Morbid exercise behaviour and eating disorders: A meta-analysis

MANUEL ALCARAZ-IBÁÑEZ[1,2] ⓘ, ADRIAN PATERNA[1,2]* ⓘ, ÁLVARO SICILIA[1,2] ⓘ and MARK D. GRIFFITHS[3]

[1] Department of Education, University of Almería, Almería, Spain

[2] Health Research Centre, University of Almería, Almería, Spain

[3] Psychology Department, Nottingham Trent University, Nottingham, UK

REVIEW ARTICLE

ABSTRACT

Background and aims: This study examined the relationship between self-reported symptoms of morbid exercise behaviour (MEB) and eating disorders (ED) using meta-analytic techniques. *Methods:* We systematically searched MEDLINE, PsycINFO, Web of Science, SciELO and Scopus. Random effects models were used to compute pooled effect sizes estimates (r). The robustness of the summarized estimates was examined through sensitivity analyses by removing studies one at a time. *Results:* Sixty-six studies comprising 135 effect-sizes ($N = 21,816$) were included. The results revealed: (a) small-sized relationship in the case of bulimic symptoms ($r = 0.19$), (b) small- ($r = 0.28$) to medium-sized relationships ($r = 0.41$) in the case of body/eating concerns, and (c) medium-sized relationships in the case of overall ED symptoms ($r = 0.35$) and dietary restraint ($r = 0.42$). Larger effect sizes were observed in the case of overall ED symptoms in clinical, younger, and thinner populations, as well as when employing a continuously-scored instrument for assessing ED or the Compulsive Exercise Test for assessing MEB. Larger effect sizes were also found in female samples when the ED outcome was dietary restraint. *Conclusions:* The identified gaps in the literature suggest that future research on the topic may benefit from: (a) considering a range of clinical (in terms of diagnosed ED) and non-clinical populations from diverse exercise modalities, (b) addressing a wide range of ED symptomatology, and (c) employing longitudinal designs that clarify the temporal direction of the relationship under consideration.

INTRODUCTION

The potential health benefits of physical exercise has been well documented (Béland et al., 2019; Lin et al., 2015). However, it has also been demonstrated that some individuals exercise to the point of losing control over such a behaviour, and persist on exercising even when this interferes with their professional and/or social responsibilities, or even being injured (Cook, Hausenblas, & Freimuth, 2014b; Szabo, Demetrovics, & Griffiths, 2018).

Despite the potential harmful health implications of problematic exercise, it has not yet been included as a psychiatric disorder within any officially recognized psychological or medical diagnostic frameworks (Berczik et al., 2012). One of the main reasons for this may be the lack of consensus on its conceptualization and assessment. Consequently, this form of exercise has been named using a range of terms that may ultimately refer to a related but somewhat differentiated phenomena (e.g., excessive exercise, compulsive exercise, exercise dependence, and exercise addiction) (Berczik et al., 2012; Cook, Hausenblas, et al., 2014b; Cunningham, Pearman, & Brewerton, 2016; Szabo et al., 2018). For instance, the term

*Corresponding author's e-mail: a.paterna@ual.es

"compulsive" implies an urge to engage in a non-necessarily pleasant behaviour (in this case, exercising) in order to prevent a perceived negative consequence if the behaviour is not performed (Cook, Hausenblas, et al., 2014b; Starcevic, 2016). Conversely, the term "addiction" here refers to a process that implies repeated failures to control exercise behaviour in spite of its negative consequences and provides either relief from experienced distress or even pleasure itself (Berczik et al., 2012). In the present paper, we use the generic term "morbid exercise behaviour" (MEB) proposed to encompass all these conceptualizations according to its common characteristics (i.e., the presence of an increasingly uncontrollable exercise-related behaviour that, regardless of the effective time spent exercising, involves physical and/or psychological harm) (Szabo et al., 2018).

