQ scores would also report high levels of exercise overvaluation, the OEQ does not include items assessing the extent to which one bases their self-esteem on exercise.

Another measure related to exercise behavior is the Exercise Identity Scale (EIS; Anderson & Cychosz, 1994), which was developed to assess the degree to which an individual's participation in exercise was descriptive of their self-concept. While there are similarities between the EIS and exercise overvaluation, the EIS emphasizes an individual identifying with the role of exerciser and how that individual's behaviors attempt to fulfill that role (e.g., "Physical exercise is a central factor to my self-concept," "For me, being an exerciser means more than just exercising"). As part of its assessment of role-identity, the EIS does not include items assessing the extent to which one bases their self-esteem on exercise. Someone who identifies as an exerciser may report a high or low level of exercise overvaluation, depending on the role exercise plays in their life. For instance, someone who exercises frequently for enjoyment or health reasons, rather than for the purposes of maintaining self-esteem, would likely report a low-to-moderate level of exercise overvaluation and few (if any) negative mental or physical health outcomes, but someone who identifies as an exerciser for the purposes of maintaining self-esteem or as part of underlying psychopathology (e.g., an eating disorder) would likely report a higher level of exercise overvaluation. The measure we developed would allow for a more nuanced exploration of these relationships.

Self-Esteem, Exercise, and Disordered Eating

As highlighted previously, there are no existing measures designed to assess the extent to which someone derives their self-esteem primarily from their exercise performance. Thus, it is unsurprising that the available literature on self-esteem and exercise focuses on the relationship between global self-esteem and exercise. A systematic review of studies examining the relationship between global self-esteem and exercise in adults showed a small but statistically significant relationship between engaging in exercise and improved self-esteem (Spence et al., 2005). Other studies of adult and college student samples reported similar findings, with engagement in even relatively small amounts of exercise associated with modest increases in self-esteem and improved quality of life (e.g., Kell & Rula, 2019; Li et al., 2014; Moore et al., 2007). The relationship between exercise and self-esteem has also been documented in athletes. In general, athletes report higher self-esteem than nonathletes (Armstrong & Oomen-Early, 2009). In a study of Spanish adults, not only did athletes report greater self-esteem than nonathletes, but athletes who participated in individual sports reported greater self-esteem than those who participated in team sports, regardless of gender (Laborde et al., 2016). This may demonstrate the effect of teammate competition on self-esteem. Team sport athletes not only compete with their opponents during competition, but also compete with their teammates for starting positions, which may place a larger emphasis on their exercise perceptions. For example, on some soccer teams, there are two starting spots as a striker. That same team may have four or five athletes who play in the striker position. The athletes' starting spot (one of the two strikers) is based on their weekly practice, weightlifting workouts, speed, and agility training sessions with the team. Because all of these forms of exercise are critical to gaining a starting position, team sport athletes may place a larger emphasis on exercise when reporting about their self-esteem.

Though much of the literature suggests a positive relationship between global self-esteem and exercise, there are exceptions. Several studies have shown no relationship between exercise and self-esteem, including in Norwegian adults participating in CrossFit (Köteles et al., 2016), adult women (Levy & Ebbeck, 2005), and male and female college students (e.g., Russell, 2002; Russell & Cox, 2003). Two other studies suggested that higher self-esteem was related to maladaptive exercise behaviors, with global self-esteem predicting symptoms of exercise addiction in a sample of college students after controlling for body shame (Ertl et al., 2018) and higher self-esteem predicting excessive exercise in patients with anorexia nervosa (Bewell-Weiss & Carter, 2010). Still, other studies suggest that the relationship between exercise and self-esteem may differ by gender. In a study that followed women and men over 3 weeks, high levels of self-reported exercise were associated with increased self-esteem and decreased depression in women, but with decreased self-esteem and increased depression in men (Joiner & Tickle, 1998). A study of Spanish college students showed that, while both women and men who were more active reported greater vitality, higher self-esteem was observed for men only (Molina-Garcia et al., 2011). Another study showed higher self-esteem in men who exercised, but women did not report higher self-esteem and, instead, reported higher levels of body image disturbance and disordered eating. It is important to note, however, that these studies employed global measures of self-esteem, rather than domain-specific measures.

Taking other variables into account may help to explain why some studies show differing relationships between exercise and self-esteem. One such variable is the reason exercise occurs. Several studies have shown that exercising for weight or appearance-related reasons is associated with lower self-esteem (e.g., Maltby & Day, 2001; Strelan et al., 2003). Some studies have also shown that weight or appearance-related motives for exercise are associated with body image disturbance and disordered eating behavior, and this is true of both women and men (e.g., DiBartolo et al., 2007; Gonçalves & Gomes, 2012; Grieve et al., 2008; Hubbard et al., 1998). The extent to which one bases their self-esteem on exercise as measured by the Exercise Overvaluation Scale (EOS) may also help to explain some of the differing relationships between exercise, self-esteem, and other variables related to psychological well-being. Specifically, it may help researchers determine the point at which valuing or being invested in exercise is maladaptive for some individuals.

The Current Studies

Taken together, the existing literature on exercise and well-being suggests that exercise is generally related to improved self-esteem, but that there are circumstances under which this is not the case. Some people may grow to develop a sense of moderate investment

in exercise as part of their self-esteem through experiences with sports and athletic participation. For these individuals, participating in exercise may contribute to a more positive self-view. Other individuals may become overly invested in exercise as the primary determinant of their self-esteem, something that may be linked to outcomes like excessive exercise behavior and disordered eating behavior. The purpose of the current studies was to develop and psychometrically evaluate the EOS as a tool for understanding the extent to which one's view of exercise is related to their well-being in positive and negative ways. Understanding and measuring this construct would facilitate a more meaningful exploration of the relationships between body image, disordered eating, exercise, and self-esteem in athletes and nonathletes alike.

Study 1

The aims of Study 1 were to identify manifestations of exercise overvaluation, develop candidate items to reflect exercise overvaluation, and identify a subset of items with sufficient common covariance among item responses to enable valid and reliable measurement. Low correlations between scores derived from these items and scores on measures of global self-esteem and state self-esteem were expected to serve as discriminant validity evidence. Moderate to high correlations with contingent self-esteem, eating disorder symptoms, and obligatory exercise were expected to serve as convergent validity evidence.

