9-1-2015

Validity and Reliability of Eating Disorder Assessments Used with Athletes: A Review

Zachary Pope
Boise State University

Yong Gao
Boise State University

Nicole Bolter
Boise State University

Mary Pritchard
Boise State University

HOSTED BY

ELSEVIER

CrossMark

Available online at www.sciencedirect.com

ScienceDirect

Journal of Sport and Health Science 4 (2015) 211–221

www.jshs.org.cn

Review

Validity and reliability of eating disorder assessments used with athletes: A review

Zachary Pope*, Yong Gao, Nicole Bolter, Mary Pritchard

Department of Kinesiology, Boise State University, Boise, ID 83725, USA

Received 11 October 2013; revised 1 February 2014; accepted 12 May 2014
Available online 20 September 2014

Abstract

Background: Prevalence of eating disorders (EDs) among college-aged athletes has risen in recent years. Although measures exist for assessing EDs, these measures have not been thoroughly reviewed in athletes. This study reviewed the validity and reliability evidence of the commonly used measures for assessing EDs in athlete populations aged 18–26 years.
Methods: Databases were searched for studies of regarding ED on male and/or female athletes. Inclusion criteria stated the study (a) assessed EDs in an athlete population 18–26 years of age and (b) investigated EDs using a psychometric measure found valid and/or reliable in a non-athlete population and/or athlete population.
Results: Fifty studies met the inclusion criteria. Seven and 22 articles, respectively, studied EDs behaviors in male and female athletes whereas 21 articles studied EDs in combined-gender samples. The five most commonly used measures were the Eating Attitudes Test (EAT), Eating Disorder Inventory (EDI), Bulimia Test-Revised (BULIT-R), Questionnaire for Eating Disorder Diagnosis (QEDD), and the Eating Disorder Examination Questionnaire (EDE-Q).
Conclusion: Only seven studies calculated validity coefficients within the study whereas 47 cited the validity coefficient. Twenty-six calculated a reliability coefficient whereas 47 cited the reliability of the ED measures. Four studies found validity evidence for the EAT, EDI, BULIT-R, QEDD, and EDE-Q in an athlete population. Few studies reviewed calculated validity and reliability coefficients of ED measures. Cross-validation of these measures in athlete populations is clearly needed.

Keywords: Athletes; Eating disorders; Psychometrics; Reliability; Validity

1. Introduction

Eating disorders (EDs) encompass abnormal eating and weight control patterns, such as caloric restriction, excessive exercise, binging and/or purging, and abnormal body dissatisfaction, over a prolonged period of time.[1] According to the *Diagnostic and Statistical Manual of Mental Disorders: Fifth Edition* (DSM-5), common EDs include anorexia nervosa (AN), bulimia nervosa (BN), binge eating disorder (BED), and other specified feeding or eating disorders (OSFEDs). Anorexia nervosa is characterized by a severe limitation in caloric intake despite being severely underweight whereas BN features periods of abnormally high caloric intake in a short, distinct period of time (i.e., 2 h) during which the individual feels they have no control over their feeding behaviors followed by extreme purging measures (i.e., laxative use, vomiting, high amounts of exercise).[1] Individuals with BED experience the same period of abnormally high caloric intake and lack of control over their feeding behaviors as seen in BN but do not engage in extreme purging measures following the binge episode.[1] OSFEDs refers to an ED category wherein the individual meets a portion of the criteria for AN, BN, and/or BED but does not meet enough of these criteria to qualify as a clinical ED.[1] EDs occur among females and males in non-athlete populations and are

* Corresponding author.
E-mail address: zacharypope@u.boisestate.edu (Z. Pope)
Peer review under responsibility of Shanghai University of Sport.

2095-2546/$ - see front matter
http://dx.doi.org/10.1016/j.jshs.2014.05.001

concerning because of their negative effect on physical and mental health.[1] Given the danger EDs pose to a person's physical and mental health, assessing an individual's risk for EDs is vital for non-athletes as well as athlete populations.

EDs have been observed among female athletes and, more recently, some male athletes.[2–6] Sanford-Martens et al.[7] found 21.2% of a male athlete sample and 14.5% of female athletes possessed ED behaviors. In the seminal study of EDs in a large sample of Division-I athletes (n = 1445; 562 females, 883 males), Johnson et al.[3] found 13.02% of males and 10.85% of females engaged in binge eating at least once per week. Additionally, 5.52% and 2.04% of the female and male athletes, respectively, carried out some type of purging behavior on a weekly basis (i.e., use of laxatives, excessive exercise, vomiting). Two landmark studies on EDs in male athletes from a wide array of sports found 16.6%−19.2% to display ED behaviors.[4,6] The preceding findings indicate EDs occur in athlete populations and that both male and female athletes are affected.

Male and female athletes engaging in ED behaviors such as binging/purging, laxative use, or excessive exercise are putting both their athletic performance and health in serious jeopardy. For example, Sundgot-Borgen and Torstveit[8] state prolonged periods of caloric restriction cannot only degrade physical/psychological performance (e.g., strength production, fatigue levels, concentration, mental acuity) but also put the athlete in danger of serious health problems. Endocrine, cardiovascular, reproductive, and central nervous systems maladaptations, as well as gastrointestinal and renal problems, are all potential complications.[8] Thus, a need exists to properly assess EDs in male and female athletes to minimize any negative athletic performance or health consequences.

Gender is an essential consideration when one examines why male and female athletes engage in ED behaviors. Society's body ideals for each gender, and how these ideals affect athletes, may determine whether or not an athlete engages in ED behaviors. The "thin ideal" society projects upon female athletes may predispose them to engagement in weight control practices (e.g., excessive exercise, vomiting, use of laxatives) to lose weight—even if the loss in weight does not aid performance. In the hopes of achieving this thin, athletic body, some female athletes put themselves at risk for the female athlete triad (i.e., disordered eating, amenorrhea, and low bone mineral density)—a dangerous health condition.[9] In fact, 29.4%−57.1% of female athletes (varied based on the classification of the sport as aesthetic, endurance, team, and anaerobic) reported a bone injury during their collegiate career. Over 30% displayed ED behaviors and 31% report an irregular menstrual cycle without the use of oral contraceptives during this same time period.[9]

Conversely, some male athletes strive to reduce body fat while increasing muscle mass—producing the muscular, lean figure society deems most attractive.[10] This desire for a lean, muscular figure can predispose male athletes to ED behaviors such as binge eating, excessive exercise, and laxative use to build muscle but reduce body fat which may or may not be advantageous to the athlete's sport.[10,11] Not only are male and female athletes trying to conform to society's "ideal" body type, these individuals are also striving to achieve the body type which enhances sport performance.[10,12]

Male and female athletes are predisposed to engage in ED behaviors because of the sport context.[13] There can be sport-specific weight restrictions[14,15] and negative comments by coaches and teammates[16,17] that make athletes susceptible to the development of EDs. Furthermore, research suggests that EDs may be reinforced as coaches, teammates, and spectators comment upon changes in body type and performance that more closely align with how an athlete in said sport should appear or perform, respectively.[8]

Age and competitive level can also play a role in the onset of EDs. Woodside and Garfinkel[18] report individuals between the age of 18 and 26 years are more susceptible to ED (see also, Wright et al.[19]). This increased susceptibility to engagement in ED behaviors can arise due to the stress associated with a lack of structure and boundaries, moving away from home, and becoming more independent when young adults attend a college or university.[19] The preceding age range also corresponds to a time when athletes are often at higher levels of sport competition (e.g., collegiate, national, or international competitions). Athletes at higher levels of competition are exposed to even greater sport pressures (e.g., weight restrictions imposed by sport or coach, the need to conform to the "ideal" body type for a specific sport, belief weight reduction will enhance performance), which further predispose them to the development of EDs.[20,21] Given that athletes are under significant societal and sport pressures (e.g., sport-specific weight restrictions, pressure from coaches/teammates, conforming to both the male/female body ideal of society and sport), it is important for sports psychologists to have the tools necessary to assess EDs in this population.

