Trinity University

Digital Commons @ Trinity

Psychology Faculty Research

Psychology Department

9-2017

The Female Athlete Body (FAB) Study: Rationale, Design, and Baseline Characteristics

Tiffany M. Stewart

Tarryn Pollard

Tom Hildebrandt

Robbie Beyl

Nicole Wesley

See next page for additional authors

Follow this and additional works at: https://digitalcommons.trinity.edu/psych_faculty

Part of the Psychology Commons
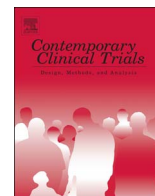
Publication Details

Repository Citation

Authors

Tiffany M. Stewart, Tarryn Pollard, Tom Hildebrandt, Robbie Beyl, Nicole Wesley, Lisa S. Kilpela, and Carolyn Becker

CrossMark

The Female Athlete Body (FAB) study: Rationale, design, and baseline characteristics☆

Tiffany M. Stewart[a],*, Tarryn Pollard[a], Tom Hildebrandt[b], Robbie Beyl[a], Nicole Wesley[a], Lisa Smith Kilpela[c], Carolyn Black Becker[d]

[a] Pennington Biomedical Research Center, Baton Rouge, LA, United States
[b] The Mount Sinai Hospital, United States
[c] University of Texas Health Science Center at San Antonio, Antonio, TX, United States
[d] Trinity University, San Antonio, TX, United States

A B S T R A C T

Background: Eating Disorders (EDs) are serious psychiatric illnesses marked by psychiatric comorbidity, medical complications, and functional impairment. Research indicates that female athletes are often at greater risk for developing ED pathology versus non-athlete females. The Female Athlete Body (FAB) study is a three-site, randomized controlled trial (RCT) designed to assess the efficacy of a behavioral ED prevention program for female collegiate athletes when implemented by community providers. This paper describes the design, intervention, and participant baseline characteristics. Future papers will discuss outcomes.
Methods: Female collegiate athletes ($N = 481$) aged 17–21 were randomized by site, team, and sport type to either FAB or a waitlist control group. FAB consisted of three sessions (1.3 h each) of a behavioral ED prevention program. Assessments were conducted at baseline (pre-intervention), post-intervention (3 weeks), and six-, 12-, and 18-month follow-ups.
Results: This study achieved 96% ($N = 481$) of target recruitment ($N = 500$). Few group differences emerged at baseline. Total sample analyses revealed moderately low baseline instances of ED symptoms and clinical cases.
Conclusions: Health risks associated with EDs necessitate interventions for female athletes. The FAB study is the largest existing RCT for female athletes aimed at both reduction of ED risk factors and ED prevention. The methods presented and population recruited for this study represent an ideal intervention for assessing the effects of FAB on both the aforementioned outcomes. We anticipate that findings of this study (reported in future papers) will make a significant contribution to the ED risk factor reduction and prevention literature.

1. Introduction

Despite recent advances in Eating Disorder (ED) treatment, approximately 50% of those with EDs remain symptomatic over time [12]. Research indicates that even the most efficacious ED treatments leave a significant number of individuals symptomatic and/or at risk for relapse [28]. Together, these findings highlight prevention of EDs as a key mental health goal.

Female collegiate athletes represent an important ED prevention cohort. This group is often at higher risk for developing EDs versus non-athlete females [18,26]. Research indicates that as many as 70% of NCAA Division (D)-I female athletes consume insufficient calories to support daily energy needs [11]. Insufficient caloric intake, which is associated with disordered eating, increases risk for the Female Athlete

Triad and subsequent injury [19]. The Triad (low energy availability, menstrual disorders, and decreased bone mineral density) increases risk for serious long-term health consequences, such as osteoporosis, reproductive disorders, and cardiovascular disease [6]. Although the International Olympic Committee recently renamed and expanded the Triad under the heading of Relative Energy Deficiency in Sport [14], this change was released during the course of the current trial and most existing literature focuses on the Triad. Thus, this paper uses the Female Athlete Triad construct.

Athletic departments represent the type of community that may sustain and fully integrate programs aimed at reducing ED risk factors and/or preventing EDs. Athletic departments have an incentive to maintain the health of their athletes both for ethical and pragmatic (i.e., performance) reasons. This incentive, in conjunction with the typical

high level of organization and available resources found in an athletic department, produces an ideal environment for introducing and maintaining efficacious ED prevention programs.

A pilot/feasibility study (R15MH077659) tested the Female Athlete Body Project (FAB) program (formerly called Athlete-Modified Healthy Weight Intervention {AM-HWI}) in a female athlete sample [3]. This study provided the first evidence that a relatively short intervention could reduce ED risk factors and symptoms at one-year follow-up in female athletes. For more details, please see Becker et al. [3].

The present study (1 RO1 MH094448-01) examined the FAB program in a large-scale, three site, two-arm randomized controlled trial (RCT) using group (cluster) randomization. The overarching objective of the study was to test the efficacy of FAB on both primary (e.g., ED symptoms and body image variables) and secondary outcomes (e.g., triad knowledge, healthcare utilization, negative affect, and seeking help for the triad). This study also sought to test the efficacy when FAB was implemented under sustainable conditions by having the program delivered by community members (i.e., peer leaders). Thus, this trial is best conceptualized as an efficacy/effectiveness hybrid trial. The present paper outlines the study design/method, intervention, and baseline sample characteristics. Study outcomes will be discussed in future publications.

2. Method

2.1. Study design

The FAB study is a 3-site RCT in which female collegiate athletes were assigned by entire team (group/cluster randomization) to either the FAB intervention or a waitlist brochure control condition. In the FAB condition, participants received one peer-led 1.3 h session per week for 3 weeks. Those in the waitlist brochure control condition received a brochure at baseline containing information on the Female Athlete Triad. Follow-up assessments were conducted for both groups at 3 weeks (immediately post-intervention for those in the FAB condition), 6, 12, and 18 months post-enrollment date in the study (see Fig. 1). Athletic departments often chose to provide programming to all athletes on a semi-required basis (e.g., required unless given an excused absence). In accordance with community participatory research (CPR) methodology, which promotes shared decision making between researchers and communities as well as respect for core community values, we felt it important to respect this community value. Thus, the "program", which consisted of randomized delivery of FAB or waitlist brochure control, was separated from the research study, which consisted of only the assessments (self-report questionnaires and phone interviews). This allowed athletics staff the option of semi-requiring athletes' participation in the program, while allowing student athletes to opt out of the associated research study.