Another controversial issue concerning MEB is whether this could be considered as an independent psychopathological entity or, on the contrary, a compensatory behaviour that emerges in the course of psychiatric illness such as eating disorders (ED) (Berczik et al., 2012; Starcevic & Khazaal, 2017). Indeed, "excessive exercise" (defined as that which interferes with important activities, occurs at inappropriate times/settings, or is carried out despite injury or other medical complications) is listed in the latest (fifth) edition of the Diagnostic and Statistical Manual of Mental Disorders (DSM-5) as a diagnostic feature for both bulimia nervosa and the restricting subtype of anorexia nervosa (American Psychiatric Association, 2013). In support of this conception, it has been argued that individuals suffering from bulimia nervosa may be worried about the prospect of undesired changes in appearance that, as results of their binge eating, may occur in absence of exercise (Cook, Hausenblas, et al., 2014b). In the case of the restricting type of anorexia nervosa, excessive exercise may serve the purpose of weight loss implied in the former eating pathology. Conversely, MEB has been also proposed as a pathological behaviour with a potentially relevant role in the onset and maintenance of ED (Cook & Hausenblas, 2008, 2011; Cook, Hausenblas, Crosby, Cao, & Wonderlich, 2015).

Evidence linking MEB and ED has been summarized in several narrative and systematic reviews (Fietz, Touyz, & Hay, 2014; Meyer, Taranis, Goodwin, & Haycraft, 2011; Starcevic & Khazaal, 2017). Findings from these studies suggest that MEB tends to be positively associated with (a) overall ED symptoms, (b) symptoms of specific disorders such as bulimia nervosa, and (c) a diagnostic feature of anorexia nervosa such as body/eating concerns. However, these findings have not been without limitations. Examples of the latter are (a) being derived from a low number of primary studies, (b) being restricted in some cases to very specific populations (e.g., adolescents clinically diagnosed with an ED), (c) conceptualizing MEB in terms of the time devoted to the activity or its intensity, and (d) not examining the relationships under consideration by distinguishing between the different conceptualizations and tools proposed for the assessment of MEB. Consequently, it remains unclear whether MEB may be differently related with specific ED or its associated diagnostic features.

An analytical approach that could contribute to overcoming some of these limitations and, by extension, to a better understanding of the relationship between MEB and ED is meta-analysis. By combining data from a series of studies, meta-analytic techniques allow for obtaining population estimates on the level of association between two variables, as well as examining moderators of such a relationship (Borenstein, Hedges, Higgins, & Rothstein, 2009). Despite this, there is only one previous meta-analysis examining the association between MEB and ED (Trott et al., 2020). Findings from this study showed that the risk of MEB was approximately four times higher among individuals with ED vs. those without ED, and that these differences may vary according to the instrument used for the assessment of MEB. However, the conclusions drawn by Trott et al. (2020) are arguably limited in scope due to at least three major constraints: (i) the very specific nature of the study population (i.e., adults with a non-clinically diagnosed ED); (ii) the small number of studies included in the subgroup analyses, a factor that could call into question the accuracy of the results obtained (Fu et al., 2011), and (iii) the fact of operationalizing MEB in terms of its prevalence according to cut-off points that, in absence of official recognition of this potential disorder and its diagnostic criteria, have not been clinically validated (and which could therefore be considered arbitrary).

An updated and comprehensive meta-analytic examination of evidence linking MEB and ED may contribute to a better understanding of such a relationship and to identify potentially insightful avenues of research. Therefore, the aim of the present study was to examine the relationship between self-reported symptoms of MEB and ED using meta-analytic techniques, further considering potential moderators of such a relationship.

METHOD

The systematic review and meta-analysis was conducted in accordance with the checklist from Preferred Reporting Items for Systematic Reviews and Meta Analyses (PRISMA) (Moher, Liberati, Tetzlaff, & Altman, 2009) and was registered on PROSPERO (CRD42019119413) (see Appendix A).

Locating information

Electronic bibliographic databases MEDLINE, PsycINFO, Web of Science, SciELO and Scopus were searched for eligible studies using the following search terms: ("morbid exercise" OR "exercise dependence scale" OR "exercise addiction inventory" OR "exercise addiction" OR "exercise dependence" OR "compulsive exercise test" OR "compulsive exercise" OR "physical activity compulsive" OR "obligatory exercise questionnaire" OR "obligatory exercise" OR "commitment to exercise scale" OR "commitment to exercise" OR "excessive exercise scale" OR "excessive exercise") AND ("eating disorders" OR eating OR "eating behaviour" OR "eating behaviour" OR "eating pathology" OR

"disordered eating" OR bulimi* OR anorexi* OR dietary OR restraint OR "dietary restraint" OR binge OR pica OR rumination OR "drive for thinness" OR "weight concerns" OR "shape concerns" OR "eating concerns"). The search was restricted to studies published between 1988 (the date in which the first psychometric instrument assessing MEB was developed) (Pasman & Thompson, 1988) and 2018 inclusive. Reference lists of retrieved studies were manually reviewed to identify further potentially eligible studies.