EDs can be assessed via various psychometric measures. Through the use of these measures, psychologists can assess the severity of ED behaviors an athlete might engage in such as caloric restriction, binging/purging, and excessive exercise. Some examples of measures used for assessment of EDs include the Eating Attitudes Test (EAT),[22] the Eating Disorder Inventory (EDI),[23] the Bulimia Test-Revised (BULIT-R),[24] Questionnaire for Eating Disorder Diagnoses (QEDD),[25] and the Eating Disorder Examination Questionnaire (EDE-Q).[26] The EAT and EDI have multiple versions. The EAT has been shortened from its original 40-item version to a 26-item version, the EAT-26.[27] The EDI has two subsequent versions, the EDI-2[28] and EDI-3,[29] which have been modified to reflect the most current definitions of EDs.

These five measures are similar in that the questionnaires use dichotomous (i.e., yes/no) and/or Likert-type formatting to assess EDs (e.g., anorexic and bulimic behaviors, dangerous weight control behaviors) present in the individual being evaluated. The QEDD, EAT, EDI, and BULIT-R were developed from pre-existing definitions of EDs in the DSM.[18–20,25,26] The EDE-Q was also based upon the definitions of EDs from the DSM but was developed first into a structured interview format and then converted to a questionnaire.[26] Each EDs measure aims to assess specific types of

ED behaviors. For instance, the BULIT-R was developed to assess the degree of bulimic behavior present in an individual whereas the EAT was developed to gauge the severity of anorexic behavior.[18,20] Still other questionnaires, such as the EDI, QEDD, and EDE-Q, have subscales encompassing the assessment of both bulimic and anorexic tendencies.[19,25,26] The EAT, EDI, BULIT-R, QEDD, and EDE-Q are all capable of being completed within 10−15 min and yield preliminary evidence as to the severity of ED and weight control behaviors present in an individual. These questionnaires are cheaper and more time efficient than structured psychological interviews and, therefore, are used when there is a need to test a large group of individuals at once. Scores are most often summed and compared to cut-off scores (e.g., scoring a 20 on the EAT-26 is indicative of an ED). It is important to note that although it is common to assess EDs using the preceding questionnaires, these assessments alone cannot be used to make an official diagnosis of EDs. Official diagnoses of EDs must take place via structured clinical interviews.

The EAT, EDI, BULIT-R, QEDD, and EDE-Q were all developed and validated for measuring EDs in non-athlete populations. However, it is unclear whether these measures are valid for the assessment of EDs in male and female athletes. Petrie and Greenleaf[30] state the study of EDs in athlete populations is negatively impacted because many researchers use measures with "questionable psychometric properties". In line with Petrie and Greenleaf's observation, Hagger and Chatzisarantis[31] suggest one of the major problems in sport psychology research is researchers look to use measures validated in one population and administer these same measures to different populations. When a measure validated in one population is used with a new population without proper validation, the results of the study can be brought into question and the generalization of those results can be difficult.[31]

To ensure the trustworthiness of a study's results, researchers need to discern whether the measure they are using on a new population is valid and reliable. Validity refers to whether a test/instrument measures what it is supposed to measure (i.e., does an ED measure accurately assess the severity of eating disorder behaviors in athletes?) and can be measured in a number of ways (e.g., concurrent, predictive, convergent).[32] The validity of a measure can be further evaluated via tests of measurement invariance to determine whether an instrument measures the same construct (e.g., severity of ED behaviors) across different groups (e.g., male/female, cycling/swimming).[33] Reliability refers to the consistency of the measurement scores on a test/instrument measuring a certain attribute (e.g., if the same individual is administered an ED assessment twice, does the score remain the same and/or have very little variation?) and can also be measured in several ways (e.g., test-retest reliability, internal consistency).[34]

To date, little is known about whether ED measures are valid and reliable in both male and female athlete populations. Therefore, the purpose of this study was two-fold: (1) gather information about which ED measures are most commonly used with male and female athletes and (2) review the validity and reliability evidence of the various psychometric measures used for assessing EDs in male and female athlete populations 18−26 years of age. To our knowledge, no other review has undertaken this task. Ensuring valid and reliable ED assessments in athlete populations will allow for the accurate measurement and potential treatment of EDs among athletes.

2. Methods

The databases searched were SPORTDiscus, CINAHL, and PsycINFO. The search process was completed using the keywords "validity", "reliability", "eating disorders", "disordered eating", "college", and "athletes" in varying combinations from September 1990 to June 2012. Disordered eating refers to an individual possessing a disruption in feeding behaviors that does not meet the criteria for a clinical ED diagnosis.[1,35] It was included as a search term because the focus of the current study was on ED assessments, many of which are not only used to assess EDs, but also commonly used to concurrently examine disordered eating in the literature.

Three inclusion criteria were designated. First, the study had to be an original research article written in English. Second, the study must have assessed EDs in an athletic population of 18−26 years of age. The age range of 18−26 years was chosen because this is a period in an athlete's life when she/he is competing in the highest level of sport competition (i.e., college, national, or international) as well as the time period when individuals are most susceptible to EDs.[21,36] Third, the study had to investigate EDs using a measure found to be valid and/or reliable in non-athlete and/or athlete populations. Aside from the preceding inclusion criteria, no exclusion criteria were present as the researchers were open to studies of any sport, gender, ED assessment, and sample size.

Upon retrieving the articles that met the inclusion criteria, the following components of each article were stratified within a Microsoft Excel Spreadsheet (Microsoft, Inc., Redmond, WA, USA): ED assessment used, study name, study authors, year published, publishing journal, gender of population studied (female athletes, male athletes, combined male/female athlete population), sample size, sport, major findings (quantitative *vs*. qualitative), and statistical methods used. Most importantly, both validity and reliability coefficients for each ED measure were recorded within the spreadsheet. These coefficients were further delineated as one of two types: (1) values calculated directly from the current study or (2) values cited from another study. The type of validity and/or reliability calculated/cited was also recorded. These excel data were then used to (a) surmise the most commonly used ED assessments, (b) observe which studies calculated/cited the validity and reliability of the ED measure(s) used in studies investigating EDs behaviors in the male and female athletes, and (c) assess the type of validity and reliability used. This methodology allowed the researchers to make suggestions about ED assessments needing additional validity and reliability when investigating EDs among male and female athletes while also suggesting which measures have demonstrated adequate validity and reliability in this population.

Table 1
Validity and reliability coefficients for measures used with male athletes only.