2.2. Study sites

Study sites were not included or excluded based on National Collegiate Athletic Association (NCAA) Multidivisional Classification. Multidivisional classification is a system by which schools are categorized into "divisions" based on level of sports competition, number of sports teams, and the amount of sports-related scholarships they are permitted to offer students. These divisions are ranked from one to three. Schools in division one (D-I) are the most competitive and sports-focused, while division three (D-III) schools are the least competitive, focusing more on student-athlete experience [16]. Of note is that study sites are given generic names in this paper to allow the specific participating universities to remain anonymous.

2.2.1. Baton Rouge, LA

The Pennington Biomedical Research Center (Pennington Biomedical) served as the coordinating center for the study. A large
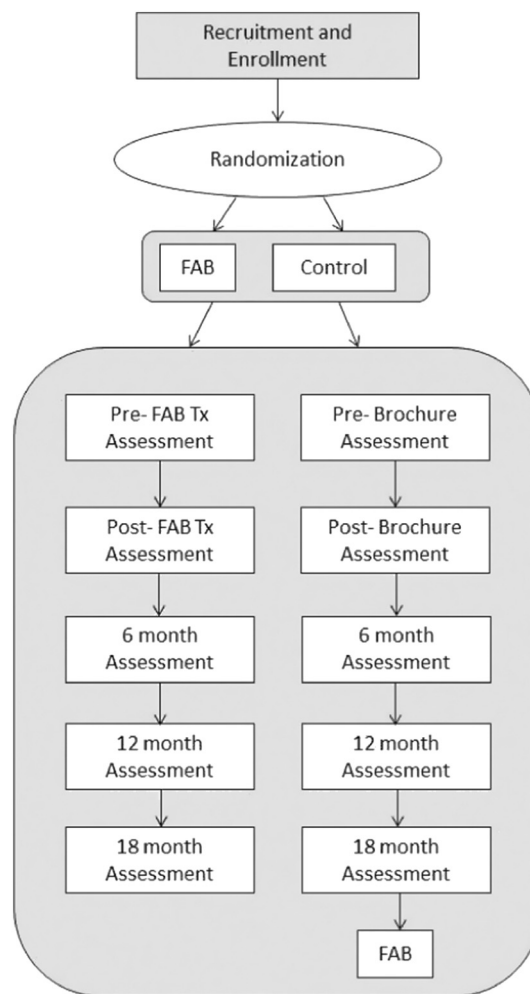

Fig. 1. Study design.

state university, "Louisiana University" (LAU), which consists of 10 NCAA D-I women's teams, served as a participant study site (see Table 1 for team list).

2.2.2. San Antonio, TX

A small private college with nine D-III women's athletic teams served as the anchor participant site in Texas, "Texas University 1" (TXU1), and two teams at from a second small school, "Texas University 2" (TXU2), in the same city served as a secondary data collection site (Table 1). During the course of the study, the second school transitioned from a D-II to a D-I university.

2.2.3. Washington, DC

A mid-size private university, "Washington DC University" (WDCU), served as the participant site in Washington, DC. Participants included seven D-I women's athletic teams.

2.3. Institutional Review Board Approval

The Institutional Review Board (IRB) at Pennington Biomedical served as the IRB of reference. All sites' IRBs approved the study.

2.4. Study population and recruitment

2.4.1. Eligibility

Eligibility criteria remained constant across all sites. Criteria required participants to be: a) female, b) a member of a university-sponsored athletic team, c) willing and able to provide informed

Table 1
Study recruitment results by team and school at baseline.

	LAU	TXU1	AU	TXU2	Total
Athletes eligible for study by team[a]					
Basketball	10	19	12	0	41
Cheer	21	13	0	0	34
Field hockey	0	0	18	0	18
Golf	6	6	0	0	12
Gymnastics	16	0	0	0	16
Lacrosse	0	0	20	0	20
Soccer	20	26	27	0	73
Softball	19	22	0	0	41
Swim/dive	31	27	26	0	84
Tennis	6	13	0	0	19
Track, field, cross country	45	34	17	24	120
Volleyball	14	16	17	14	61
Total	188	176	137	38	539
Recruited peer leaders					
FAB (no control)	14	20	11	4	49
Athletes enrolled in FAB by team					
Basketball	0	12	0	0	12
Field hockey	0	0	14	0	14
Golf	6	0	0	0	6
Soccer	19	25	0	0	44
Softball	18	20	0	0	38
Swim/dive	0	23	27	0	50
Tennis	5	13	0	0	18
Track, field, cross country	42	0	0	24	66
Volleyball	0	0	15	0	15
Total athletes in FAB	90	93	56	24	263
Athletes enrolled in waitlist control by team					
Basketball	9	0	13	0	22
Cheer	20	13	0	0	33
Golf	0	6	0	0	6
Gymnastics	14	0	0	0	14
Lacrosse	0	0	22	0	22
Soccer	0	0	16	0	16
Swim/dive	23	0	0	0	23
Track, field, cross country	0	31	10	0	41
Volleyball	14	13	0	14	41
Total athletes in control	80	63	61	14	218
Total athletes enrolled	170	156	117	38	481

Note. Achieved 90% of target enrollment. After cluster randomization, FAB participants numbered 263, waitlist control participants numbered 218.

[a] Number available after peer leader numbers removed from total available.

consent, and d) able to complete questionnaires at multiple time points and comply with the study protocol. To maximize generalizability, there were no exclusion criteria.

2.4.2. Program recruitment

Athletics staff at all sites chose to semi-require participation in the program (i.e., receiving FAB or brochure), while the research study (i.e., the assessments) remained optional. This design has been used in similar research with other hierarchical communities [2,5]. Athletic staff notified participants about the program, as well as the program attendance expectations, during team meetings set up by coaches or training staff.

2.4.3. Peer-leader recruitment

Coaches and/or head athletic trainers nominated student-athletes as potential FAB program peer-leaders based on perceived reliability, leadership skills, and potential to be good role models. To encourage participation, experimenters explained the benefits of peer leadership for both résumé building and contributing to athletics. Peer leaders also received incentives (iPod shuffle, tote bag, water bottle) for participation. At least one, and ideally two, peer-leaders from each team were recruited. Additionally, in order for 2–3 peer-leaders to run all FAB groups, we supplemented groups with peer-leaders from other sports. Thus, each group was run by at least one peer-leader from the same

sport augmented by 1–2 peer leaders from another sport. For example, when soccer teams completed FAB, at least one peer-leader was a soccer athlete.