Method

Participants. A total of 1,053 participants were recruited from Amazon's Mechanical Turk; 169 were excluded for failing attention checks, and 64 were excluded for incomplete responses. The participants were required to be 18 years of age or older, reside in the United States, and speak English. The remaining 820 participants were then randomly assigned into either the developmental or the validation data sets to enable cross-validation of the dimensional structure of item responses. The demographic characteristics of each sample are presented in Table 1. The average age of the participants was 36.80 years ($SD = 12.20$), and the majority were female (55.0%; $n = 451$) and White (79.9%; $n = 655$). Twelve percent ($n = 99$) reported being a member of a sports team.

Procedure. All study procedures were approved by the university's institutional review board. Informed consent was obtained, and the participants completed an online battery composed of demographic information, exercise habits (i.e., minutes of exercise per week), the candidate items for the EOS, and measures of self-esteem, disordered eating, and obligatory exercise. All survey data were collected and stored anonymously.

Measures. *Exercise overvaluation:* A phenomenological approach to item development was adopted by the coauthors, including one clinical psychologist content expert to maximize content validity. Semistructured interviews were carried out with a convenience sample of 10 student-athletes (seven women and three men; mean age = 20.0, $SD = 1.1$), including five National Collegiate Athletic Association Division I student-athletes, two club sports athletes, and three nonsport interviewees. The athletes reported a mean of 418.0 min of exercise per week ($SD = 25.5$). The interviews included questions related to global self-esteem (e.g., "How do you know your self-esteem is high?") and the relationship between self-esteem and exercise (e.g., "Describe how exercise changes when you're with others versus by yourself" and "If you haven't exercised in a while, describe how you would feel"). After the interviews, a thematic content analysis was carried out separately by the coauthors, with the first and second authors each conducting their own round of open-coding on the transcripts to identify emergent themes. Discrepancies in themes were resolved through discussion between the authors until consensus was achieved. Next, the first and second authors independently conducted a round of thematic coding, sorting individual comments under each theme. As with theme identification, discrepancies in the thematic sorting of comments were resolved through consensus-based discussion. This process yielded six themes: confidence, competition, feedback from others, self-image, reliance on routine, and negative affectivity. Supplementary Material 1 (available online) presents each of the themes, example comments that informed thematic identification, and the number of comments coded under each theme. Guided by the themes, a total of 27 candidate items were developed for pilot testing (See Supplementary Material 2 [available online] for the list of candidate items). Each of the items was then reviewed for content validity by a content expert (third author), who was blind to the item generation process. Considerations included face validity and content representativeness of each item to its respective theme as well as to the exercise overvaluation construct more broadly. Responses to the items ranged from 1 (*strongly disagree*) to 4 (*strongly agree*).

Global self-esteem: Rosenberg's Generalized Self-Esteem Scale was used to assess global self-esteem (Rosenberg, 1965). The scale includes 10 statements, with response options ranging from 1 (*strongly agree*) to 5 (*strongly disagree*). Examples of items include "On the whole I am satisfied with myself" and "I am able to do things as well as most other people." Higher sum scores indicate better self-esteem. Internal consistency reliability for the present study was high ($\alpha = .92$).

State self-esteem: The State Self-Esteem Scale measures an individual's self-esteem at a given time point (Heatherton & Polivy, 1991). The scale consists of 20 items prompting participants to rate how true the statement is of them right now. The State Self-Esteem Scale is composed of three subscales: performance self-esteem, social self-esteem, and appearance self-esteem. Ratings are based on a 5-point Likert-type scale ranging from 1 (*not at all*) to 5 (*extremely*). Examples of items include "I am worried about whether I am regarded as a success or a failure" and "I am pleased with my appearance right now." Sum scores are calculated for each subscale, and higher scores indicate higher self-esteem. Internal consistency reliability for the present study was high for scores from the composite scale ($\alpha = .93$) and the performance ($\alpha = .83$), social ($\alpha = .89$), and appearance ($\alpha = .86$) subscales.

Contingent self-esteem: The Contingent Self-Esteem Scale assesses an individual's self-esteem based on reflected appraisals (i.e., what respondents think other people think of them; Paradise & Kernis, 1999). The scale prompts participants to rate the extent to which each of 15 items are true of them on a 5-point Likert scale ranging from 1 (*not at all*) to 5 (*extremely*). Example items include "Even in the face of failure, my feelings of self-worth remain unaffected" and "If I get along well with somebody, I feel better about myself overall." Higher sum scores indicate higher contingent self-esteem. Internal consistency reliability for the present study was high ($\alpha = .88$).

Disordered eating: The Eating Disorder Examination-Questionnaire (EDE-Q) assesses the features and symptoms of eating disorders that individuals reported within the last 4 weeks (Fairburn & Beglin, 1994). The EDE-Q includes 28 questions assessing

Table 1 Demographics

	Combined, $N = 820$		Developmental sample, $n = 409$		Validation sample, $n = 411$		Test–retest sample, $n = 134$	
Demographic characteristic	**Measure**	**%**	**Measure**	**%**	**Measure**	**%**	**Measure**	**%**
Mean age in years (*SD*)	36.8 (12.2)		37.3 (12.3)		36.2 (12.0)		19.7 (2.08)	
Median age in years (range)	34 (18–79)		34 (18–76)		33 (19–79)		(18–23)	
Missing	0		0		0			
Gender								
Female	451	55.0	230	56.2	221	52.8	63	47
Male	367	44.8	177	43.3	190	46.2	71	53
Other	2	0.2	2	0.5	0			
Race								
Asian/Pacific Islander	54	6.6	18	4.4	36	8.8	4	3
African descent/Black	63	7.7	32	7.8	31	7.5	17	12.7
Native American	15	1.8	8	2.0	7	1.7	1	0.7
Caucasian/White	655	79.9	336	82.2	319	77.6	104	77.6
Other	33	4.0	15	3.7	18	4.4	8	6
Ethnicity								
Hispanic/Latino	89	10.9	42	10.3	47	11.4	11	8.2
Not Hispanic/Latino	731	89.1	367	89.7	364	33.6	121	90.3
Member of a sports team								
Yes	99	12.1	48	11.7	51	12.4		
No	721	87.9	361	88.3	360	87.5		
Mean minutes of exercise (*SD*)	182.8 (198.3)		187.8 (243.0)		177.8 (140.6)		659.16 (448.23)	
Median minutes (range)	120 (0–4,000)		120 (0–4,000)		120 (0–840)		(0–2,500)	
Missing	1		1		0			
Year in school								
Freshman							36	26.9
Sophomore							38	28.4
Junior							31	23.1
Senior							27	20.1
Missing							2	1.5
Level of competition								
Collegiate							96	71.6
Club/Intramural							38	28.4
Mode of participation								
Team sports							87	64.9
Solo sports							47	35.1

eating disorder symptoms within the past 28 days. The response items range from "no days" to "every day." Examples of items include "Have you been deliberately trying to limit the amount of food you eat to influence your shape or weight (whether or not you have succeeded)?" and "Over the past 28 days, how many times have you eaten what other people would regard as an unusually large amount of food (given the circumstances)?" There are four subscales, including restraint, eating concern, shape concern, and weight concern. The mean subscale scores are obtained, and an overall score is calculated by averaging the subscale scores. Higher scores indicate more severe eating disorder symptoms. Internal consistency reliability for the present study was high for scores from the composite scale ($\alpha = .94$) and the restraint ($\alpha = .83$), eating ($\alpha = .84$), shape ($\alpha = .90$), and weight ($\alpha = .84$) concerns subscales.