ED assessment used	Authors[a]	Sample size	Sport type[b]	Validity				Reliability			
				Calculated validity coefficient (*r*)	Cited validity coefficient (*r*)	Citation for validity[a]	Population used for validation	Calculated reliability coefficient (α)	Cited reliability coefficient (α)	Citation for reliability[a]	Population used for reliability
EAT-26	Ferrand, 2004[44]	42	Cyclists	–	0.60–0.93	Garner, 1982[27]	Women	0.61–0.77	0.90	Garner, 1982[27]	Women
	Riebl, 2007[47]	124	Cyclists	–	0.60–0.93	Garner, 1982[27]	Women	–	0.90	Garner, 1982[27]	Women
EDI-2	Goldfield, 2006[45]	74	Bodybuilders	–	0.43–0.68	Garner, 1983[23]	Women	–	0.83–0.93	Garner, 1983[23]	Women
	Hallsworth, 2005[46]	83	Bodybuilders, weightlifters	–	0.83	Garner, 1991[28]	Women	0.59–0.85	0.83–0.93	Garner, 1991[28]	Women
QEDD/BULIT-R	Petrie, 2007[4]	199	Football, basketball, track and field	−0.51–0.70	0.73–0.90	QEDD: Mintz, 1997;[25] BULIT-R: Thelen, 1996[50]	QEDD: women; BULIT-R: women	0.87	0.85–0.98	QEDD: Mintz, 1997;[25] BULIT-R: Thelen, 1996[50]	QEDD: women; BULIT-R: women
	Petrie, 2008[6]	203	Football, basketball, track and field	−0.51–0.70	0.73–0.90	QEDD: Mintz, 1997;[25] BULIT-R: Thelen, 1996[50]	QEDD: women; BULIT-R: women	0.92	0.85–0.98	QEDD: Mintz, 1997;[25] BULIT-R: Thelen, 1996[50]	QEDD: women; BULIT-R: women
WPSS-MA	Galli, 2011[39]	203	Football, baseball, track and field	0.25–0.39	−0.51–0.70	BULIT-R comparison: Thelen, 1991[24]	Male athletes	0.91–0.92	0.87	BULIT-R comparison: Thelen, 1991[24]	Male athletes

[a] Citations have been listed by the first author's last name in order to conserve space.
[b] If study had more than three sports wherein ED behaviors were observed, only the three major sports studied are listed to conserve space.
Abbreviations: ED = eating disorder; EAT = Eating Attitudes Test; EDI = Eating Disorder Inventory; QEDD = Questionnaire for Eating Disorder Diagnosis; BULIT-R = Bulimia Test-Revised; WPSS-MA = Weight Pressures in Sport Scale for Male Athletes.

Table 2
Validity and reliability coefficients for measures used with female athletes only.

ED assessment used	Authors[a]	Sample size	Sport type[b]	Validity				Reliability			
				Calculated validity coefficient (*r*)	Cited validity coefficient	Citation for validity[a]	Population used for validation	Calculated reliability coefficient (α)	Cited reliability coefficient (α)	Citation for reliability[a]	Population used for reliability
AQ	Hinton, 2005[40]	167	Volleyball, swimming, basketball	0.46–0.79	–	N/A	Female athletes	0.77–0.91	–	N/A	Female athletes
BULIT-R	Petrie, 1993[58]	215	Gymnastics	–	0.99	Thelen, 1991[24]	Women	–	0.95	Thelen, 1991[24]	Women
EAT-26	Doninger, 2005[42]	207	Rowing, soccer, track and field	0.18–0.88	0.60–0.93	Garner, 1982[27]	Women	0.70–0.88	0.90	Garner, 1982[27]	Women
	Haase, 2009[59]	137	Netball, soccer, sport aerobics	–	0.60–0.93	Garner, 1982[27]	Women	0.74–0.89	0.90	Garner, 1982[27]	Women
	Haase, 2011[60]	136	Soccer, rowing, cycling	–	0.60–0.93	Garner, 1982[27]	Women	0.91	0.90	Garner, 1982[27]	Women
	Jankauskiene, 2012[61]	305	Track and field, cycling, ball sports	–	0.60–0.93	Garner, 1982[27]	Women	0.89	0.90	Garner, 1982[27]	Women
	Kirk, 2001[62]	403	Soccer, lacrosse, tennis	–	0.60–0.93	Garner, 1982[27]	Women	0.51–0.87	0.90	Garner, 1982[27]	Women
	Torres-McGehee, 2011[63]	138	Equestrian	–	0.60–0.93	Garner, 1982[27]	Women	0.89	0.90	Garner, 1982[27]	Women
EDE-Q	Beals, 2006[64]	112	Field hockey, cross country	–	0.78–0.85	Fairburn, 1994[26]	Women	–	–	–	Women
	de Bruin, 2011[65]	52	Dance, volleyball, football	–	0.78–0.85	Fairburn, 1994[26]	Women	–	–	–	Women
	Muscat, 2008[17]	223	Volleyball, soccer, running	–	0.78–0.85	Fairburn, 1994[26]	Women	0.93	0.85–0.92	Fairburn, 1994[26]	Women

EDI-2	Malinauskas, 2007[66]	115	Not specified	–	0.83	Garner, 1991[28]	Women	–	0.83–0.93	Garner, 1991[28]	Women
	Reinking, 2005[67]	146	Swimming, basketball, volleyball	–	0.83	Garner, 1991[28]	Women	–	0.83–0.93	Garner, 1991[28]	Women
	Torstveit, 2008[68]	331	Not specified	–	0.83	Garner, 1991[28]	Women	0.89–0.90	0.83–0.93	Garner, 1991[28]	Women
QEDD	Sears, 2012[12]	423	Track and field, swimming, cross country	–	0.90	Mintz, 1997[25]	Women	0.64–0.94	0.85–0.94	Mintz, 1997[25]	Women
EAT-26/EDI-2	Beals, 2002[9]	425	Basketball, cross country, track and field	–	0.43–0.93	EAT-26: Garner, 1982;[27] EDI-2: Garner, 1983[23]	EAT-26: women; EDI-2: women	–	0.83–0.93	EAT-26: Garner, 1982; EDI-2: Garner, 1983[23]	EAT-26: women; EDI-2: women
	Picard, 1999[69]	109	Cross country, rowing, basketball	–	0.43–0.93	EAT-26: Garner, 1982;[27] EDI-2: Garner, 1983[23]	EAT-26: women; EDI-2: women	–	0.83–0.93	EAT-26: Garner, 1982;[27] EDI-2: Garner, 1983[23]	EAT-26: women; EDI-2: women
QEDD/BULIT-R	Anderson, 2012[70]	414	Gymnastics, swimming, diving	–	0.73–0.90	QEDD: Mintz, 1997;[25] BULIT-R: Thelen, 1996[50]	QEDD: women; BULIT-R: women	–	–	QEDD: Mintz, 1997;[25] BULIT-R: Thelen, 1996[50]	QEDD: women; BULIT-R: women
	Greenleaf, 2009[71]	204	Soccer, softball, swimming	–	0.60–0.90	QEDD: Mintz, 1997;[25] BULIT-R: Thelen, 1996[50]	QEDD: women; BULIT-R: women	–	0.85–0.98	QEDD: Mintz, 1997;[25] BULIT-R: Thelen, 1996[50]	QEDD: women; BULIT-R: women
	Reel, 2007[72]	451	Gymnastics, basketball, golf	–	0.90	QEDD: Mintz, 1997;[25] BULIT-R: Thelen, 1996[50]	QEDD: women; BULIT-R: women	–	0.85–0.94	QEDD: Mintz, 1997;[25] BULIT-R: Thelen, 1996[50]	QEDD: women; BULIT-R: women
AMDQ/EDI-2 /BULIT-R	Nagel, 2000[41]	149	Basketball, cheerleading, gymnastics	0.55–0.62	0.60–0.65	BULIT-R: Brelsford, 1992[73]	AMDQ: female athletes; EDI-2: women; BULIT-R: women	0.77–0.95	0.90–0.93	BULIT-R: Brelsford, 1992[73]	AMDQ: female athletes; EDI-2: women; BULIT-R: women
EDI-2/EAT	Warren, 1990[37]	126	Gymnastics, cross country, basketball	–	0.43–0.87	EDI-2: Garner, 1983;[23] EAT: Garner, 1979[22]	EDI-2: women; EAT: women	–	0.83–0.94	EDI-2: Garner, 1983;[23] EAT: Garner, 1979[22]	EDI-2: women; EAT: women

[a] Citations have been listed by the first author's last name in order to conserve space.