Retrieved references were managed in Endnote 7.1. Studies were selected in two stages by the first two authors, specifically by (a) screening the title and abstracts for relevance, and (b) reviewing the full-texts taking into account the inclusion/exclusion criteria. Disagreements between both reviewers were discussed and resolved on a consensual basis with the assistance of the third author.

Corresponding authors of the studies included in the review were approached to request unpublished data that may be potentially eligible according to the objectives of the present study. When relevant information for a given retrieved study was missing (e.g., BMI, age, or that needed to calculate effect-sizes), they were also asked if they could provide it. When no response was received within one month, the authors were contacted again. The response rates (i.e., the percentage of authors that provided data that were effectively analysed) were 3% (unpublished data) and 43% (missing relevant information).

Eligibility criteria

The present study gathered data that provided evidence on the association between MEB and ED, in both cases, as assessed by self-report instruments. With a focus on avoiding publication bias, the intention was not only to retrieve published studies but also to retrieve data from unpublished quantitative research providing relevant effect sizes such as dissertations, non-significant findings excluded from publications, data already collected but not yet published, and data not included in original publications (e.g., those corresponding to sub-populations or reflecting sub-domain scores).

Inclusion criteria. Studies were considered eligible if the following criteria were met: (a) at least one validated self-reported instrument assessing MEB was used; (b) at least one validated self-reported instrument assessing symptoms of a specific ED or diagnostic criteria of the latter proposed in the DSM-5 (American Psychiatric Association, 2013) was used; (c) studies written in English or Spanish (the languages spoken by the authors), although no restrictions in terms of country of origin were considered; and (d) sufficient data to calculate effect size were available.

Exclusion criteria. Studies were excluded on the basis of the following criteria: (a) MEB was only addressed in terms of exercise volume/intensity (e.g., frequency or hours of practice); (b) only composite scores comprising two or more instruments assessing MEB were provided so that individual scores were not available; (c) specific items or factors were excluded when obtaining global scores for MEB and sub-domains scores were not available; (d) scores on MEB were obtained using a partially/completely altered factorial structure from the one originally proposed for the instrument; (e) data concerning clinical and non-clinical population in terms of ED were not available segmented according to this condition; and (f) scores for a given ED outcome were derived from full instruments or sub-factors not consistent with the diagnostic features proposed in the DMS-5 (e.g., those assessing body dissatisfaction).

Coding procedure

A coding frame was firstly developed taking into account the common features of the studies retrieved in a preliminary search, which was subsequently pilot-tested. The resulting coding sheet was used by the first two authors when extracting the relevant data from the studies included in the systematic review (see Appendix B). Disagreements between both reviewers were discussed and resolved on a consensual basis. To determine the inter-coder reliability, Cohen's Kappa (range: 0.63 to 0.84, percent agreement: 87–94%) was calculated using the reliability calculator 'ReCal' (Freelon, 2013). The following coding categories were considered: (a) citation and year of publication; (b) sample size; (c) sample type (clinical vs non-clinical); (d) sex; (e) age; (f) study quality; (g) body mass index (BMI); (h) publication status (published vs. unpublished); (i) MEB measure; (j) ED measure; (k) ED outcome; (l) ED assessment (i.e., continuous vs. categorical); and (m) effect size of the correlation between MEB and ED. These coded features were considered for descriptive purposes, as well as potential moderator variables (Rosenthal, 1995).

Risk bias

Assessment of risk bias was based on the Quality Assessment Tool for Quantitative Studies developed by the Effective Public Health Practice Project (Thomas, Ciliska, Dobbins, & Micucci, 2004). Consequently, the retrieved studies were qualitatively termed as weak, moderate or strong depending, respectively, on the presence of two or more, one, or no weak scores in any of the following components: (a) selection bias; (b) study design; (c) confounders; (d) blinding; (e) data collection methods; and (f) withdrawals and dropouts. The assessment of risk-bias was conducted by the first two authors, with disagreements between both reviewers being discussed and resolved on a consensual basis with the assistance of the third author. As a result of this procedure, 21 studies were rated as strong, 39 as moderate, and 6 as weak in terms of risk bias.