Obligatory exercise: The OEQ measures athletes' feelings about needing to exercise (Pasman & Thompson, 1988). The scale consists of 20 items where participants indicate how often they engage in each behavior. The items are based on a 4-point scale ranging from 1 (*never*) to 4 (*always*). Examples of items include "I engage in exercise on a daily basis," and "I will engage in other forms of exercise if I am unable to engage in my usual form of exercise." Higher sum scores indicate a higher obligation to exercise feelings and behaviors. Internal consistency reliability for the present study was high ($\alpha = .87$).

Data Analysis. All statistical analyses were carried out with STATA for Mac (version 13; StataCorp, 2013) and M*plus* for Mac (version 7.11; Muthén & Muthén, 1998–2012). One-half of the respondents from the Mechanical Turk sample were randomly assigned to a developmental data set to facilitate identification of the latent factor structure using principal component analysis and exploratory factor analysis. The remaining participants were

assigned to a validation data set to enable cross-validation of the factor structure using confirmatory factor analysis. Identification of the factor structure was guided by Kaiser's (i.e., eigenvalues greater than one) and Cattell's (i.e., the elbow rule) criteria, factor loadings, and interpretability of factor solutions. Upon identification of the optimal factor solution, a confirmatory factor analysis was carried out in the validation data set with a geomin rotation. Traditional fit indices guided evaluations of model fit (Hu & Bentler, 1999), and all factor loadings and residual correlations were inspected to identify potentially problematic local dependency. Correlations were estimated between scores reflecting overvaluation of exercise and scores from theoretically related measures of general self-esteem, state self-esteem, contingent self-esteem, eating disorder symptoms, and obligation to exercise.

Results

The principal component analysis carried out on the developmental sample indicated a minimum of two components using Cattell's criterion and a maximum of four using the Kaiser criterion (Figure 1). The two-factor solution of the exploratory factor analysis was the most substantively interpretable using a factor loading threshold of 0.65 or higher for indicators of latent factors (Table 2). The first factor consisted of nine items that reflect internal standards, or the extent to which an individual's feelings about themselves are governed by standards that they have set for their exercise behavior. The second factor consisted of five items reflecting external standards, or the extent to which an individual's feelings about themselves are governed by external benchmarks for exercise behavior, whether real or imagined. A total of nine items were removed for weak factor loadings (i.e., <.65), three items were removed for content redundancy, and one item was removed for both a weak factor loading and content redundancy. Fit indices for the two-factor confirmatory factor analysis conducted with the validation sample indicated that the data fit the two-factor model adequately (comparative fit index = .92, Tucker–Lewis index = .90, and root mean square error of approximation = .13), with a correlation of .61 between factors. Of the 91 residual correlations, only three exceeded .20 with a maximum residual correlation of −.25. Internal consistency reliability was high for the composite scores ($\alpha = .89$) and the internal standards ($\alpha = .89$) and external standards ($\alpha = .81$) subscale scores. The EOS composite and subscale scores range from 1 to 4. The mean score on the internal standards subscale was 2.59 ($SD = .61$), the mean score on the external standards subscale was 2.04 ($SD = .62$), and the mean composite score on the EOS was 2.32 ($SD = .53$).

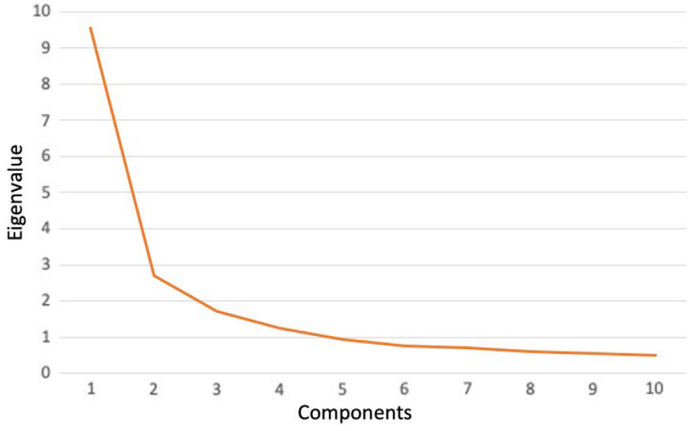

Figure 1 — Components of exercise overvaluation.

Table 3 presents correlations between the EOS (composite score and scores on the internal and external standards subscales) and measures of theoretically related and unrelated constructs as validity evidence. Discriminant and convergent validity evidence were strong, as indicated by correlations between composite scores reflecting overvaluation of exercise and scores reflecting generalized self-esteem ($r = -.14$, $p < .001$), state self-esteem ($r = -.25$, $p < .001$), contingent self-esteem ($r = .33$, $p < .001$), eating disorder symptoms ($r = .31$, $p < .001$), and obligatory exercise ($r = .69$, $p < .001$). To assess whether measurement error may have attenuated observed relationships among the EOS and measures of convergent and discriminant validity, we also calculated attenuation-corrected coefficients using the standard double-correction method outlined by Muchinsky (1996). The differences between attenuated and unattenuated correlations did not suggest substantively different conclusions (Supplementary Material 3 [available online]); indeed, the largest difference between uncorrected and corrected correlations was +0.1. Finally, we also investigated whether scores from the EOS were linearly related to scores from the other measures through a visual inspection of scatterplots, which did not reveal any nonlinear trends (see Figure 1 in Supplementary Material 4 [available online]).