[b] If study had more than three sports wherein ED behaviors were observed, only the three major sports studied are listed to conserve space.

Abbreviations: ED = eating disorder; EAT = Eating Attitudes Test; EDI = Eating Disorder Inventory; QEDD = Questionnaire for Eating Disorder Diagnosis; BULIT-R = Bulimia Test-Revised; AQ = ATHLETE Questionnaire; EDE-Q = Eating Disorder Examination Questionnaire; AMDQ = Athletic Milieu Direct Questionnaire.

Table 3
Validity and reliability coefficients for measures used with combined male/female athlete samples.

ED assessment used	Authors[a]	Sample size	Sport type[b]	Validity				Reliability			
				Calculated validity coefficient (r)	Cited validity coefficient (r)	Citation for validity[a]	Population assessment was found valid	Calculated reliability coefficient (α)	Cited reliability coefficient (α)	Citation for reliability[a]	Population assessment was found reliable
EAT-26	Blackmer, 2011[74]	103	Not specified	–	0.90	Mintz, 2000[88]	Women	–	0.90	Garner, 1982[27]	Women
	Costarelli, 2009[75]	60	Taekwondo, judo	–	0.87	Garner, 1979[22]	Women	–	0.93	Garner, 1979[22]	Women
	Lane, 2003[76]	165	Football, ice hockey, swimming	–	–		Women	0.79	0.93	Garner, 1979[22]	Women
	Milligan, 2006[77]	176	Basketball, track and field, wrestling	–	0.90	Mintz, 2000[88]	Women	0.96	0.51–0.87	Kirk, 2001[62]	Female athletes
	Pritchard, 2007[78]	194	Not specified	–	0.90	Mintz, 2000[88]	Women	–	0.93	Garner, 1979[22]	Women
	Pritchard, 2007[79]	354	Soccer, track and field, baseball	–	0.90	Mintz, 2000[88]	Women	0.94	0.51–0.87	Kirk, 2001[62]	Female athletes
	Rouveix, 2006[15]	55	Judo	–	0.87	Garner, 1979[22]	Women	0.63–0.76	0.93	Garner, 1979[22]	Women
	Sykora, 1993[14]	162	Rowing	–	0.87	Garner, 1979[22]	Women	–	0.93	Garner, 1979[22]	Women
	Van Zyl, 2012[38]	272	Not specified	–	0.60–0.93	Garner, 1982[27]	Women	0.78	0.90	Garner, 1982[27]	Women
EAT-40	Haase, 2002[80]	316	Waterpolo, cricket, wrestling	–	0.87	Garner, 1979[22]	Women	0.88	0.93	Garner, 1979[22]	Women
EDE-Q	Gomes, 2011[81]	290	Basketball, karate, swimming	–	0.78–0.85	Fairburn, 1994[26]	Women	0.94	0.81–0.92	Fairburn, 1994[26]	Women
EDI-2	Engel, 2003[82]	1445	Football, track and field, swimming	–	0.83	Garner, 1991[28]	Women	–	0.83–0.93	Garner, 1991[28]	Women
	Gapin, 2011[83]	179	Runners	–	0.43–0.68	Garner, 1983[23]	Women	0.75–0.83	0.83–0.93	Garner, 1983[23]	Women
	Johnson, 1999[3]	1445	Football, track and field, swimming	–	0.83	Garner, 1991[28]	Women	–	0.83–0.93	Garner, 1991[28]	Women
	Johnson, 2004[2]	1445	Football, track and field, swimming	–	0.83	Garner, 1991[28]	Women	–	0.83–0.93	Garner, 1991[28]	Women
	Reel, 1998[84]	124	Cheerleading	–	0.83	Garner, 1991[28]	Women	–	0.83–0.93	Garner, 1991[28]	Women
	Sundgot-Borgen, 2004[5]	3316	Track and field, basketball, baseball	–	0.83	Garner, 1991[28]	Women	–	0.83–0.93	Garner, 1991[28]	Women
EDI-2/QEDD	Hausenblas, 2004[85]	412	Track and field	–	0.98	QEDD: Mintz, 1997[25]	EDI-2: women; QEDD: women	0.69–0.85	0.83–0.93	EDI-2: Garner, 1991;[28] QEDD: Mintz, 1997[25]	EDI-2: women; QEDD: women
EAT-26/EDE-Q	Hopkinson, 2004[86]	250	Swimming, cross country, soccer	–	–	–	EAT-26: women; EDI-2: women	0.81–0.94	0.81–0.92	EAT-26: Garner, 1982;[27] EDE-Q: Fairburn, 1994[26]	EAT-26: women; EDE-Q: women
EDI-2/EAT	Marchand, 2007[87]	305	Football, gymnastics, rowing	–	0.60–0.93	EDI-2: Garner, 1991;[28] EAT: Garner, 1982[27]	EDI-2: women; EAT: women	–	0.82–0.90	EDI-2: Garner, 1991;[28] EAT: Garner, 1982[27]	EDI-2: women; EAT: women
QEDD/BULIT-R	Sanford-Martens, 2005[7]	489	Gymnastics, track and field, volleyball	−0.51–0.70	0.90	QEDD: Mintz, 1997[25]	QEDD: women; BULIT-R: women	–	0.85	QEDD: Mintz, 1997[25]	QEDD: women; BULIT-R: women

[a] Citations have been listed by the first author's last name in order to conserve space.
[b] If study had more than three sports wherein ED behaviors were observed, only the three major sports studied are listed to conserve space.
Abbreviations: ED = eating disorder; EAT = Eating Attitudes Test; EDI = Eating Disorder Inventory; QEDD = Questionnaire for Eating Disorder Diagnosis; BULIT-R = Bulimia Test-Revised; EDE-Q = Eating Disorder Examination Questionnaire.

3. Results

3.1. Characteristics of selected studies

Out of 450 articles identified, 50 met the inclusion criteria. The earliest study retrieved using the search terms listed and databases queried was from 1990,[37] whereas June 2012 was the most current study analyzed.[38] Sample sizes ranged from 17 to 3316 participants ($\overline{X} = 327$). Common individual sports studied were track and field and swimming whereas soccer and volleyball were the most frequent team sports to be examined. The percentage of athletes with EDs ranged from 7.1% to 60.0% ($\overline{X} = 23.9\%$) in these studies. In terms of gender, seven and 22 articles, respectively, evaluated exclusively male athletes or female athletes, whereas 21 articles assessed ED behaviors of male and female athletes within the same study. Tables 1–3 categorize articles by exclusively male, exclusively female, and combined-gender athlete studies, respectively.

3.2. Most commonly used ED assessments

The five most commonly used measures were the EAT ($n = 2$; EAT-26: $n = 21$), EDI ($n = 2$; EDI-2: $n = 15$), BULIT-R ($n = 9$), QEDD ($n = 8$), and the EDE-Q ($n = 5$). Of importance is that some studies ($n = 14$) included multiple ED measures (e.g., evaluated athletes with the EAT and EDI). None of the preceding five measures were developed for use in athlete populations.