2.4.4. Randomization procedures

Group (cluster) randomization was utilized. Twenty-eight athletic teams across the three sites received random assignment to either FAB or waitlist brochure control. Because some sports are at higher risk for ED symptoms and body dissatisfaction (e.g., gymnastics at higher risk than volleyball; [18], we stratified randomization by sport type to ensure that the highest risk teams were not all randomized to one condition. Thus, randomization was stratified by site, sport type (i.e., higher risk versus lower risk), and team size (to balance numbers to each condition). The SAS procedure PROC PLAN was used to generate pseudo-random numbers via the SAS RANUNI function and the randomization schedule.

2.4.5. Study participant recruitment

Study staff recruited study participants. Although athletics staff coordinated meetings with study staff, no athletics staff members were present during study recruitment, in order to reduce coercion. Experimenters repeatedly reminded athletes that study participation was both voluntary and anonymous. Coaches never learned whether athletes opted in or out of study participation, and there was no consequence for not participating. When filling out self-report measures, all team members sat in a circle with their backs to one another to reduce coercion from teammates and the likelihood that teammates knew who was or was not participating. During this time, athletes choosing to enroll in the study filled out the forms. Athletes who opted out of enrollment sat quietly while pretending to fill out the forms. Then, all athletes returned the completed or blank surveys. In addition, athletes who could not attend in person sessions were sent a link to complete questionnaires electronically. To further insure anonymity, phone interviewers were blinded to participants' identities. Finally, participants created self-generated ID codes so that data was kept as anonymous as possible.

Provided incentives included Amazon gift cards valued at $20 for completion of each questionnaire packet and $30 for each phone assessment at each of the five time points. Thus, participants could earn a total of $250 if they completed all assessments at all of the time points. This is in compliance with NCAA Bylaw 16.11.1.10.2, which allows institution-based research studies to compensate student-athletes for participation in a research study involving only student athletes [15].

2.4.6. FAB content

The FAB program content was created by significantly modifying the Healthy Weight intervention originally developed by Stice and colleagues (see [23] for research support) to meet the unique needs of female athletes. Over the course of 3 weekly, 80 min group sessions consisting of 5 to 8 athletes, participants received information and guidance to help them engage in a lifestyle that promotes attainment of a healthy weight via pursuit of the healthy-ideal, rather than the sport-specific thin ideal. The healthy ideal is defined as: "however your unique body appears when you are doing the behaviors necessary to appropriately and simultaneously maximize your physical health, mental health, quality of life and athletic performance". Core elements of the FAB program include: a) defining and identifying differences between the healthy-ideal and the appearance ideals for women in everyday life and in sport; b) providing education about the Female Athlete Triad, nutrition, the concept of balancing input and output to achieve and maintain the healthy-ideal, sleep, and exercise (particularly out of athletic season); c) identifying healthy and unhealthy patterns of behavior; d) goal setting to help apply session information in daily life; and e) three body image exercises from the intervention in the pilot study (e.g., practicing assertively challenging negative body statements in the locker room). "Homework assignments" facilitated the

application of the information between sessions. The specific topics covered in each session are listed below:

2.4.6.1. Session one. a) The concept of embracing the way one looks at optimal health; b) defining the traditional thin-ideal Western standard of female beauty, the sport-specific thin-ideal appearance standard, and the healthy ideal; c) identifying benefits of pursuing the healthy ideal, d) discussing how small but consistent lifestyle changes to balance energy intake and energy output can help maintain a healthier body weight/satisfaction; e) defining and discussing the Female Athlete Triad and associated health consequences; f) a public voluntary commitment to make lasting diet and activity level changes to achieve a healthy body weight; g) advantages of increasing the nutrient density of one's diet; h) identifying small healthy lifestyle changes as a goal; and i) homework to complete a 2-day food diary and exercise log as well as a mirror exposure task aimed at enhancing body image.

2.4.6.2. Session two. a) A review of the homework and identification of target behaviors for change; b) identifying solutions to behavioral change barriers; c) identifying ways to increase the nutrient density of a meal; d) the benefits of exercise; e) benefits of getting enough sleep; f) setting goals for FAB program homework; g) additional information on nutrient-dense foods; and h) writing a letter to a younger athlete encouraging her to give up pursuit of the sport-specific thin-ideal.

2.4.6.3. Session three. a) A review of the homework; b) introduction of the concept of athlete fat talk (e.g., "Do I look fat in this uniform" and "If I lost weight, I'd perform better athletically") and role plays to practice refuting it c) setting goals; d) strategies for healthy eating when traveling to competitions; e) generating a list of personal reasons to pursue the healthy ideal; f) barriers to a healthy lifestyle; and g) brainstorming team efforts towards health.

2.4.7. Study safety

Participating athletic departments reported confidence in their ability to identify cases of EDs with their existing contingency plans and did not want to use the study to augment those plans. Further, they expressed a strong desire to make data collection as anonymous as possible to reduce coercion. Thus, consistent with principles of CPR, instead of using data to identify ED cases to be reported back to athletics, we prioritized anonymity in data collection to reduce possibility of coercion. Participants were provided with local referral information for ED treatment providers to provide a second layer of support beyond athletic department contingency plans. The present study was monitored annually by a Data and Safety Monitoring Board at Pennington Biomedical.

2.4.8. Peer-leader training

Training occurred during athletes' off-season. Peer-leaders attended two four-hour experiential training sessions; under supervision, they rotated through leading an abbreviated version of the sessions. Each peer-leader led a session, received supervision, and participated in sessions while other peer-leaders led and received supervision. This training model has been successfully used in other studies [2–5]. The project manager for each site scheduled peer-leader training; prepared and organized training materials; and led structured/scripted peer leader training sessions.

2.5. Measures

2.5.1. Demographics

Participants self-reported demographics at baseline. Data collection included age, height, weight, race, ethnicity, parental education.

2.5.2. Questionnaires
2.5.2.1. Eating disorder examination questionnaire (EDE-Q). The EDE-Q

[9] is a self-report version of the Eating Disorder Examination (EDE). The EDE-Q assesses eating attitudes and behaviors over a 28-day period. It has 4 subscales: restraint, weight concern, eating concern, and shape concern. This scale demonstrates internal consistency and reliability [24]. Internal consistency for the present sample was high for all subscales and the total score (α range = 0.77–0.94). One month diagnoses can also be generated.