Discussion

In Study 1, we sought to develop the EOS and assess initial evidence for the reliability and validity of scale scores. This study yielded a 14-item measure of exercise overvaluation with two subscales, internal standards and external standards. Items on the internal standards subscale are related to one's own standards for their exercise behavior as a part of their self-esteem (e.g., "Working out is a part of who I am" and "I become very upset if I cannot exercise on a day I was planning to"). The external standards subscale items are related to judgments against some kind of external standard, whether real or imagined, and comparisons based on one's workout quality (e.g., "I would worry what other people think about me if I missed a workout" and "I keep track of how well my workout is going compared to others"). Initial psychometric evaluation provided evidence for the internal consistency reliability and validity of test scores. The emergence of the internal and external domains in Study 1 likely has important implications for well-being. In this study, scores on both the internal and external standards subscales were correlated with obligatory exercise, but, as shown in Table 3, the correlation was much stronger for the internal standards subscale. Correlations between the internal standards subscale and the EDE-Q subscales were small, with the exception of the restraint subscale ($r = .26$), whereas correlations between the external standards subscale and the EDE-Q Eating Concern, Shape Concern, and Weight Concern subscales were moderate in strength.

Study 2

The aims of Study 2 were to provide additional reliability and validity evidence of EOS scores using a sample of undergraduate athletes. We hypothesized that EOS scores would be stable over time for individuals who reported no change in the way their exercise habits made them feel about themselves, providing test–retest reliability evidence of EOS scores. We also hypothesized that

Table 2 Exercise Overvaluation Scale Items and Factor Loadings for Study 1 and Study 2

		Study 1 (EFA)		Study 2 (CFA)	
No.	Item	Factor 1	Factor 2	Factor 1	Factor 2
2	My day is only complete if I have a good workout.	**.77**	–.11	**.75**	
10	Working out is a part of who I am.	**.83**	–.29	**.75**	
14	Exercise is critical to my self-esteem.	**.74**	.01	**.80**	
16	It's important to me to spend a certain amount of time exercising each week.	**.87**	–.37	**.54**	
18	I get frustrated when I cannot follow my workout schedule.	**.80**	.01	**.78**	
19	I become very upset if I cannot exercise on a day I was planning to.	**.74**	.10	**.83**	
20	I have a strict workout schedule.	**.76**	–.14	**.67**	
25	Exercise is an important measure of my self-worth.	**.66**	.13	**.78**	
27	I get anxious when I don't have time for a full workout.	**.69**	.17	**.82**	
4	I feel worse about myself when I notice someone having a better workout than me.	–.13	**.88**		**.74**
3	I have less respect for myself after a bad workout.	.10	**.69**		**.74**
12	I tend to notice the quality of other people's workouts.	.00	**.69**		**.76**
13	I keep track of how well my workout is going compared to others.	–.01	**.74**		**.86**
26	I would worry what other people think about me if I missed a workout.	.08	**.69**		**.70**

Note. EFA = exploratory factor analysis; CFA = confirmatory factor analysis. Bolded factor loadings denote factor assignment.

Table 3 Correlations Between the Exercise Overvaluation Scale and Measures of Convergent and Discriminant Validity

Measure	Internal standards	External standards	Composite
GSES	.03	–.27***	–.14***
SSES	–.07*	–.35***	–.25***
Performance	–.05	–.39***	–.26***
Social	–.17***	–.41***	–.34***
Appearance	.06	–.09*	–.01
CSES	.24***	.32***	.33***
EDE-Q	.21***	.32***	.31***
Restraint	.26***	.23***	.28***
Eating concern	.14***	.32***	.27***
Shape concern	.17***	.27***	.26***
Weight concern	.19***	.31***	.29***
OEQ	.72***	.47***	.69***

Note. GSES = Rosenberg Generalized Self-Esteem Scale; SSES = State Self-Esteem Scale; CSES = Contingent Self-Esteem Scale; EDE-Q = Eating Disorder Examination-Questionnaire; OEQ = Obligatory Exercise Questionnaire.
*$p < .05$. ***$p < .001$.

collegiate athletes would have higher EOS scores than casual sports athletes, given that collegiate athletes are likely to have a sense of competence about their sport and that collegiate athletes may face unique pressures to achieve a high level of performance that may lead them to develop overvalued ideas about exercise. Finally, we hypothesized team sports athletes would have higher EOS scores than solo sports athletes, based on the small negative correlation between global self-esteem and EOS observed in Study 1 and because team sport athletes have been shown to have lower general self-esteem than solo sport athletes (Laborde et al., 2016). Results consistent with these hypotheses would provide concurrent criterion validity evidence of EOS scores.

Method

Participants. A total of 134 participants were recruited from a small liberal arts university in the southeastern United States via convenience sampling. Preliminary power analysis was carried out using G*Power: Statistical Power Analyses for Windows and Mac (version 3.1.9.2; Buchner et al., 2007), which indicated that a minimum of 128 participants were needed to detect statistically significant moderate effects with alpha = .05 and power = .80. The participants came from National Collegiate Athletic Association Division I collegiate sports teams, club sports teams, and intramural sports teams. They were required to be 18 years of age or older, reside in the United States, and speak English. Collegiate sport team coaches, club sports presidents, and intramural sports captains were emailed invitations to participate, and when a team opted to participate, a meeting was scheduled with the team to obtain informed consent and collect data. The average age of the participants was 19.7 years ($SD = 2.1$). The majority were male (53.0%; $n = 71$) and White (77.60%; $n = 104$). With respect to the level of competition, 96 (71.64%) were members of a collegiate sports team, and 38 (28.36%) were members of a casual sports team (i.e., club and/or intramural). With respect to the mode of participation, 87 (64.90%) participants reported being a member of a team sport (e.g., soccer, football, and lacrosse), with the remainder (35.10%; $n = 47$) reporting membership in a solo team sport (e.g., cross country, tennis). On average, the participants self-reported 659.2 min ($SD = 448.2$) of exercise per week, which amounts to approximately 11 hr. The demographic characteristics of this sample are presented in Table 1.

Procedure. All study procedures were approved by the university's institutional review board. At the time of participation, the coaches were asked to leave the room. The participants provided informed consent, and then survey packets were distributed that included a demographics questionnaire and the EOS. Upon completion, the participants placed individual surveys into a large envelope to ensure anonymity. Approximately four days later (range: 2–7 days), the participants completed the EOS for a second time, followed by an additional item, asking participants whether

their feelings about their exercise habits were worse, better, or unchanged since the first administration. At the baseline (T0) and follow-up (T1), survey completion took approximately 7 min.