Three ED assessments used were developed specifically for administration with athletes. These athlete-specific questionnaires were the Weight Pressures in Sport Scale for Male Athletes (WPSS-MA),[39] the ATHLETE Questionnaire (AQ),[40] and the Athletic Milieu Direct Questionnaire (AMDQ).[41] These assessments were each used once in the studies reviewed.

3.3. Calculated versus cited validity and reliability coefficients

Tables 1–3 indicate the eating disorder measure(s) used, the authors, sample sizes, sport, and whether the coefficient was calculated and/or cited for each male athlete, female athlete, and combined-gender athlete study, respectively. Only seven studies calculated validity coefficients within the study ($r = -0.51$–0.88) whereas 47 cited the validity coefficient ($r = -0.51$–0.99) of the measure of that was established in studies with non-athlete populations. Twenty-six calculated a reliability coefficient ($\alpha = 0.51$–0.96) while 47 cited the reliability ($\alpha = 0.81$–0.98) of measures obtained, again, from studies on non-athlete populations.

Three studies attempted to validate the less frequently used WPSS-MA,[39] the AQ,[40] and the AMDQ[41] in an athlete population. Of the studies using the EAT, EDI, BULIT-R, QEDD, and EDE-Q, only four studies reported both validity and reliability coefficients in these commonly used questionnaires when assessing athletes for ED. Doninger et al.[42] found validity evidence for the EAT in a population of female athletes while Petrie et al.[4,6] found validity evidence of the QEDD in populations of male athletes. Finally, Sanford-Martens et al.[7] also found validity evidence for the QEDD in a population of male and female athletes.

3.4. Types of validity and reliability calculated versus cited

Table 4 presents the frequency with which different types of validity and reliability coefficients were calculated/cited for the 50 selected studies. Regarding the seven studies that calculated validity coefficients, a majority of these studies calculated convergent validity ($n = 5$). Studies that cited validity coefficients primarily did so with convergent validity as well ($n = 31$) followed by citations of discriminant validity ($n = 13$). Although convergent validity and discriminant validity were the most calculated/cited types of validity coefficients, concurrent, predictive, and other indices of construct validity were also reported/cited.

Across gender, only three studies evaluated the same ED measure (i.e., BULIT-R) and calculated convergent validity coefficients within the study; two studies used male college athletes[4,6] with validity coefficients ranging from −0.51 to 0.70 (Table 1), and one used female athletes with validity coefficients ranging from 0.55 to 0.62[41] (Table 2). Although many studies included athletes from different sports, no study examined the validity of an ED measure in a specific sport type or compared validity across sport types. Moreover, no study examined measurement invariance to determine whether the same ED measure assesses similar ED construct across athlete groups (e.g., male/female, cycling/swimming).

Reliability coefficients were much more uniform in the type calculated/cited. Most often researchers who calculated a reliability coefficient did so with internal consistency ($n = 24$) whereas two studies also reported test-retest reliability. Researchers choosing to cite the reliability used in previous studies did so more often with internal consistency ($n = 46$) while nine studies also referenced test-retest reliability of the eating disorder measure used.

4. Discussion

Three major findings were unveiled as a result of this review. Although not surprising, the first finding of this review

Table 4
Types of calculated/cited validity and reliability coefficients.

Type	Calculated	Cited
Validity	7	47
Predictive	1	12
Criterion	–	8
Construct	1	1
Convergent	5	31
Discriminant	–	13
Concurrent	–	5
Reliability	26	47
Internal consistency	24	46
Test-retest reliability	3	9

Note: some studies calculated and/or cited more than one type of validity and/or reliability coefficient.

revealed that the number of studies (14% of the 50 studies reviewed) completed on exclusively male athletes was much lower in comparison to those conducted on exclusively female athletes. Secondly, this review found eight different measures were used in the assessment of EDs in athletes. Specifically, the use of the EAT, EDI, QEDD, BULIT-R, and EDE-Q questionnaires, developed for non-athlete populations, was much more prevalent than the use of psychometric measures assessing the same EDs constructed specifically for athletes—namely the WPSS-MA, AQ, and AMDQ. Finally, this review found a majority of the literature available examining EDs in athletes to cite the validity and reliability of EDs questionnaires reported in previous studies but fewer calculated their own validity and reliability coefficients with the athletic population they studied.

4.1. Gender of athlete population studied

This review found research on EDs in exclusively male athlete populations is less prevalent than research examining these same behaviors in female athlete samples. Hudson et al.[43] found rates of anorexia and bulimia to be significantly higher in non-athlete males than previously thought. The same is true within male athlete populations as the prevalence of ED behaviors has also been increasing in this population.[4,6,39,44–47] This increase in the prevalence of ED behaviors indicates the need for sports psychologists to validate ED assessments in this population to gain further knowledge of predisposing factors that might be unique to the development of EDs in male athletes. Despite the limited amount of research on exclusively male athlete samples and EDs, the QEDD has been found to be a valid psychometric measure for the assessment of these abnormal behaviors in this population.[4,6]

One barrier to studying EDs in male athletes might be that EDs have largely been considered a "woman's problem" and, therefore, the development of psychometric measures for EDs has been tailored toward the "thinness" ideal some women engage in ED behaviors to achieve.[48] Male athletes are more often concerned with increasing muscularity than with losing body fat, as doing so projects the epitome of masculinity/male athletes in contemporary culture.[10,11] Although measures such as the EAT, EDI, BULIT-R, QEDD, and EDE-Q may assess some of the factors related to female athletes engaging in ED behaviors to lose weight, these measures poorly assess male athletes' desire to engage in behaviors to become larger and, thus, bring into question the reliability and validity of these measures for male athletes. Furthermore, many of the ED measures available were developed over 20 years ago when the study of males in non-athlete populations, not to mention male athletes, was not a common topic to be studying. Therefore, the ED measures may not accurately account for factors contributing to male patterns of EDs. Although new ED measures such as the Eating Disorder Assessment for Men[49] (EDAM) are being developed to better account EDs among men, this measure has yet to be used to examine EDs among male athletes. All of the preceding factors suggest the study of EDs among male athletes and the further validation of the EAT, EDI, QEDD, BULIT-R, and EDE-Q for assessment of EDs in this population vital.

4.2. Most commonly used ED assessments

The second major finding of this review was that the use of EAT, EDI, BULIT-R, QEDD, and EDE-Q was much more frequent when assessing ED in athletes than the use of measures developed specifically for administration to athletes—WPSS-MA, AQ, and AMDQ. Only three studies, one for each questionnaire, used the WPSS-MA, AQ, and AMDQ. The lack of studies using the WPSS-MA, AQ, and AMDQ is not surprising considering these three ED measures are much newer in relation to the EAT, EDI, BULIT-R, QEDD, and EDE-Q (e.g., the AQ and WPSS-MA were developed/validated 8 and 2 years ago, respectively) and, thus, have not been used with enough frequency for researchers to realize these measures are available. Additionally, the lack of use of the WPSS-MA, AQ, and AMDQ might also be a result of the fact the EAT, EDI, BULIT-R, QEDD, and EDE-Q have always been available for use in the assessment of EDs in athlete samples, despite the fact these ED measures may not be valid in this population.

4.3. Calculated versus cited validity and reliability coefficients

Given the EAT, EDI, BULIT-R, QEDD, and EDE-Q are most frequently used within the literature to assess EDs in athletes, it is important to know which ED measure are best suited (i.e., have adequate validity and reliability in assessing EDs in athlete populations) for administration to male and female athletes. This review found approximately half the selected studies calculated a reliability coefficient within the athlete population ($n = 26$) and only seven studies calculated a validity coefficient, three of which were calculated for the infrequently used WPSS-MA, AQ, and AMDQ questionnaires. Not only have the EAT, EDI, BULIT-R, QEDD, and EDE-Q scarcely been validated in athlete populations, these five questionnaires have been validated almost exclusively in non-athlete populations with samples of women (EAT,[27] EDI,[19,28] BULIT-R,[50] QEDD,[25] EDE-Q[26]). Only four studies found validity evidence for the EAT, EDI, BULIT-R, QEDD, and EDE-Q in an athlete population.