Statistical analysis

Effect size calculations. Pearson's correlation (r) was employed as the effect size index. In the case of studies providing effect-sizes considering scores for all the sub-domains proposed for a given instrument, these were joined to allow for obtaining also effect-sizes corresponding

to global scores. In the case of studies examining the relationship between MED and ED considering the latter as a dichotomous variable (e.g., for the SCOFF, individuals being classified as in risk of an eating pathology on the basis of a number of positive responses ≥2), effect size was derived from mean scores, standard deviations, and sample size. Estimated effect-sizes were r-to-z transformed before conducting the statistical analyses. To facilitate interpretation of the results, effect-sizes and its 95% confidence intervals (CIs) were z-to-r transformed (Borenstein et al., 2009).

Assuming that variations between effect-sizes are due both to variations in its distribution and sampling errors, a random effect model was used to calculate the pooled effect sizes (Pigott, 2012). This model was chosen due to the expected heterogeneity between studies in terms of participants' characteristics (e.g., sex or clinical condition) and exposure/outcomes (Mueller et al., 2018) in order to facilitate the generalization of the obtained findings (Borenstein et al., 2009). Statistical heterogeneity was assessed using the I^2 statistic, with values of 25%, 50%, and 75% suggesting low, moderate, and high heterogeneity, respectively (Higgins, Thompson, Deeks, & Altman, 2003). The robustness of the summarized estimates was examined utilizing a sensitivity analyses (i.e., systematic reanalysis while removing studies one at a time). Results from sensitivity analyses were considered meaningful when corrected estimates were beyond the 95%CI of the original ones. This analysis was also employed to examine whether a particular study may be accounting for a large proportion of heterogeneity.

As long as at least four effect-sizes were available (Fu et al., 2011), moderator analyses employing a mixed-effects model (Borenstein et al., 2009) were conducted for the following categorical variables: ED assessment (categorical/continuous), sex (men/women/both), sample type (clinical/non-clinical), publication status (published/unpublished), study quality (strong/moderate/weak), and MEB measure. In these analyses, analogue to ANOVA analyses were employed to examine both within-group variance (Q_{within}) and between-group variance ($Q_{between}$), with a significant $Q_{between}$ suggesting meaningful differences in the true effect size between groups. Continuous covariates (BMI and age) were examined as potential sources of variance as long as at least 10 effects sizes were available (Fu et al., 2011).

As long as at least 10 effect sizes were available (Page, Higgins, & Sterne, 2019), publication bias was judged on the basis of the visual inspection of a funnel plot (i.e., a lack of symmetry) and Egger's test ($P > 0.10$). Publication bias for a given outcome was corrected using the 'trim and fill' procedure (i.e., by imputing "missing" studies and recalculating a new unbiased summary effect) Duval & Tweedie, 2000 a, b.

Point mean estimates for the estimated effect-sizes were interpreted according to the following guidelines: 0.00 to 0.10 trivial effect; 0.10 to 0.30 small effect; 0.30 to 0.50 moderate effect; and >0.50 large effect (Cohen, 1988). The described statistical analyses were conducted in Comprehensive Meta-Analysis, Version 2.2 (Borenstein, Hedges, Higgins, & Rothstein, 2005).

Dependence. Most of the statistical procedures involved in a meta-analysis require the independence of the considered effect-sizes (Becker, 2000a). In this sense, one of the most common forms of dependence consists of considering multiples effect-sizes from a single sample (Hedges, 2009). Since this was the case in the present paper, this potential source of dependence was treated as follows: (a) the effect-sizes were grouped according to specific outcomes so that differentiated meta-analyses were conducted for each of them (Borenstein et al., 2009); (b) when MEB was assessed using multiple instruments (Cunningham et al., 2016), and given that subgroup analyses according to this feature were planned, random removal of effect sizes was conducted until just one effect size remained (Cheung, 2014); (c) when different effect sizes were provided for several groups in a same study (e.g., men/women), each of these was treated individually (Cheung, 2014); and (d) when a given ED outcome was assessed using multiple instruments (LePage, Price, O'Neil, & Crowther, 2012), the dependent effect sizes were averaged within their respective studies before conducting the analysis (Cheung, 2014).