Data Analysis. All statistical analyses were carried out with STATA for Mac (version 13; StataCorp, 2013). In order to evaluate the test–retest reliability of EOS scores, intraclass correlation coefficient (ICC) of ICC(3,1) was selected because it is appropriate for a two-way mixed-effects design (McGraw & Wong, 1996; Shrout & Fleiss, 1979). Differences in EOS scores for the study participants who indicated no change in the way their exercise habits made them feel about themselves between the survey administrations would serve as indicators of measurement error. In order to evaluate concurrent criterion validity, a 2×2 factorial design was selected. The first predictor was the level of competition with two levels (i.e., collegiate or casual), and the second predictor was the mode of participation with two levels (i.e., team sports or solo sports). While data from both the baseline and follow-up assessments were utilized for estimation of the ICC the data from the follow-up assessment were selected for criterion validity so that model estimates were not contaminated with dependency from repeated assessments.

Results

Test–Retest Reliability. The participants completed the EOS at the baseline (T0) and again approximately four days later (T1). At the end of the survey at T1, the participants were asked to compare their feelings about their exercise habits at that exact moment to how they felt at T0. Options included feeling worse at T0 than at T1, better at T0 than T1, or having no change in feeling between the time points. Of the 134 participants who completed Study 2, 81 (60.4%) indicated no change in the way their exercise habits made them feel about themselves. A total of 48 participants indicated feeling better (35.8%), while five participants indicated feeling worse (3.7%). As hypothesized, test–retest reliability was good for composite scores, ICC(3,1) = .90, 95% CI [.84, .93]; internal standards subscale scores, ICC(3,1) = .89, 95% CI [.84, .93]; and external standards subscale scores, ICC(3,1) = .82, 95% CI [.74, .88].

Criterion Validity. A 2×2 factorial analysis of variance was carried out to test differences in EOS scores among four distinct types of athletes (i.e., casual solo sports, casual team sports, collegiate solo sports, and collegiate team sports). The EOS scores were normally distributed at T1, as indicated by histogram inspection. As hypothesized, the EOS composite scores for collegiate sport athletes were higher ($M = 2.74$, $SE = .05$) than for casual athletes ($M = 2.50$, $SE = .10$), and this difference was statistically significant, $F(1, 130) = 7.28$, $p = .008$. With respect to EOS subscales, the differences in internal standards of EOS between collegiate and casual athletes were statistically significant, $F(1, 130) = 5.21$, $p = .02$; collegiate athletes had an average score of 2.65 ($SE = .06$), and casual athletes had an average score of 2.45 ($SE = .11$). The differences in external standards of EOS between collegiate and casual athletes were also statistically significant, $F(1, 130) = 8.05$, $p < .01$. Collegiate athletes had an average score of 2.82 ($SE = .06$), and casual athletes had an average score of 2.55 ($SE = .10$).

Consistent with our hypotheses, the difference in the EOS composite scores for team sports athletes and solo sports athletes was statistically significant, $F(1, 130) = 10.33$, $p < .01$, with team sports athletes scoring higher on the EOS ($M = 2.77$, $SE = .06$) than solo sports athletes ($M = 2.49$, $SE = .08$). Team sports athletes scored higher on internal standards ($M = 2.67$, $SE = .06$) than solo sports athletes ($M = 2.47$, $SE = .09$), and this difference was statistically significant, $F(1, 130) = 5.82$, $p = .02$. The difference in external standards between team and solo sports athletes was also statically significant, $F(1, 130) = 13.64$, $p < .01$. Team sports athletes scored an average of 2.87 ($SE = .06$), and solo sports athletes scored an average of 2.51 ($SE = .09$). No statistically significant interactions between the level of competition and the mode of participation were observed for tests of composite scores, internal standards subscale scores, or external standards subscale scores.

Discussion

Study 2 provided additional evidence of the psychometric properties of the EOS, including test–retest reliability and criterion validity, through the use of known group comparisons. All study hypotheses were supported. While collegiate athletes scored higher on the EOS than casual athletes and team sports athletes scored higher on the EOS than solo sports athletes, the mean scores for all groups of athletes fell between 2.47 and 2.87 on a 4-point scale, corresponding to a moderate level of investment. This moderate level of investment, as stated previously, is expected in those who participate in sports on a regular basis.

General Discussion

The purpose of the present studies was to describe the construct of exercise overvaluation, develop a new instrument to measure it, and evaluate the psychometric performance of scale scores. Exploratory factor analysis in Study 1 yielded a 14-item, two-factor measure of exercise overvaluation. The nine items that reflected an internalized need and desire to exercise to maintain self-esteem were labeled the internal standards subscale, and the five items that reflected the extent to which real or imagined external standards for exercise shape one's self-esteem were labeled the external standards subscale. Correlations between EOS scores and scores on measures of general self-esteem, contingent self-esteem, state self-esteem, eating disorder symptoms, and obligatory exercise provided evidence of the scale's discriminant and convergent validity. Study 2 provided test–retest reliability and concurrent criterion validity evidence in a sample of National Collegiate Athletic Association and casual college student-athletes.

In the current study, exercise overvaluation was weakly correlated with global self-esteem, and correlations with contingent self-esteem were small to moderate. Exercise overvaluation and obligatory exercise were strongly correlated, which is not surprising. Both measures are related to exercise, but the EOS focuses on how someone evaluates themselves in light of their exercise performance, while the OEQ measures exercise behaviors and emotional responses that, in combination with one another, may suggest underlying pathology (e.g., exercising despite an injury, exercising to compensate for eating, and ruminating about exercise; Pasman & Thompson, 1988). Based on item content and the existing literature on excessive or pathological exercise, we would expect that a person high in exercise overvaluation is likely to exercise in a driven manner, as measured by the OEQ. While the OEQ taps into emotional responses to exercise, especially negative emotional responses when one's rules about exercise cannot be followed, it does not directly elicit the ways exercise contributes to self-esteem. The EOS does assess the extent to which one's exercise behaviors contribute to their self-esteem, and thus it can be used together with the OEQ and other measures to shed

important light on the broader psychological processes that underlie observed relationships between exercise and well-being.

The findings from Studies 1 and 2 suggested that our participants reported a moderate level of exercise overvaluation. Of note, the mean scores on all subscales across all samples fell between 2 and 3 on a 4-point scale. The sample in Study 1, which consisted largely of nonathletes, had the lowest scores on the EOS, particularly on the external standards subscale. The samples of athletes in Study 2 reported higher mean scores than in Study 1, and among Study 2 participants, collegiate sports athletes had higher scores on the EOS than casual sports athletes. This suggests that collegiate sports athletes rely on their exercise habits as a larger part of their self-esteem than casual sports athletes, which is consistent with the level of time, energy, and commitment required of students to compete at the collegiate level. Likewise, we found that team sports athletes had higher scores on the EOS than solo sports athletes, suggesting that team sports athletes seem to rely on their exercise habits as a larger part of their self-esteem than solo sports athletes. This difference was especially large for the external standards subscale. It is logical that, in a team sports environment, where the performance of multiple athletes working together is required for a win, athletes may perceive more pressure from others to maintain their exercise routine.