Doninger et al.[42] found validity evidence for the EAT-26 in a population of 207 female athletes with convergent validity coefficients reported at $r = 0.18–0.88$ when evaluated alongside the EDI-2 and internal consistency reliability coefficients (α) reported ranging from 0.70 to 0.88. The r value of 0.18 reported in the study was a product of the ability of the EAT-26 to look at others' perceptions of an individual's eating behaviors when the EAT-26 was compared to the EDI. The EDI does not assess others' perceptions of eating, leading to the low r value.[42] However, the preceding results regarding the convergent validity between the EAT-26 and EDI demonstrate very little common variance between the two measures in

assessing others' perceptions of eating behaviors, pointing to the need for validation via other statistical methods besides convergent validity.

Additionally, evidence for the validity of the QEDD has also been found with an athlete population in three separate studies. Petrie et al.[4,6] examined ED with the QEDD in a population of 199 and 203 male athletes from both team and individual sports, respectively. Sanford-Martens et al.[7] also found evidence of the validity of the QEDD when studying a combined sample of 325 male and female athletes (159 females, 166 males). Between the three studies, convergent validity coefficients for the QEDD were $r = -0.51-0.70$ with internal consistency reliability coefficients reported at $\alpha = 0.87$. The negative r value is expected because it was obtained through correlating a subscale within QEDD that assesses body satisfaction with BULIT-R,[4] indicating athletes with higher body satisfaction were less likely to have bulimic tendencies. Given the QEDD assesses the degree of both bulimic and anorexic behaviors and demonstrated moderate to good validity and good reliability in both male and female athletes, the QEDD appears best equipped to gauge ED in athletes of both genders.

Once again, the studies researchers cite are most often the validation studies conducted with non-athlete populations. In this accord, one needs to question the accuracy of the measure with athlete populations.[31] As stated above, only Doninger et al.,[42] Petrie et al.,[4,6] and Sanford-Martens et al.[7] have found validity evidence for two of the five most commonly used ED measures (EAT, EDI, QEDD, BULIT-R, and EDE-Q), specifically the EAT and QEDD. Because questionnaires developed specifically for athletes (i.e., WPSS-MA, AQ, and AMDQ) are used much less frequently than the EAT, EDI, QEDD, BULIT-R, and EDE-Q among the literature, calculating, and reporting the validity and reliability coefficients of the EAT, EDI, QEDD, BULIT-R, and EDE-Q with athlete populations is needed.

4.4. Types of validity and reliability calculated or cited

The results of this review indicate studies that did calculate validity and reliability coefficients for the ED assessment used to observe athletes did so with traditional psychometrical/statistical methods (e.g., classic test theory (CTT)). While the use of CTT is widespread in research, CTT does present some inadequacies for constructing/validating psychometric measures; for example, CTT does not measure latent variables such as ED severity adequately and is both sample- and item-dependent, which increases the error of the measurement.[51,52]

Specific to the results of this review regarding validity and reliability, it is concerning the most frequently calculated/cited type of validity was convergent validity. Convergent validity is typically employed when a researcher wishes to draw a correlation between a specific field measure and another field measure within an area of research.[32] A field measure is typically less accurate when used to assess an attribute compared with a "gold standard" because field tests usually contain more errors. For example, when measuring aerobic fitness, a 12-min running test (a field test) is often less accurate than a laboratory test (a gold standard) because many factors such as running efficiency, road condition, or temperature can introduce measurement errors in the 12-min running test. In this regard, convergent validity is less preferred if criterion-related validity, determined by correlating scores from a field measure to those from a gold standard measure, of a measure can be established. Regarding ED assessments, no measure is considered the "gold standard" within the field, which renders the measurement of criterion-related validity inadequate for ED measures.

Regarding reliability, 24 studies calculated internal consistency while only three studies calculated test-retest reliability. Although measures of internal consistency are commonly used, further tests of other types of reliability might be advocated for reliability checking. Specifically, it may be worthwhile to include test-retest reliability in the evaluation of eating disorder measurements in athlete populations to assess whether or not athletes achieve approximately the same EDs score during multiple assessments[53] and/or to ensure that changes of EDs scores over time are not the result of measurement property change of the eating disorder measures.

4.5. Limitations and future directions

A limitation of the current study is that this review could not present sufficient information about the validity and reliability of ED measures across genders and sport types. As such, these results cannot provide recommendations on whether an ED measure assesses EDs similarly across different athlete groups because such information on comparing groups and measurement invariance was unavailable in the current literature. If researchers want to make meaningful comparisons about the prevalence and severity of EDs across athlete groups (e.g., male/female, different sport types), it is essential to establish that a measure is not only valid across groups but also evaluates the groups in the same way prior to comparing mean scores. Otherwise, differences in EDs between groups cannot be appropriately interpreted.

Future research on EDs in athletes should also look to validate the psychometric measure used via advanced psychometrical approaches in addition to the use of CTT. Advanced modern measurement approaches such as Item Response Theory[54] and Rasch Modeling[55] have been demonstrated to assess latent variables (e.g., abilities, attributes) better than CTT,[51,52,56,57] which might provide additional insight into a better investigation of ED assessments and help discern whether these psychometric measures are valid and reliable in an athlete population. Additionally, further validation of ED measures in a male athlete population may allow for the development of rehabilitative programs to help these individuals with EDs as well as contribute to the small amount of literature examining exclusively male athletes for EDs. Finally, future studies might look to complete an analysis of measures that assess "drive for muscularity". This form of body dysmorphia is becoming more prevalent in both male and female athletes but was not analyzed in this study.[11]

5. Conclusion

With the increasing prevalence of EDs in athlete populations, valid and reliable ED assessment tools will be paramount to detecting and providing treatment for ED behaviors some athletes possess before these behaviors progress to a clinical level. This review provides clear evidence that the most commonly used ED assessment tools (EAT, EDI, BULIT-R, QEDD, and EDE-Q) need to be more thoroughly assessed for validity and reliability among athlete populations. Perhaps more modern measurement approaches (e.g., Rasch Modeling) can provide a better understanding than traditional methods typically employed. Furthermore, research is needed with male athletes as this cohort of the athlete population has not been studied nearly as extensively as female athletes. Ensuring valid and reliable eating disorder assessment tools are being used when evaluating EDs in athlete populations is important. Not only will valid and reliable eating disorder assessments enable sport psychologists to identify those athletes at risk for a clinical diagnosis, it will also allow trainers and coaches to work together with sport psychologists to develop rehabilitative treatment processes for these harmful behaviors within the context of sport.