2.5.2.2. Ideal-body stereotype scale-revised (IBSS-R). Internalization of the traditional thin-ideal was assessed with the IBSS-R [22]. Participants responded to questions assessing their perceptions of the ideal body type. This scale demonstrates internal consistency, 2-week test-retest reliability, predictive validity for bulimic symptom onset [23], and sensitivity to detecting intervention effects [20]. Internal consistency for the present sample was high (α = 0.84).

2.5.2.3. Internalization of sport-specific thin-ideal. Because athletes often face pressure to obtain a sport-specific body type, the IBSS-R [22] was modified to be suitable for the sport-specific thin-ideal. In consultation with a leading expert in the field of EDs and athletes, we created a 19-item measure modeled off the IBSS-R. Items focused on the perceived benefits of particular physical features with regards to sport performance. Internal consistency in the current sample was high (α = 0.84).

2.5.2.4. Positive and negative affect scale-revised (PANAS-X)\. The sadness, guilt, and fear/anxiety subscales of the PANAS-X [27] were used to assess the intensity of negative emotional states. This scale demonstrates internal consistency and 2-month test-retest reliability [27]. Internal consistency in the present sample was high (α = 0.93).

2.5.2.5. Health survey utilization scale (HSUS). The HSUS [13,23] assesses the frequency of one's usage of health and mental health services. Participants reported number of hours spent speaking to a variety of providers. Providers were split into categories for physical and mental health, eating disorders, weight problems, and other personal problems. Participants provided responses for the previous month and then for the previous year. The scale has demonstrated acceptable reliability and 20-week test-retest reliability [23]. Internal consistency in the present sample for both monthly and yearly totals (α = 0.77–0.87) was high.

2.5.2.6. Perceived credibility and expectancy. We assessed perceived intervention credibility and expectancy for improvement via a 5 of 6 items of the Credibility/Expectancy Questionnaire (CEQ; [7]). For the present study one total score was created. The CEQ has good internal consistency and test-retest reliability [7]. Internal consistency in this sample was high (α = 0.89).

2.5.2.7. Knowledge of the female athlete triad. This measure assesses knowledge of the Female Athlete Triad with a set of 10 true/false or multiple-choice items. Percentage of correct answers assessed knowledge of the Triad. Internal consistency in the present sample was low (α = 0.49). This is likely due to the fact that this measure does not measure a single construct but rather assesses knowledge across the multiple domains of the Triad.

2.5.2.8. Self-reported weight. Participants provided self-reported height and weight. Although collection of objective height/weight is considered optimal, we chose self-report due to coaches/trainers' concerns about weighing athletes.

2.5.3. Other measures
2.5.3.1. Eating disorder examination (EDE). A trained, masters level, clinical psychology research associate with extensive clinical assessment experience administered the EDE [10] for the study via

telephone. The EDE interview is widely viewed as the most reliable and valid diagnostic assessment of EDs. The interviewer used a brief, adapted form consisting of the diagnostic items of the EDE to assess DSM-IV ED symptoms over the previous month, rather than the previous three months [8]. Internal consistency in the present sample was high ($\alpha = 0.80$). The EDE also provides diagnoses; it demonstrates high test-retest reliability for threshold or subthreshold diagnoses of anorexia nervosa, bulimia nervosa, and binge eating disorder ($r = 0.96$), as well as high inter-rater agreement ($\kappa = 0.86$; [9]). Fortunately, the diagnostic items from the shortened, one-month DSM-IV version can be used to generate diagnoses from DSM-5, which came out during the course of the study. Thus, diagnoses were ultimately based on DSM-5.

2.5.3.2. Female athlete triad identification. The number of athletes who self-identified with the Triad was assessed by asking athletics staff how many athletes had come forward with concerns about the Triad. To get a baseline measure of Triad identification, any incidences of athletes voicing these concerns to the coaches over the previous year were discussed.

2.5.3.3. Intervention fidelity. To promote adherence, each session followed a detailed intervention manual. All sessions were audiotaped using digital voice recorders, and two independently trained raters reviewed and rated a randomly selected 50% of sessions for adherence to the intervention protocol. Raters were required to rate a series of training tapes to establish inter-rater reliability ($k > 0.85$) before rating sessions. Each session's protocol adherence was measured via a detailed session-specific checklist for the concepts, skills, and exercises outlined in each session. Each item was rated for fully completed, mostly completed, somewhat completed, and did not complete at all on a 1–4 Likert Scale. Past studies have found this scale to show inter-rater agreement (ICC = 0.72; [17]).

We also assessed competence with a 12-item measure developed and used by Stice and colleagues in their randomized prevention trials with non-athletes [21]. Items such as "leaders express ideas clearly and at an appropriate pace," "leaders keep group members on task during discussion," and "leaders solicit feedback" were rated on a 1–5 Likert scale anchored by poor (rating of one: "leaders do not ask for feedback to determine member's understanding of, and response to, the session. They also are not able to respond to feedback throughout the session") and superior (rating of five: "leaders always solicit feedback from several group members to ensure that material is clearly understood and respond expertly to verbal and non-verbal feedback throughout the session").

2.5.3.4. Assessment schedule. Participants completed assessments at baseline, post-intervention (3 weeks), and at the 6-, 12-, and 18-month follow-ups. At each assessment, participants completed self-report measures, as well as the EDE telephone interview. All assessments were conducted at each time point with the following exceptions: demographics at baseline only; intervention suitability/expectations at baseline and post-test only; qualitative feedback on programs and cross contamination checks at post-test only; and CBIQA and TAS at time points 12 and 18 months only.

2.6. Statistical analyses

2.6.1. Analysis of baseline assessment data

A random intercept linear mixed effects model with Team as the cluster variable and age, race/ethnicity and BMI as covariates was used to estimate differences in group means for each outcome variable. Race and BMI are commonly used as covariates to account for possible differences in responses. To test for significant differences between the means of each group, we used Z-tests for beta weights of treatment effect on dependent variables at level 2 using $p < 0.05$. For nominal

variables, we used logistic adaption of the same model, and Wald tests were used to assess model adjusted risk (binomial proportion).