RESULTS

Selection of studies

A total of 2,029 studies were identified from multiple database search. As a result of the study selection procedure (see Fig. 1), 67 studies were included in the systematic review and 66 in the meta-analysis. The decision of excluding one study from the meta-analysis (Gianini et al., 2016) was adopted in light of its specific research design (i.e., longitudinal). More specifically, given that the low number featuring this condition (i.e., $K < 4$) did not allow for examining study design as a potential source of heterogeneity (Fu et al., 2011).

Description of studies

The study characteristics and their corresponding effect-sizes were grouped taking into account the specific ED and associated diagnostic features proposed in the DSM-5 (American Psychiatric Association, 2013). The following potential groups of ED outcomes were identified: (a) overall ED symptoms, defined as those scores that derived both from full instruments or aggregated factors cover more than one kind of symptom or both symptoms and diagnostic features; (b) symptoms of specific disorders; and (c) single diagnostic features. From the potential ED outcomes identified (see Appendix B), only the following were present in the retrieved studies: (a) overall ED symptoms, (b) bulimic symptoms (a specific disorder), (c) dietary restraint (a diagnostic feature of anorexia nervosa), and (d) body/eating concerns (a diagnostic feature of anorexia nervosa and bulimia). Given that this final diagnostic feature was found
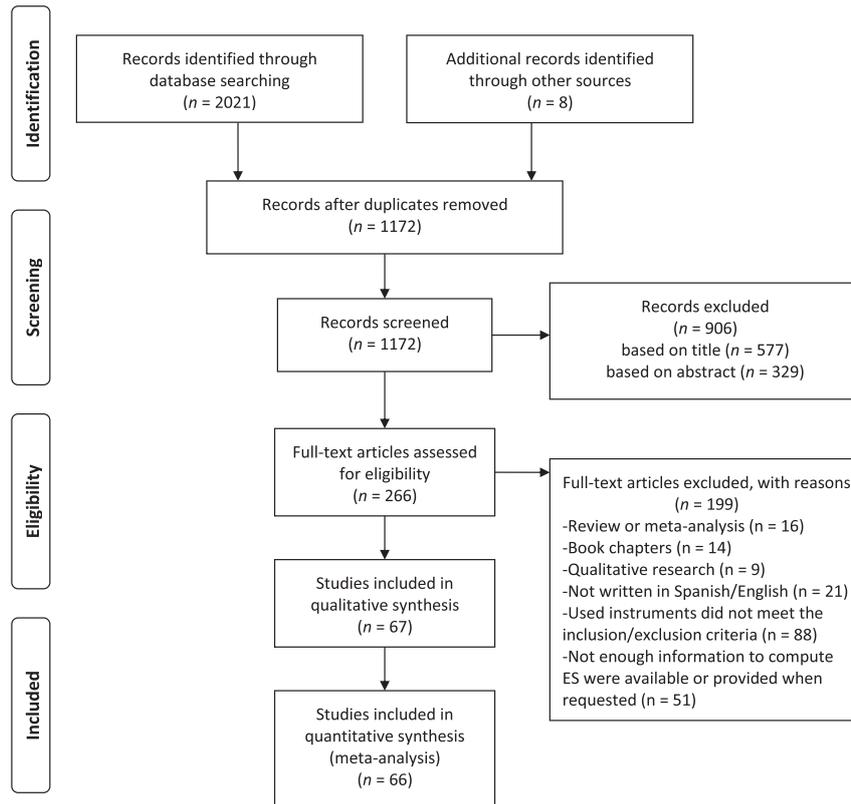
Fig. 1. PRISMA flow diagram of study selection

to be differentially operationalised and termed (i.e., as *drive for thinness* in the *Eating Disorders Inventory* [EDI], and *weight, shape and eating concerns* in the *Eating Disorder Examination Questionnaire*, [EDE-Q]), we opted for defining them as two differentiated outcomes. In the case of the EDE-Q, it was observed that the content alluding to weight and shape was not restricted to their corresponding factor in the EDE-Q. Therefore, the average score of the three factors (weight, shape, and eating) was employed. Consequently, a total of 135 effect-sizes from 66 studies (N = 21,816) were examined in five different meta-analyses (see Tables 1–5). In addition to the above, the relationship between the specified ED outcomes and specific features of MEB was examined through a differentiated set of meta-analyses when enough data were available (i.e., $K > 4$), as it was the case for *Compulsive Exercise Test* (CET) and the *Exercise Dependence Scale-Revised* (EDS-R) (see Table 7).