Though we did not specify a factor structure for the EOS a priori, the emergence of the internal and external subscales holds interesting possibilities for researchers who wish to better study exercise attitudes and behavior in athletes and nonathletes alike. The internal standards subscale may help determine the extent of an individual's internalization of their exercise behaviors and performance. When researchers can identify the degree of internalization, it can help to predict if athletes are more likely to engage in the maladaptive exercise, along with other comorbid behaviors (e.g., eating disorder symptoms) to fulfill their self-esteem needs. If an individual reports high levels of exercise overvaluation on the EOS internal standards subscale, practitioners and coaches would have the opportunity to implement prevention and safety programs for the athlete. The external subscale may help determine the extent to which an individual places emphasis on comparison to others during performance. Athletes who report high levels on the EOS external standards subscale may adopt unhealthy performative behaviors to feel better about their workout in comparison with others' workouts. Athletes often experience overuse injuries that result from repeated stress on tissue without allowing for the proper healing time (Hogan & Gross, 2003). Overuse makes up 50% of all athletic injuries (Baquie & Bruker, 1997). Athletes who feel they do not measure up to teammates in practice and, therefore, feel a decrease in their self-esteem, may continue to increase exercise outside of mandatory sessions. The excessive use of their tissues has the potential to result in an overuse injury taking the athlete away from the sport until the tissue heals. For an athlete who reports high levels of exercise overvaluation, this has the potential to negatively impact their self-esteem and result in further psychopathology (e.g., depression, anxiety). In response to high external standards subscale scores, coaches and practitioners would have the opportunity to take preventative measures to reduce risks to the athlete.

In addition to preventing maladaptive behavior and reducing risk, measuring the extent to which people view exercise as a significant contributor to their self-esteem has important implications for understanding how and why people engage in healthy exercise patterns. Developing some level of investment and competence in exercise can, as Crocker and Wolfe (2001) suggested, motivate behavior. This investment may prompt people to engage in physical activity across the lifespan, something that is key for maintaining overall health and aging successfully (United States Department of Health and Human Services, 2018). It may also encourage people to make regular time for exercise, despite having a busy schedule. According to the United States Centers for Disease Control and Prevention (2017), only 23.2% of United States adults meet the recommended guidelines for aerobic and strength-training activity. Understanding the exercise attitudes and habits of those with a moderate level of investment in exercise as measured by the EOS could inform interventions to increase physical activity in the general population. Existing applications of SDT to exercise facilitate this understanding to a certain extent. Ryan and Deci (2000) delineated four regulation styles characteristic of extrinsic motivation, and two of these are most relevant to our work on exercise overvaluation. In introjected regulation, an individual engages in behavior in order to avoid negative emotions, like guilt or anxiety, or to experience positive emotions, like pride. In external regulation, an individual engages in the behavior simply to comply with external standards. These two forms of regulation, particularly in combination with the other forms as described by Ryan and Deci (i.e., integrated regulation and identified regulation) do indeed help to describe how or why people are motivated to exercise. The EOS extends beyond measures of SDT constructs, as they relate to exercise by more directly exploring how exercise behavior determines self-esteem, rather than the forces involved in the regulation of exercise behavior. In addition, most applications of SDT to exercise focus on predictors of healthy exercise behavior (see Standage et al., 2019, for a review), whereas the construct of exercise overvaluation inherently explores a more pathological attitude toward engagement in exercise that may better predict negative outcomes (e.g., excessive exercise).

Limitations and Future Directions

The study has important limitations. First, we used convenience sampling for both studies, and the majority of our samples were disproportionately White, which limits the generalizability of these findings. In addition, though we exceeded the minimum sample size from our power calculations in Study 2, casual athletes were underrepresented relative to collegiate athletes. Second, while evidence from Studies 1 and 2 provide broad support for the validity of scale scores, scale items were only reviewed for content validity by the study's principal investigators. While additional validation by outside content experts would have been beneficial, one of the investigators (D. Lindner) is a content expert who was involved in neither the thematic content analysis nor the item generation process. Finally, the use of self-report data presumes that participants are (a) aware of their own thoughts, feelings, and behaviors and (b) willing to share those with researchers. We took steps to encourage participants to answer questions openly and honestly; the Study 1 data were collected anonymously online, and in Study 2, the coaches left the room and questionnaires were separated from consent forms.

Nevertheless, the findings of this study suggest that EOS is a useful tool for measuring exercise overvaluation. In particular, this measure supplements the existing measures of psychological constructs related to exercise (e.g., types of self-regulation, exercise identity, and obligatory exercise) and may help researchers better understand previously established relationships. For instance, the EOS may help researchers identify the point at which investment in exercise becomes problematic, and this, in turn, can inform the

development of health and wellness programming and coaching practices designed to foster a healthier, more adaptive investment in exercise and exercise behaviors that contribute to wellness rather than psychopathology. The ability to identify this "tipping point" may be especially useful in the study of the female athlete triad, where some female athletes who ostensibly began sport participation with healthier habits eventually develop maladaptive behaviors leading to low energy availability (which may or may not be accompanied by disordered eating behavior), decreased bone density, and amenorrhea or other menstrual dysfunction (Nazem & Ackerman, 2012). The female athlete triad can have long-term consequences for women's reproductive and skeletal systems, and these consequences could limit a woman's future exercise or sport participation (Thein-Nissenbaum, 2013). Exercise overvaluation may serve as a risk factor for the development of the female athlete triad and would likewise be an important factor to consider when developing treatment. In addition to exploring the relationship between exercise overvaluation and maladaptive behaviors, we also encourage researchers to examine exercise overvaluation in other contexts, including learning more about the connection between exercise overvaluation, motivation, and exercise habits to inform interventions to increase physical activity and foster general health and well-being in the general population.

Acknowledgments

This project was partially supported by the Stetson University College of Arts and Sciences Dean's Fund and the Stetson University Psychology Department.