2.6.2. Case status description

Case status was determined using both EDE and EDE-Q data. Clinical diagnoses were determined using DSM-5 criteria, while sub-threshold definitions were based on Taylor et al. [25]. The EDE classifies "binge eating" into two categories: subjective binge episodes (SBE), defined as episodes of eating in which a subjectively large amount of food is consumed, accompanied by feelings of loss of control; and objective binge episodes (OBE) in which an objectively large amount of food is consumed, accompanied by feelings of loss of control. The EDE criteria for subthreshold bulimia nervosa (BN) were defined as meeting all objective binge episode and compensatory criteria for the previous one or two months but not the third. Weight or shape concern scores meeting diagnostic threshold (i.e., 4 or greater), six episodes of objective or subjective bulimic episodes across three months, and any compensatory behavior were also categorized as subthreshold BN diagnosis.

Subthreshold binge eating disorder (BED) was defined similarly to BN. Those diagnosed as subthreshold BED were those that met the full BED diagnostic criteria for the previous one or two months, but not the third; or scored at or above a four on weight or shape concerns and had six episodes of binge eating over the previous three months. Subthreshold purging disorder was defined as those who scored greater than or equal to four on weight/shape concern, and whose number of compensatory episodes over the previous three months totaled at six or more. Disordered eating was defined as those who scored a one or two on dietary restriction, reported any objective or subjective bulimic episodes, any vomiting, any laxative use, or any diuretic use over the previous three months.

For the EDE-Q, the same criteria applied for subthreshold diagnoses; however, the EDE-Q measures behaviors only over the previous 28-day period. Therefore, all requirements were halved (e.g., the criteria for BN was reduced to three binges/compensatory behaviors over one month, etc.), and the driven exercise criteria was removed as this is harder to assess via self-report in athletes.

3. Results

3.1. Recruitment

At baseline, 584 athletes were available across all sites for participation in the program as either participants or peer leaders (see Fig. 2 for baseline consort results). After peer-leader recruitment, 539 athletes remained eligible for program participation (Table 1). Of these athletes, 482 athletes participated in the program, and 481 participated in the study; this represents 96% of targeted enrollment. See Table 1 for the number of athletes enrolled in the study by team and site. After cluster randomization, FAB participants numbered 263 and control numbered 218.

3.2. Baseline demographics

Analyses of demographic data revealed no significant differences between FAB and control groups on age, height, and weight (Table 2). For BMI, a small, non-significant difference emerged (Table 2). Race and ethnicity frequencies were not significantly different between groups and no significant differences emerged on levels of parental education (Table 3).

3.3. Baseline values for assessment measures

Using Z-tests from a linear model, with age, race/ethnicity and BMI as covariates, the control group and FAB demonstrated no significant differences on the EDE-Q, the IBSS-R, and Internalization of the Sport
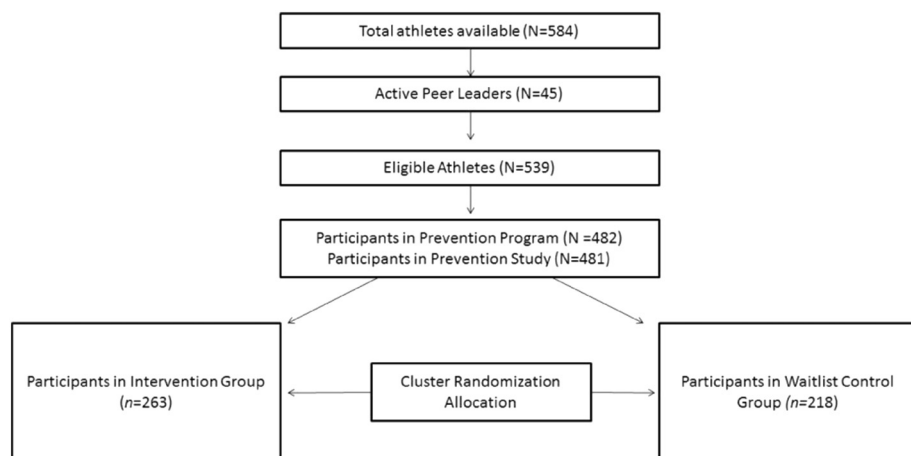
Fig. 2. Consort diagram.

Table 2
Mean differences between groups on demographic measures.

Measures	Overall	FAB	Waitlist control	p
Age (years)	19.35 (1.22)	19.29 (1.25)	19.42 (1.18)	0.262
Height (inches)	66.05 (3.68)	65.95 (3.39)	66.18 (4.00)	0.618
Weight (pounds)	139.90 (23.04)	140.80 (22.47)	138.80 (23.72)	0.969
BMI	22.50 (2.86)	22.74 (3.01)	22.23 (2.65)	0.260

Note. Standard deviations appear in parentheses below means. Wald tests *p* values reported for treatment effect controlling for race and ethnicity.

Specific Thin Ideal at baseline. The FAB group scored significantly higher on the HSUS subsection of eating disorder for the previous month and weight services usage for the previous year. No significant differences were found for the EDE-Q total score, the remaining EDE-Q subscales, or any of the other self-report measures at baseline (Table 4). This lack of differences demonstrates that both groups showed similar levels of ED behaviors, beliefs about ideal bodies in daily life and in

sport, negative affect, and knowledge of the Triad at baseline.

The EDE and EDE-Q were utilized to assess the frequencies of various ED behaviors among participants. Tests of equality between frequencies in each group revealed no significant differences (Table 5). Significantly more Objective Binge Episodes (OBEs) were reported by FAB participants than in the control condition on the EDE-Q; on the EDE, significantly more control participants met diagnostic criteria for weight and shape concerns compared to FAB participants. DSM-5 ED cases based on the EDE and EDE-Q are presented in Table 5.

Overall, participants reported more ED behaviors on the EDE-Q than the EDE; similarly, more participants met criteria for an ED on the EDE-Q. In sum, OBEs and weight/shape concerns differed significantly by group on one measure, but the lack of significant differences for sub-threshold or full diagnoses indicates an adequate randomization of ED behaviors across both groups.

3.4. Intervention fidelity

All rated sessions produced a median adherence rating of 4. Adherence means for session 1 ranged from 3.56–3.94; mean ratings for session 2 ranged from 3.43–3.87 with the exception of one outlier

Table 3
Ethnicity and race response frequencies by group & parental education.