Overall ED symptoms

The analysis examining the relationship between MEB and overall ED symptoms (see Appendix C) included 60 effect-sizes from 48 studies (N_{total} = 16,421). Findings from the random effects model showed a medium effect size ($r = 0.35$, $P < 0.001$; 95%CI = 0.30 to 0.40). The high heterogeneity observed ($I^2 = 91.37$) suggested the presence of potential moderators. As regards to categorical variables, findings from analogue to ANOVA analyses employing a mixed-effects

model (see Table 6) showed significant differences between groups in three cases: (a) ED assessment ($Q_{between}$ [1] = 12.97; $P < 0.001$), with effect sizes ranging from small in the case of studies considering a categorical assessment of ED ($K = 4$; $r = 0.21$; $P < 0.001$) to medium in the case of studies considering a continuous assessment of ED ($K = 56$; $r = 0.36$; $P < 0.001$); (b) sample type ($Q_{between}$ [1] = 16.46; $P < 0.001$), the effect size being larger for the group comprising clinical samples ($K = 5$; $r = 0.60$; $P < 0.001$) than for the group comprising non-clinical samples ($K = 55$; $r = 0.33$; $P < 0.001$); and (c) MEB measure ($Q_{between}$ [4] = 55.35; $P < 0.001$), with effect sizes ranging from small in the case of the *Exercise addiction inventory* (EAI) ($K = 11$; $r = 0.15$; $P = 0.001$) and the *Compulsive exercise test* (CES) ($K = 6$; $r = 0.28$; $P < 0.001$) to large in the case of the CET ($K = 8$; $r = 0.56$; $P < 0.001$). Conversely, no significant differences between groups were found according to sex ($Q_{between}$ [2] = 2.81; $P = 0.246$), study quality ($Q_{between}$ [2] = 4.54; $P = 0.103$), or publication status ($Q_{between}$ [1] = 0.43; $P = 0.511$). Regarding continuous variables, after removing effect-sizes for which no mean age ($K = 5$) or BMI ($K = 21$) were available, findings from the random model meta-regression analysis (see Appendix F) showed age ($K = 55$; slope = -0.011; $SE = 0.004$; $P = 0.012$) and BMI ($K = 39$; slope = -0.015; $SE = 0.006$; $P = 0.020$) as significant moderators.

Additionally, the relationship between MEB and overall ED symptoms was examined considering the specific features included in the different instruments assessing MEB. This was possible just in the case of features included in the

Table 1. Study's characteristics and effect sizes (overall ED symptoms)