References

Anderson, D.F., & Cychosz, C.M. (1994). Development of an exercise identity scale. *Perceptual and Motor Skills, 78*(3, Pt. 1), 747–751. https://doi.org/10.1177/003151259407800313

Armstrong, S., & Oomen-Early, J. (2009). Social connectedness, self-esteem, and depression symptomatology among collegiate athletes versus nonathletes. *Journal of American College Health, 57*(5), 521–526. PubMed ID: 19254893 https://doi.org/10.3200/JACH.57.5.521-526

Baquie, P., & Bruker, P. (1997). Injuries presenting to an Australian sports medicine centre: A 12-month study. *Clinical Journal of Sport Medicine, 7*(1), 28–31. PubMed ID: 9117522 https://doi.org/10.1097/00042752-199701000-00006

Bewell-Weiss, C.V., & Carter, J.C. (2010). Predictors of excessive exercise in anorexia nervosa. *Comprehensive Psychiatry, 51*(6), 566–571. PubMed ID: 20965301 https://doi.org/10.1016/j.comppsych.2010.03.002

Buchner, A., Erdfelder, E., Faul, F., & Lang, A.G. (2007). G*Power: Statistical power analyses for Windows and Mac (version 3.1.9.2). http://www.gpower.hhu.de/en.html

Crocker, J., Luhtanen, R.K., Cooper, M.L., & Bouvrette, A. (2003). Contingencies of self-worth in college students: Theory and measurement. *Journal of Personality and Social Psychology, 85*(5), 894–908. PubMed ID: 14599252 https://doi.org/10.1037/0022-3514.85.5.894

Crocker, J., & Wolfe, C.T. (2001). Contingencies of self-worth. *Psychological Review, 108*(3), 593–623. PubMed ID: 11488379 https://doi.org/10.1037/0033-295X.108.3.593

Deci, E.L., & Ryan, R.M. (1985). *Intrinsic motivation and self-determination in human behavior*. Plenum Press.

Deci, E.L., & Ryan, R.M. (1995). Human autonomy: The basis for true self-esteem. In M.H. Kernis (Ed.), *Efficacy, agency, and self-esteem – Plenum series in social/clinical psychology* (pp. 31–49). Plenum Press.

DiBartolo, P.M., Lin, L., Montoya, S., Neal, H., & Shaffer, C. (2007). Are there 'healthy' and 'unhealthy' reasons for exercise? Examining individual differences in exercise motivations using the Function of Exercise Scale. *Journal of Clinical Sport Psychology, 1*(2), 93–120. https://doi.org/10.1123/jcsp.1.2.93

Dishman, R.K., Hales, D.P., Pfeiffer, K.A., Felton, G.A., Saudners, R., Ward, D.S, ... Pate, R.R. (2006). Physical self-concept and self-esteem mediate cross-sectional relations of physical activity and sport participation with depression symptoms among adolescent girls. *Healthy Psychology, 25*(3), 396–407. PubMed ID: 16719612 https://doi.org/10.1037/0278-6133.25.3.396

Ertl, M.M., Longo, L.M., Groth, G.H., Berghuis, K.J., Prout, J., Hetz, M.C., & Martin, J.L. (2018). Running on empty: High self-esteem as a risk factor for exercise addiction. *Addiction Research and Theory, 26*(3), 205–211. https://doi.org/10.1080/16066359.2017.1347257

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

Fairburn, C.G., & Beglin, S.J. (1994). Assessment of eating disorders: Interview or self-report questionnaire? *International Journal of Eating Disorders, 16*(4), 363–370. PubMed ID: 7866415 https://doi.org/10.1002/1098-108X(199412)16:4<363:AID-EAT2260160405>3.0.CO;2-#

Gonçalves, S.F., & Gomes, A.R. (2012). Exercising for weight and shape reasons vs. health control reasons: The impact on eating disturbance and psychological functioning. *Eating Behaviors, 13*(2), 127–130. PubMed ID: 22365795 https://doi.org/10.1016/j.eatbeh.2011.11.011

Grieve, F.G., Jackson, L., Reece, T., Marklin, L., & Delaney, A. (2008). Correlates of social physique anxiety in men. *Journal of Sport Behavior, 31*(4), 329–337.

Heatherton, T.F., & Polivy, J. (1991). Development and validation of a scale for measuring state self-esteem. *Journal of Personality and Social Psychology, 60*(6), 895–910. https://doi.org/10.1037/0022-3514.60.6.895

Hogan, K.A., & Gross, R.H. (2003). Overuse injuries in pediatric athletes. *Orthopedic Clinics of North America, 34*(3), 405–415. https://doi.org/10.1016/S0030-5898(03)00006-3

Holm-Denoma, J.M., Scaringi, V., Gordon, K.H., Van Orden, K.A., & Joiner, T.E. (2009). Eating disorder symptoms among undergraduate varsity athletes, club athletes, independent exercisers, and nonexercisers. *International Journal of Eating Disorders, 42*(1), 47–53. PubMed ID: 18651619 https://doi.org/10.1002/eat.20560

Hrabosky, J.I., Cash, T.F., Veale, D., Neziroglu, F., Soll, E.A., Garner, D.M., ... Phillips, K.A. (2009). Multidimensional body comparisons among patients with eating disorders, body dysmorphic disorder, and clinical controls: A multisite study. *Body Image, 6*(3), 155–163. PubMed ID: 19410528 https://doi.org/10.1016/j.bodyim.2009.03.001

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

Hubbard, S.T., Gray, J.J., & Parker, S. (1998). Differences among women who exercise for 'food related' and 'non-food related' reasons. *European Eating Disorders Review, 6*(4), 255–265. https://doi.org/10.1002/(SICI)1099-0968(199812)6:4<255::AID-ERV262>3.0.CO;2-G

Joiner, T.E., & Tickle, J.J. (1998). Exercise and depressive and anxious symptoms: What is the nature of their interrelations? *Journal of*

Occupational Rehabilitation, 8(3), 191–198. https://doi.org/10.1023/A:1021326406744

Kell, K.P., & Rula, E.Y. (2019). Increasing exercise frequency is associated with health and quality of life benefits for older adults. *Quality of Life Research, 28*(12), 3267–3272. PubMed ID: 31410639 https://doi.org/10.1007/s11136-019-02264-z

Kernis, M.H. (2003). Toward a conceptualization of optimal self-esteem. *Psychological Inquiry, 14*(1), 1–26. https://doi.org/10.1207/S15327965PLI1401_01

Köteles, F., Kollsete, M., & Kollsete, H. (2016). Psychological concomitants of CrossFit training: Does more exercise really make your everyday psychological functioning better? *Kinesiology, 48*(1), 39–48. https://doi.org/10.26582/k.48.1.13