	FAB		Waitlist control		β_{logit} (SE)
Ethnicity response					
Hispanic or Latino	37 (15.29%)		27 (12.86%)		−0.32(0.47) p = 0.49
Not Hispanic or Latino	205 (84.71%)		183 (87.14%)		
Race response					
American Indian or Alaska Native	2 (0.81%)		3 (1.45%)		−0.60(0.99) p = 0.55
Asian	7 (2.82%)		2 (0.97%)		1.14(0.82) p = 0.21
Black or African American	43 (17.34%)		28 (13.53%)		0.31(0.61) p = 0.62
Native Hawaiian or Other Pacific Islander	3 (1.21%)		3 (1.45%)		−0.15(0.75) p = 0.84
Caucasian	193 (77.82%)		171 (82.61%)		
Parental education response	FAB mother	FAB father	Waitlist mother	Waitlist father	
High School Graduate	12 (5.5%)	20 (9.2%)	20 (7.6%)	34 (12.9%)	
Some College	51 (23.4%)	42 (19.3%)	70 (26.6%)	50 (19.0%)	
Bachelors	86 (39.4%)	73 (33.5%)	85 (32.3%)	68 (25.9%)	
Some Graduate School	11 (5.0%)	9 (4.1%)	9 (3.4%)	11 (4.2%)	
Masters	46 (21.1%)	43 (19.7%)	47 (17.9%)	56 (21.3%)	
Doctorate	9 (4.2%)	28 (12.8%)	20 (7.6%)	35 (12.4%)	

Note. Ethnicity: *B* = − 0.32 (*SE* = 0.47), *p* = 0.49. Race: *B* = 1.34, (*p* = 0.18). Mother Education: X^2 = 11.29, df = 6, *p* = 0.08. Father Education: X^2 = 7.62, df = 6, *p* = 0.27. Numbers in parentheses indicate column percentages.

Table 4
Baseline model of primary measures using mixed effects linear model.

Measure	FAB	Waitlist control	β_{diff}	p
EDE-Q weight concern	1.37 (1.32)	1.61 (1.36)	−0.18(0.13)	0.151
EDE-Q eating concern	0.59 (0.82)	0.68 (0.92)	−0.03(0.08)	0.731
EDE-Q shape concern	1.73 (1.34)	1.98 (1.36)	−0.15(0.14)	0.280
EDE-Q restraint	1.32 (1.30)	1.38 (1.32)	02 (0.11)	0.855
EDE-Q total score	1.25 (1.05)	1.41 (1.12)	−0.08(0.11)	0.440
IBSS-R short	3.57 (0.68)	3.70 (0.64)	−0.09(0.07)	0.200
Internalization of sport thin ideal	2.59 (0.59)	2.81 (0.64)	−0.22(0.11)	0.053
PANAS-X average	1.64 (0.64)	1.70 (0.64)	−0.06(0.06)	0.331
HSUS previous month				
# Physical absence	1.22 (2.33)	1.00 (1.59)	0.09(0.15) −0.30(0.30)	0.564 0.319
#Mental absence	0.52 (1.84)	0.28 (0.95)	0.08(0.26) −0.60(0.42)	0.758 0.150
#Weight absence	0.18 (1.24)	0.18 (0.75)	−0.07(0.41) 0.28(0.59)	0.861 0.642
#Eating disorders absence	0.17 (1.49)	0.07 (0.59)	−0.83(0.38) −1.26(1.01)	0.027* 0.213
#Other absence	0.32 (1.02)	0.33 (1.02)	0.01(0.20) 0.13(0.37)	0.983 0.725
HSUS previous year				
#Physical absence	3.50 (3.89)	3.18 (3.11)	08(0.10) −0.21(0.36)	0.430 0.572
#Mental absence	1.07 (2.79)	0.63 (1.85)	0.19(0.19) −0.58(0.31)	0.312 0.062
#Weight absence	0.44 (1.54)	0.30 (1.13)	0.86(0.38) 0.10(0.53)	0.023* 0.856
#Eating disorders absence	0.32 (1.37)	0.14 (0.94)	0.36(0.59) −0.71(0.59)	0.539 0.229
#Other absence	0.85 (1.91)	0.56 (1.52)	0.19(0.18) −0.27(0.32)	0.281 0.410
Perceived credibility & expectancy	28.69 (9.06)	25.47 (8.81)	2.97(1.61)	0.065
Knowledge of triad	8.29 (1.58)	8.29 (1.45)	0.25(0.16)	0.111

Note. p value set to ≤ 0.05. All β adjusted for BMI, Age, Race, Ethnicity, and cluster level for team. Count variables estimated using Zero Inflated Poisson Regression. Logit scale used for absence.

 * Represents significant differences between groups ($p < 0.05$).

session that was rated 3.17; and mean ratings for session 3 ranged from 3.59–4.00 with one outlier of 3.03. On the basis of these ratings, we deemed overall adherence as acceptable.

For competence, all rated sessions received a median rating of 3 or 4 with the exception of one session which received a median rating of 5. These scores are equivalent to "good/average" and "excellent/above average." Mean competence ratings for session 1 ranged from 3.00–4.73, with the exception of a single session (M = 2.73). For session 2, mean ratings ranged from 3.18–4.36 with one exception (M = 2.91), and mean ratings for session 3 ranged from 3.00–4.36.

4. Discussion

The present paper outlines the study design, methods, and baseline participant characteristics for the Female Athlete Body Project (FAB) study. The study investigates the efficacy of a behavioral intervention designed to improve body satisfaction, promote awareness and prevention of the Female Athlete Triad, and reduce evidence-based, modifiable risk factors for EDs among female collegiate athletes. This study is best classified as an efficacy/effectiveness hybrid trial given

that we sought to evaluate efficacy under sustainable conditions by having low-cost community members (i.e., peer leaders) implement the actual intervention. Future papers will present the outcomes of the FAB intervention for female collegiate athletes.

The present study achieved 96% of the goal sample recruitment. We also recruited an adequate number of peer-leaders to successfully deploy FAB across multiple sites. We utilized CPR research methodology in this program of research, which may have contributed to the high level of cooperation by participating athletics departments as well as successful recruitment. As noted above, CPR methods involve sharing power and decision making with community members. As such, the trial was designed in collaboration with athletic department staff and several core features were included in response to staff opinions. Example features included making data anonymous and using a peer-leader delivery model. Athletics wanted peer delivery for two reasons. First, peers are low cost which makes implementation more financially feasible. Second, peer delivery creates leadership opportunities for student athletes, which fits with part of the educational mission of collegiate athletics.