References

1. American Psychiatric Association. *Diagnostic and statistical manual of mental disorders*. 5th ed. Arlington, VA: American Psychiatric Association; 2013.
2. Johnson C, Crosby R, Engel S, Mitchell J, Powers P, Wittrock D, et al. Gender, ethnicity, self-esteem and disordered eating among college athletes. *Eat Behav* 2004;**5**:147–56.
3. Johnson C, Powers PS, Dick R. Athletes and eating disorders: the national collegiate athletic association study. *Int J Eat Disord* 1999;**26**:179–88.
4. Petrie T, Greenleaf C, Carter J, Reel J. Psychosocial correlates of disordered eating among male collegiate athletes. *J Clin Sport Psychol* 2007;**1**:340–57.
5. Sundgot-Borgen J, Torstveit MK. Prevalence of eating disorders in elite athletes is higher than in the general population. *Clin J Sport Med* 2004;**14**:25–32.
6. Petrie T, Greenleaf C, Reel J, Carter J. Prevalence of eating disorders and disordered eating behaviors among male collegiate athletes. *Psychol Men Mascul* 2008;**9**:267–77.
7. Sanford-Martens T, Davidson M, Yakushko O, Martens M, Hinton P, Beck N. Clinical and subclinical eating disorders: an examination of collegiate athletes. *J Appl Sport Psychol* 2005;**17**:79–86.
8. Sundgot-Borgen J, Torstveit M. Aspects of disordered eating continuum in elite high-intensity sports. *Scand J Med Sci Sports* 2010;**20**(Suppl. 2):112–21.
9. Beals K, Manore M. Disorders of the female athlete triad among collegiate athletes. *Int J Sport Nutr Exerc* 2002;**12**:281–93.
10. Galli N, Reel J. Adonis or hephaestus? Exploring body image in male athletes. *Psychol Men Mascul* 2009;**10**:95–108.
11. Murray S, Baghurst T. Revisiting the diagnostic criteria for muscle dysmorphia. *Strength Cond J* 2013;**35**:69–74.
12. Sears L, Tracy K, McBrier N. Self-esteem, body image, internalization, and disordered eating among female atheltes. *Athlet Train Sports Health Care* 2012;**4**:29–37.
13. Thompson R, Sherman R. *Eating disorders in sport*. New York, NY: Routledge; 2010.
14. Sykora C, Grilo CM, Wilfrey DE, Brownell KD. Eating, weight, and dieting disturbances in male and female lightweight and heavyweight rowers. *Int J Eat Disord* 1993;**14**:203–11.
15. Rouveix M, Bouget M, Pannafieux C, Champely S, Filaire E. Eating attitudes, body esteem, perfectionism and anxiety of judo athletes and nonathletes. *Int J Sports Med* 2006;**28**:340–5.
16. Arthur-Cameselle J, Quatromoni P. Factors related to the onset of eating disorders by female collegiate athletes. *Sport Psychol* 2010;**25**:1–17.
17. Muscat A, Long B. Critical comments about body shape and weight: disordered eating of female athletes and sport participants. *J Appl Sport Psychol* 2008;**20**:1–24.
18. Woodside D, Garfinkel P. Age of onset in eating disorders. *Int J Eat Disord* 1992;**12**:31–6.
19. Wright F, Bewick BM, Barkham M, House AO, Hill AJ. Co-occurence of self reported disordered eating and self-harm in UK university students. *Br J Clin Psychol* 2009;**48**:397–410.
20. Thompson R, Sherman R. Athletes, athletic performance, and eating disorders: healthier alternatives. *J Soc Issues* 1999;**55**:317–37.
21. Thompson R, Sherman R. "Good athlete" traits and characteristics of anorexia nervosa: are they similar? *Eat Disord* 1999;**7**:181–90.
22. Garner G, Garfinkel P. The eating attitudes test: an index of the symptoms of anorexia nervosa. *Psychol Med* 1979;**9**:273–9.
23. Garner G, Olmstead M, Polivy J. Development and validation of a multidimensional eating disorder inventory for anorexia nervosa and bulimia. *Int J Eat Disord* 1983;**2**:15–34.
24. Thelen M, Farmer J, Wonderlich S, Smith M. A revision of the bulimia test: the BULIT-R. *Psychol Assess* 1991;**3**:119–24.
25. Mintz L, O'Halloran M, Mulholland A, Schneider P. Questionnaire for eating disorder diagnoses: reliability and validity of operationalizing DSM-IV criteria into a self-report format. *J Couns Psychol* 1997;**44**:63–79.
26. Fairburn C, Beglin S. Assessment of eating disorders: interview or self-report questionnaire. *Int J Eat Disord* 1994;**16**:363–70.
27. Garner G, Olmsted M, Bohr Y, Garfinkel P. The eating attitudes test: psychometric features and clinical correlates. *Psychol Med* 1982;**12**:871–8.
28. Garner D. *Eating disorder inventory-2: professional manual*. Lutz, FL: Psychological Assessment Resourses; 1991.
29. Garner D. *Eating disorder inventory-3: professional manual*. Lutz, FL: Psychological Assessment Resources; 2004.
30. Petrie T, Greenleaf C. Eating disorders in sport: from theory to research to intervention. In: Tenenbaum G, Eklund R, editors. *Handbook of sport psychology*. 3rd ed. Hoboken, NJ: John Wiley and Sons; 2007.p.362.
31. Hagger M, Chatzisarantis N. Assumptions in research in sport and exercise psychology. *Psychol Sport Exerc* 2009;**10**:511–9.
32. Borsboom D, Mellenbergh G, van Heerden J. The concept of validity. *Psychol Rev* 2004;**111**:1061–71.
33. Meredith W. Measurement invariance, factor analysis and factorial invariance. *Psychometrika* 1993;**58**:525–43.
34. Dimitrov D. Reliability: arguments for multiple perspectives and potential problems with generalization across studies. *Educ Psychol Meas* 2002;**62**:783–801.
35. Holt M, Espelage D. Problem-solving skills and relationship attributes among women with eating disorders. *J Couns Dev* 2002;**80**:346–54.
36. Breen H, Desilva P, Espelage D. Treatment of disordered eating. *Curr Med Lit Clin Nutr* 2005;**14**:40–5.
37. Warren B, Stanton A, Blessing D. Disordered eating patterns in competitive female athletes. *Int J Eat Disord* 1990;**9**:565–9.
38. Van Zyl Y, Surujlal J, Dhurup M. Eating disorders among university student-athletes. *Afr J Phys Acti Health Educ Rec Dance* 2012;**18**:267–80.
39. Galli N, Reel J, Petrie T, Greenleaf C, Carter J. Preliminary development of weight pressure in sport scale for male athletes. *J Sport Behav* 2011;**34**:47–68.
40. Hinton P, Kubas K. Psychosocial correlates of disordered eating in female collegiate athletes: validation of the ATHLETE questionnaire. *J Am Coll Health* 2005;**54**:149–56.
41. Nagel DL, Black DR, Leverenz LJ, Coster DC. Evaluation of a screening test for female college athletes with eating disorders and disordered eating. *J Athl Train* 2000;**35**:431–40.
42. Doninger G, Enders C, Burnett K. Validity evidence for eating attitudes test scores in a sample of female college athletes. *Meas Phys Educ Exerc Sci* 2005;**9**:35–49.