Laborde, S., Guillén, F., & Mosley, E. (2016). Positive personality-trait-like individual differences in athletes from individual- and team sports and in non-athletes. *Psychology of Sport and Exercise, 26,* 9–13. https://doi.org/10.1016/j.psychsport.2016.05.009

Levy, S.S., & Ebbeck, V. (2005). The exercise and self-esteem model in adult women: The inclusion of physical acceptance. *Psychology of Sport and Exercise, 6*(5), 571–584. https://doi.org/10.1016/j.psychsport.2004.09.003

Li, Y., Xu, Z., & Liu, S. (2014). Physical activity, self-esteem and mental health in students from ethnic minorities attending colleges in China. *Social Behavior and Personality, 42*(4), 529–537. https://doi.org/10.2224/sbp.2014.42.4.529

Maltby, J., & Day, L. (2001). The relationship between exercise motives and psychological well-being. *The Journal of Psychology: Interdisciplinary and Applied, 135*(6), 651–660. https://doi.org/10.1080/00223980109603726

McGraw, K.O., & Wong, S.P. (1996). Forming inferences about some intraclass correlation coefficients. *Psychological Methods, 1*(1), 30–46. https://doi.org/10.1037/1082-989X.1.1.30

Molina-Garcia, J., Castillo, I., & Queralt, A. (2011). Leisure-time physical activity and psychological well-being in university students. *Psychological Reports, 109*(2), 453–460. PubMed ID: 22238852 https://doi.org/10.2466/06.10.13.pr0.109.5.453-460

Moore, J.B., Mitchell, N.B., Kilpatrick, M.W., & Bartholomew, J.B. (2007). The physical self-attribute questionnaire: Development and initial validation. *Psychological Reports, 100*(2), 627–642. PubMed ID: 17564239 https://doi.org/10.2466/pr0.100.2.627-642

Morgan, W.P. (2000). Psychological outcomes of physical activity. In R.J. Maughan (Ed.), *Basic and applied sciences for sports medicine*. Sage.

Muchinsky, P.M. (1996). The correction for attenuation. *Educational and Psychological Measurement, 56*(1), 63–75. https://doi.org/10.1177/0013164496056001004

Muthén, L.K., & Muthén, B.O. (1998–2012). *Mplus user's guide* (7th ed.). Muthén & Muthén.

Nazem, R.G., & Ackerman, K.E., (2012). The female athlete triad. *Sports Health, 4*(4), 302–311. PubMed ID: 23016101 https://doi.org/10.1177/1941738112439685

Paradise, A.W., & Kernis, M.H. (1999). Development of the contingent self-esteem scale. Unpublished data, University of Georgia.

Pasman, L., & Thompson, J.K. (1988). Body image and eating disturbance in obligatory runners, obligatory weightlifters, and sedentary individuals. *International Journal of Eating Disorders, 7*(6), 759–769. https://doi.org/10.1002/1098-108X(198811)7:6<759::AID-EAT2260070605>3.0.CO;2-G

Rosenberg, M. (1965). *Society and the adolescent self-image*. Wesleyan University Press.

Rosenberg, M., Schooler, C., Schoenbach, C., & Rosenberg, F. (1995). Global self-esteem and specific self-esteem: Different concepts, different outcomes. *American Sociological Review, 60*(1), 141–156. https://doi.org/10.2307/2096350

Russell, W. (2002). Comparison of self-esteem, body satisfaction, and social physique anxiety across males of different exercise frequency and racial background. *Journal of Sport Behavior, 25*(1), 74–90.

Russell, W.D., & Cox, R.H. (2003). Social physique anxiety, body dissatisfaction, and self-esteem in college females of differing exercise frequency, perceived weight discrepancy, and race. *Journal of Sport Behavior, 23,* 297–381.

Ryan, R.M., & Deci, E.L. (2000). Self-determination theory and facilitation of intrinsic motivation, social development, and well-being. *American Psychologist, 55*(1), 68–78. PubMed ID: 11392867 https://doi.org/10.1037/0003-066X.55.1.68

Shrout, P.E., & Fleiss, J.L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychology Bulletin, 86*(2), 420–428. https://doi.org/10.1037/0033-2909.86.2.420

Smolak, L., Murnen, S.K., & Ruble, A.E. (2000). Female athletes and eating problems: A meta-analysis. *International Journal of Eating Disorders, 27*(4), 371–380. https://doi.org/10.1002/(SICI)1098-108X(200005)27:4<371::AID-EAT1>3.0.CO;2-Y

Spence, J.C., McGannon, K.R., & Poon, P. (2005). The effect of exercise on global self-esteem: A quantitative review. *Journal of Sport & Exercise Psychology, 27*(3), 311–334. https://doi.org/10.1123/jsep.27.3.311

Standage, M., Curran, T., & Rouse, P.C. (2019). Self-determination-based theories of sport, exercise, and physical activity motivation. In T.S. Horn& A.L. Smith (Eds.), *Advances in sport and exercise psychology* (pp. 289–311). Human Kinetics.

StataCorp. (2013). Stata statistical software (version 13). *StataCorpLB*. www.stata.com

Strelan, P., Mehaffey, S.J., & Tiggemann, M. (2003). Self-objectification and esteem in young women: The mediating role of reasons for exercise. *Sex Roles, 48*(1–2), 89–95. https://doi.org/10.1023/A:1022300930307

Thein-Nissenbaum, J. (2013). Long term consequences of the female athlete triad. *Maturitas, 75*(2), 107–112. PubMed ID: 23541905 https://doi.org/10.1016/j.maturitas.2013.02.010

United States Centers for Disease Control and Prevention. (2017). Exercise or physical activity. https://www.cdc.gov/nchs/fastats/exercise.htm

United States Department of Health and Human Services. (2018). *Physical activity guidelines for Americans* (2nd ed.). https://health.gov/paguidelines/second-edition/pdf/Physical_Activity_Guidelines_2nd_edition.pdf

Williams, K.L., & Galliher, R.V. (2006). Predicting depression and self-esteem from social connectedness, support and competence. *Journal of Social and Clinical Psychology, 25*(8), 855–874. https://doi.org/10.1521/jscp.2006.25.8.855

Woodward, H.E., Rizk, M.T., Wang, S.S., & Treat, T.A. (2014). Disordered eating links to body-relevant and body-irrelevant influences on self-evaluation. *Eating Behaviors, 15*(2), 205–208. PubMed ID: 24854805 https://doi.org/10.1016/j.eatbeh.2014.01.009