CPR methods have been successfully used to engage other communities in the delivery of prevention interventions [1] and were vital in establishing the requisite partnership for the both the pilot and present trials [3]. Because communities (including athletics departments) may be less receptive to programming that is perceived as coming from outside the community, future researchers should continue to incorporate CPR methods to enhance community stakeholder engagement and partnership. Use of CPR methods can also help researchers tailor interventions to make them more appealing to community stakeholders. In terms of group differences, BMI emerged as marginally statistically significant and not clinically significant (mean difference of 0.51 kg/m^2). The two groups were similar on all other demographic categories. Baseline ED diagnostic assessments also revealed few differences between the groups in ED behaviors and diagnoses. From an overall baseline perspective, the total sample had a relatively low percentage of ED symptoms, as well as a low incidence of full diagnostic cases given previous reports in the literature. One possible reason for this may be the fact that athletic trainers often exempted athletes from participating in the program if they had an active ED and were in treatment. Low rates do not necessarily pose a problem. Despite the low numbers of ED diagnoses, this cohort nonetheless represents a large and diverse population of individuals at risk of ED development. The FAB program is not meant to provide support to those already struggling with EDs; rather, FAB is a prevention intervention designed to mitigate ED development due to the risks of the athletic environment. This population with minimal existing diagnostic cases of EDs is ideal for examining whether the FAB intervention prevents at-risk participants from worsening in the relevant ED symptoms and risk factors over time.

It should be noted that rates of behaviors and diagnostic cases were higher on the EDE-Q than the EDE. One possible reason for this is that fewer participants completed the EDE at baseline as compared to the EDE-Q. Those with ED behaviors may have opted out of the interview. Similarly, the increased privacy associated with a self-report measure may have increased disclosure. This seems a possibility given that the interviewer reported that some students seemed to complete their interviews in less than optimally private locations (e.g., in a public place when on the phone). Lastly, the interviewer may have been more adept at distinguishing between true ED behaviors. The fact that rates of more subjective behaviors (e.g., binge episodes) showed greater discrepancy than more objective behaviors (e.g., vomiting) provides support for this hypothesis.

Strengths of the current study include the utilization of both self-report and clinical interview data, multiple sites of data collection, inclusion of diverse sports and divisions of female collegiate athletics, cluster randomization balanced by site and sport type, and follow-up through 18-months. The study also utilized CPR methods to enhance

Table 5
ED behaviors and diagnoses, frequency (percent).

	FAB	Waitlist control	β	p	β_{logit}	p
ED behaviors measured by EDE-Q						
Laxative use	2 (0.76)	10 (4.61)	− 1.70 (12.92)	0.900	0.380 (11.91)	0.980
Vomiting	9 (3.42)	7 (3.23)	− 74(4.56)	0.87	− 0.26(1.62)	0.874
Diuretic use	2 (0.76)	5 (2.30)	− 1.92 (6.99)	0.78	0.75(4.46)	0.867
Subjective bulimic episodes	46 (17.49)	36 (16.59)	− 0.10 (0.12)	0.42	0.04 (0.53)	0.943
Objective bulimic episodes	35 (13.31)	30 (13.82)	− 0.39 (0.10)	< 0.001	0.15(0.33)	0.653
Objective overeating episodes	46 (17.49)	43 (19.82)	0.94 (0.48)	0.05	1.27 (1.17)	0.278
Weight/shape scores > 4	25 (9.51)	28 (12.9)	n/a	n/a	0.11 (1.11)	0.921
ED behaviors measured by EDE						
Laxative use	0	0	NA			
Vomiting	4 (1.83)	1 (0.77)	0.117	0.732		
Diuretic use	0	0	NA			
Subjective bulimic episodes	3 (1.38)	4 (3.08)	0.489	0.485		
Objective bulimic Episodes	3 (1.38)	0	0.554	0.457		
Objective overeating episodes	2 (0.91)	0	0.131	0.717		
Weight/shape scores > 4	41 (18.81)	47 (36.15)	11.05	< 0.001		
ED diagnoses measured by EDE-Q						
Purging disorder	17 (6.46)	14 (6.45)	0.001	0.971		
Bulimia nervosa	3 (1.14)	6 (2.76)	1.03	0.308		
Anorexia nervosa	1 (0.38)	0	0.815	0.367		
Binge eating disorder	5 (1.90)	2 (0.92)	0.240	0.624		
Subthreshold binge eating disorder	13 (4.94)	8 (3.69)	0.170	0.680		
Subthreshold bulimia nervosa	2 (0.76)	0	0.320	0.572		
ED diagnoses as measured by EDE						
Purging disorder	4 (1.83)	1(0.77)	0.0.110	0.740		
Bulimia nervosa	1 (0.46)	0	0.0.598	0.439		
Anorexia nervosa	1 (0.46)	0	0.598	0.439		
Binge eating disorder	0	0	NA			
Subthreshold binge eating disorder	0	0	NA			
Subthreshold bulimia nervosa	0	0	NA			

Note. p values and β_{logit} not calculated for EDE behaviors and diagnoses and EDEQ diagnoses secondary to insufficient cases. Chi square with Yates correction calculated descriptive purposes. EDE denominator FAB = 218, WC = 130.

athletic department buy-in and to create a study that delivered the intervention in a manner consistent with how athletic departments would be likely to implement FAB, were the program to be disseminated widely (i.e., using low-cost peer-leaders). The achievement of 96% of target recruitment can be credited to the participation of athletic departments aiding in study recruitment and their semi-mandating participation in the FAB intervention itself. Going forward, engagement of athletic departments will be an invaluable resource for encouraging wide-scale athlete participation and success in this intervention. It is important to note, however, that failure to obtain such community buy-in could result in reduced success with implementation. Limitations of the present study include missing baseline data for the EDE, and inconsistent environments in which participants completed the EDE phone interview.

In conclusion, the FAB study is the largest RCT conducted to date on the prevention of EDs in female collegiate athletes. Given this, it is anticipated that the findings of this study will make a significant contribution to the ED prevention literature. Future analyses will allow for the evaluation of the efficacy of the FAB program in collegiate female athletes and to better understand the public health impact of this approach.

Disclosure statement

Dr. Becker currently serves as a co-director of the *Body Project Collaborative*, a company that facilitates implementation of prevention programs in non-athletes.

Acknowledgements

References

[1] C. Becker, E. Stice, H. Shaw, S. Woda, Use of empirically supported interventions for psychopathology: can the participatory approach move us beyond the research-to-practice gap? Behav. Res. Ther. 47 (4) (2009) 265–274.

[2] C.B. Becker, S. Bull, K. Schaumberg, A. Cauble, A. Franco, Effectiveness of peer-led eating disorders prevention: a replication trial, J. Consult. Clin. Psychol. 76 (2) (2008) 347–354, http://dx.doi.org/10.1037/0022-006X.76.2.347.