43. Hudson J, Hiripi E, Pope Jr HG, Kessler RC. The prevalence and correlates of eating disorders in the national comorbidity survey replication. *Biol Psychiatry* 2007;**61**:348–58.
44. Ferrand C, Brunet E. Perfectionism and risk for disordered eating among young french male cyclists of high performance. *Percept Motor Skill* 2004;**99**:959–67.
45. Goldfield G, Blouin A, Woodside D. Body image, binge eating, and bulimia nervosa in male bodybuilders. *Can J Psychiatry* 2006;**51**:160–8.
46. Hallsworth L, Wade T, Tiggemann M. Individual difference in male body-image: an examination of self-objectification in recreational body builders. *Br J Health Psychol* 2005;**10**:453–65.
47. Riebl SK, Subudhi AW, Broker JP, Schenck K, Berning JR. The prevalence of subclinical eating disorders among male cyclists. *J Am Diet Assoc* 2007;**107**:1214–7.
48. Andersen AE, Holman JE. Males with eating disorders: challenges for treatment and research. *Psychopharmacol Bull* 1997;**33**:391–7.
49. Stanford SC, Lemberg R. Measuring eating disorders in men: development of the eating disorder assessment for men (EDAM). *Eat Disord* 2012;**20**:427–36.
50. Thelen M, Mintz L, Vander Wal J. The bulimia test-revised: validation with DSM-IV criteria for bulimia nervosa. *Psychol Assess* 1996;**8**:219–21.
51. Bond T, Fox C. *Applying the Rasch Model: fundamental measurement in the human sciences*. 2nd ed. New York, NY: Routledge; 2007.
52. Massof RW. Understanding Rasch and item response theory models: applications to the estimation and validation of interval latent trait measures from responses to rating scale questionnaires. *Ophthalmic Epidemiol* 2011;**18**:1–19.
53. American Educational Research Association APA, National Council on Measurement in Education. *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association; 1999.
54. Hambleton R, Swaminathan H, Rogers H. *Fundamentals of item response theory*. Newbury Park, CA: Sage Press; 1991.
55. Rasch G. *Probabilisic models for some intelligence and attainment test*. Chicago, IL: University of Chicago Press; 1980.
56. Lerdal A, Kottorp A, Gay C, Aouizerat BE, Portillo CJ, Lee KA. A 7-item version of the fatigue serverity scale has better psychometric properties among HIV-infected adults: an application of the rasch model. *Qual Life Res* 2011;**20**:1447–56.
57. Muller M. Nursing competence: psychometric evaluation using Rasch modelling. *J Adv Nurs* 2013;**69**:1410–7.
58. Petrie T. Disordered eating in female collegiate gymnasts: prevalence and personality/attitudinal correlates. *J Sport Exerc Psychol* 1993;**15**:424–36.
59. Haase A. Physique anxiety and disordered eating correlates in female athletes: differences in team and individual sports. *J Clin Sport Psychol* 2009;**3**:218–31.
60. Haase A. Weight perception in female athletes: associations with disordered eating correlates and behaviors. *Eat Behav* 2011;**12**:64–7.
61. Jankauskiene R, Pajaujiene S. Disordered eating attitudes and body shame among athletes, exercisers, and sedentary female college students. *J Sport Med Phys Fit* 2012;**52**:92–101.
62. Kirk G, Singh K, Getz H. Risk of eating disorders among female college athletes and nonathletes. *J Coll Couns* 2001;**4**:122–32.
63. Torres-McGehee T, Monsma E, Gay J, Minton D, Mady-Foster A. Prevalences of eating disorder risk and body image distortion among national collegiate athletic association division I varsity equestrian athletes. *J Athl Train* 2011;**46**:431–7.
64. Beals K, Hill A. The prevalence of disordered eating, menstrual dysfunction, and low bone mineral density among US collegiate athletes. *Int J Sport Nutr Exerc* 2006;**16**:1–23.
65. de Bruin A, Oudejans R, Bakker F, Woertman L. Contextual body image and athletes' disordered eating: the contribution of athletic body image to disordered eating in high performance women athletes. *Eur Eat Disord Rev* 2011;**19**:201–15.
66. Malinauskas B, Cucchiara A, Aeby V, Bruening C. Physical activity, disordered eating risk, and anthropometric measurement: a comparison of college female athletes and nonathletes. *Coll Stud J* 2007;**41**:217–22.
67. Reinking MF, Alexander LE. Prevalence of disordered-eating behaviors in undergraduate female collegiate athletes and nonathletes. *J Athl Train* 2005;**40**:47–51.
68. Torstveit MK, Rosenvinge JH, Sundgot-Borgen J. Prevalence of eating disorders and the predictive power of risk models in female elite athletes: a controlled study. *Scand J Med Sci Sports* 2008;**18**:108–18.
69. Picard C. The level of competition as a factor for the development of eating disorders in female collegiate athletes. *J Youth Adolescence* 1999;**28**:583–94.
70. Anderson C, Petrie TA. Prevalence of disordered eating and pathogenic weight control behaviors among NCAA division I female collegiate gymnasts and swimmers. *Res Q Exerc Sport* 2012;**83**:120–4.
71. Greenleaf C, Petrie TA, Carter J, Reel JJ. Female collegiate athletes: prevalence of eating disorders and disordered eating behaviors. *J Am Coll Health* 2009;**57**:489–95.
72. Reel J, SooHoo S, Doetsch H, Carter J, Petrie T. The female athlete triad: is the triad a problem among division I female athletes. *J Clin Sport Psychol* 2007;**1**:358–70.
73. Brelsford T, Hummel R, Barrios B. The bulimia test-revised: a psychometric investigation. *Psychol Assessment* 1992;**4**:399–401.
74. Blackmer V, Searight HR, Ratwik SH. The relationship between eating attitudes, body image, and perceived family-of-origin climate among college athletes. *N Am J Psychol* 2011;**13**:435–46.
75. Costarelli V, Stamou D. Emotional intelligence, body image, and disordered eating attitudes in combat sport athletes. *J Exerc Sci Fit* 2009;**7**:104–11.
76. Lane A. Relationships between attitudes towards eating disorders and mood. *J Sci Med Sport* 2003;**6**:144–54.
77. Milligan B, Pritchard M. The relationship between gender, type of sport, body dissatisfaction, self-esteem and disordered eating behaviors in division I athletes. *Athletic Insight* 2006;**8**:32–46.
78. Pritchard M, Rush P, Milligan B. Risk factors for disordered eating in high school and college athletes. *Athlet Insight Online J Sport Psychol* 2007;**9** [no pagination specified].
79. Pritchard M, Milligan B, Elgin J, Rush P, Shea M. Comparisons of risky health behaviors between male and female college athletes and non-athletes. *Athlet Insight Online J Sport Psychol* 2007;**9**:67–78.
80. Haase A, Prapavessis H, Owens R. Perfectionism, social physique anxiety, and disordered eating: a comparison of male and female elite athletes. *Psychol Sport Exerc* 2002;**3**:209–22.
81. Gomes AR, Martins C, Silva L. Eating disorder behaviours in Portuguese athletes: the influence of personal, sport, and psychological variables. *Eur Eat Disord Rev* 2011;**19**:190–200.
82. Engel SG, Johnson C, Powers PS, Crosby RD, Wonderlich SA, Wittrock DA, et al. Predictors of disordered eating in a sample of elite division I college athletes. *Eat Behav* 2003;**4**:333–43.
83. Gapin JI, Petruzzello SJ. Athletic identity and disordered eating in obligatory and non-obligatory runners. *J Sports Sci* 2011;**29**:1001–10.
84. Reel J, Gill D. Weight concerns and disordered eating attitudes among male and female college cheerleaders. *Women Sport Phys Act J* 1998;**7**:79–94.
85. Hausenblas H, McNally K. Eating disorder prevalence and symptoms for track and feild athletes and nonathletes. *J Appl Sport Psychol* 2004;**16**:274–86.
86. Hopkinson RA, Lock J. Athletics, perfectionism, and disordered eating. *Eat Weight Disord* 2004;**9**:99–106.
87. Marchand S. *Disordered eating behaviors among collegiate athletes*. Kingston: University of Rhode Island; 2007 [ProQuest Dissertations and Theses].
88. Mintz L, O'Halloran M. The eating attitudes test: validation with DSM-IV eating disorder criteria. *J Pers Assess* 2000;**74**:489–503.