[3] C.B. Becker, L. McDaniel, S. Bull, M. Powell, K. McIntyre, Can we reduce eating disorder risk factors in female college athletes? A randomized exploratory investigation of two peer-led interventions, Body Image 9 (1) (2012) 31–42, http://dx.doi.org/10.1016/j.bodyim.2011.09.005.

[4] C.B. Becker, L.M. Smith, A.C. Ciao, Peer-facilitated eating disorder prevention: A randomized effectiveness trial of cognitive dissonance and media advocacy, J. Couns. Psychol. 53 (4) (2006) 550–565.

[5] C.B. Becker, C. Wilson, A. Williams, M. Kelly, L. McDaniel, J. Elmquist, Peer-facilitated cognitive dissonance versus healthy weight eating disorders prevention: a randomized comparison, Body Image 7 (4) (2010) 280–288, http://dx.doi.org/10.1037/0022-0167.53.4.550.

[6] M. De Souza, N. Williams, Physiological aspects and clinical sequelae of energy deficiency and hypoestrogenism in exercising women, Hum. Reprod. Update 10 (5) (2004) 433–448, http://dx.doi.org/10.1093/humupd/dmh033.

[7] G. Devilly, T. Borkovec, Psychometric properties of the credibility/expectancy questionnaire, J. Behav. Ther. Exp. Psychiatry 31 (2) (2000) 73–86, http://dx.doi.org/10.1016/S0005-7916(00)00012-4.

[8] C. Fairburn, Z. Cooper, M.E. O'Connor, Eating Disorder Examination (Edition 16.0D), Cognitive Behavior Therapy and Eating Disorders, The Guilford Press, New

York, NY, 2008, p. 324.

[9] C.G. Fairburn, S.J. Beglin, Assessment of eating disorders: interview or self-report questionnaire? Int. J. Eat. disord. 16 (4) (1994) 363–370, http://dx.doi.org/10. 1002/1098-108X(199412)16:4<363::AID-EAT2260160405>3.0.CO;2-#.

[10] C.G. Fairburn, Z. Cooper, The Eating Disorder Examination (12th Edition), in: C.G. Fairburn, G.T. Wilson (Eds.), Binge Eating: Nature, Assessment, and Treatment, Guilford Press, New York, 1993, pp. 317–360.

[11] P. Hinton, Running on empty. Training and conditioning (15). Retrieved from, http://www.training-conditioning.com/2007/03/running_on_empty.html, (2005).

[12] P. Keel, D. Herzog, Clinical Handbook of Eating Disorders: An Integrated Approach, Marcel Dekker, Inc., New York, 2004.

[13] J.M. Lane, M.E. Addis, Male gender role conflict and patterns of help seeking in Costa Rica and the United States, Psychol. Men Masculinity 6 (3) (2005) 155–168, http://dx.doi.org/10.1037/1524-9220.6.3.155.

[14] M. Mountjoy, J. Sundgot-Borgen, L. Burke, S. Carter, N. Constantini, C. Lebrun, ... A. Ljungqvist, The IOC consensus statement: beyond the Female Athlete Triad–Relative Energy Deficiency in Sport (RED-S), Br. J. Sports Med. 48 (7) (2014) 491–497, http://dx.doi.org/10.1136/bjsports-2014-093502.

[15] NCAA, 2010–11 NCAA Division III Manual, NCAA Academic and Membership Affairs Staff, Indianapolis, IN, 2010.

[16] NCAA, Divisional Differences and the History of Multidivisional Classification, NCAA, 2016.

[17] P. Rohde, G. Clarke, D. Mace, J. Jorgensen, J. Seeley, An efficacy/effectiveness study of cognitive-behavioral treatment for adolescents with comorbid major depression and conduct disorder, J. Am. Acad. Child Adolesc. Psychiatry 43 (6) (2004) 660–668.

[18] L. Smolak, S. Murnen, A. Ruble, Female athletes and eating problems: a meta-analysis, Int. J. Eat. disord. 27 (4) (2000) 371–380.

[19] I.M.C.W.G.W.i. Sport, Position Stand on the Female Athlete Triad: International

Olympic Committee, (2006).

[20] E. Stice, W.S. Agras, Predicting onset and cessation bulimic behaviors during adolescence: A longitudinal grouping analysis, Behav. Ther. 29 (1998) 257–276.

[21] E. Stice, M.L. Butryn, P. Rohde, H. Shaw, C.N. Marti, An effectiveness trial of a new enhanced dissonance eating disorder prevention program among female college students, Behav. Res. Ther. 51 (12) (2013) 862–871, http://dx.doi.org/10.1016/j. brat.2013.10.003.

[22] E. Stice, M. Fisher, E. Martinez, Eating Disorder Diagnostic Scale: additional evidence of reliability and validity, Psychol. Assess. 16 (1) (2004) 60–71, http://dx. doi.org/10.1037/1040-3590.16.1.60.

[23] E. Stice, H. Shaw, E. Burton, E. Wade, Dissonance and healthy weight eating disorder prevention programs: a randomized efficacy trial, J. Consult. Clin. Psychol. 74 (2) (2006) 263–275, http://dx.doi.org/10.1037/0022-006X.74.2.263.

[24] R. Sysko, B.T. Walsh, C.G. Fairburn, Eating Disorder Examination-Questionnaire as a measure of change in patients with bulimia nervosa, Int. J. Eat. disord. 37 (2) (2005) 100–106, http://dx.doi.org/10.1002/eat.20078.

[25] C.B. Taylor, A.E. Kass, M. Trockel, D. Cunning, H. Weisman, J. Bailey, ... D.E. Wilfley, Reducing eating disorder onsent in a very high risk sample with significant comorbid depression: a randomized controlled trial, J. Consult. Clin. Psychol. 85 (5) (2016) 402.

[26] R.A. Thompson, R.T. Sherman (Eds.), Eating Disorders in Sport, Taylor and Francis Group, New York, 2010.

[27] D. Watson, L.A. Clark, A. Tellegen, Development and validation of brief measures of positive and negative affect: the PANAS scales, J. Pers. Soc. Psychol. 54 (6) (1988) 1063–1070, http://dx.doi.org/10.1037/0022-3514.54.6.1063.

[28] S. Wonderlich, J. Mitchell, L. Swan-Kreimer, C. Peterson, S. Crow, Clinical Handbook of Eating Disorders: An Integrated Approach, Marcel Dekker, Inc., New York, 2004.