Study	n	Sex	Sample type	Age	BMI	ED measure	ED assessment	MEB measure	Publication status	Study quality	ES (r)
Alcaraz-Ibáñez et al. (2019)	266	Female	Non-Clinical	20.64	21.88	SCOFF	Continuous	EAI	Published	Moderate	0.25
Alcaraz-Ibáñez et al. (2019)	380	Male	Non-Clinical	21.68	23.57	SCOFF	Continuous	EAI	Published	Moderate	0.15
Alexander (2013)	22	Both	Clinical	15.36	13.50	EAT-26	Continuous	OEQ	Unpublished	Moderate	0.67
Becker, 2000b)	250	Female	Non-Clinical	20.07	–	EAT-26	Continuous	OEQ	Unpublished	Moderate	0.44
Bratland-Sanda et al. (2011)	43	Female	Non-Clinical	31.3	25.30	EDE-Q	Continuous	EDS-R	Published	Moderate	−0.07
Bratland-Sanda et al. (2011)	37	Female	Clinical	30.1	20.90	EDE-Q	Continuous	EDS-R	Published	Moderate	0.62
Braun (2009)	218	Female	Non-Clinical	19.28	24.27	EAT-26	Continuous	OEQ	Unpublished	Moderate	0.50
Braun (2009)	219	Male	Non-Clinical	19.28	24.74	EAT-26	Continuous	OEQ	Unpublished	Moderate	0.25
Bureau et al. (2017)	502	Both	Non-Clinical	18.91	–	EAT-26	Continuous	EAI	Published	Moderate	0.35
Compte et al. (2018)	203	Male	Non-Clinical	21.78	26.86	EDE-Q	Categorical	EDS-R	Published	Moderate	0.26
Cook & Hausenblas (2011)	387	Female	Non-Clinical	20.11	23.73	EDDS	Continuous	EDS-R	Published	Moderate	0.37
Cook et al. (2014a)	387	Female	Non-Clinical	20.11		EDDS	Categorical	EDS-R	Published	Moderate	0.24
Cook et al. (2015)	43	Female	Non-Clinical	19.95	21.61	EDDS	Continuous	EDS-R	Published	Moderate	0.56
Costa et al. (2016)	170	Female	Non-Clinical	20.57	22.63	EAT-26	Continuous	EDS-R	Published	Moderate	0.52
Costa et al. (2016)	178	Male	Non-Clinical	20.57	22.63	EAT-26	Continuous	EDS-R	Published	Moderate	0.37
Cunningham et al. (2016)	885	Female	Non-Clinical	–	–	EDE-Q	Continuous	EAI	Published	Strong	0.18
Costa et al. (2016)	608	Male	Non-Clinical	–	–	EDE-Q	Continuous	EAI	Published	Strong	0.09
Diehl et al. (1998)	160	Female	Non-Clinical	21.53	22.22	EAT-40	Continuous	OEQ	Published	Strong	0.28
Di Lodovico et al. (2018)	81	Female	Non-Clinical	28.79	–	SCOFF	Continuous	EAI	Published	Strong	0.17
Di Lodovico et al. (2018)	73	Male	Non-Clinical	30.90	–	SCOFF	Continuous	EAI	Published	Strong	0.18
Formby et al. (2014)	104	Both	Clinical	14.90	–	EDE-Q	Continuous	CET	Published	Moderate	0.68
Fortes et al. (2014)	116	Female	Non-Clinical	14.54	20.43	EAT-26	Continuous	CES	Published	Moderate	0.09
Fortes et al. (2014)	464	Male	Non-Clinical	15.05	21.28	EAT-26	Continuous	CES	Published	Moderate	0.41
Giardino & Procidano (2012)	24	Female	Non-Clinical	20.17	–	EAT-26	Continuous	EDS-R	Published	Weak	0.61
Giardino & Procidano (2012)	11	Female	Non-Clinical	22.18	–	EAT-26	Continuous	EDS-R	Published	Weak	0.35
Giardino & Procidano (2012)	43	Male	Non-Clinical	20.47	–	EAT-26	Continuous	EDS-R	Published	Weak	0.47
Giardino & Procidano (2012)	35	Male	Non-Clinical	23.34	–	EAT-26	Continuous	EDS-R	Published	Weak	0.54

(continued)

Table 1. Continued

Study	n	Sex	Sample type	Age	BMI	ED measure	ED assessment	MEB measure	Publication status	Study quality	ES (r)
Godier (2015)	126	Both	Non-Clinical	26.33	23.73	EDE-Q	Continuous	CET	Unpublished	Strong	0.53
Godier (2015)	78	Both	Clinical	27.18	18.12	EDE-Q	Continuous	CET	Unpublished	Strong	0.59
Hefner et al. (2016)	262	Female	Non-Clinical	20.48	22.63	EAT-26	Continuous	CET	Published	Moderate	0.52
Kessler (2010)	155	Female	Non-Clinical	20.86	23.76	EAT-26	Continuous	EDS-R	Unpublished	Strong	0.49
Lease & Bond (2013)	302	Female	Non-Clinical	22.30	23.20	EAT-26	Continuous	OEQ	Published	Moderate	0.67
Lease et al. (2016)	298	Female	Non-Clinical	22.3	–	EAT-26	Continuous	OEQ	Published	Moderate	0.66
LePage et al. (2012)	51	Female	Non-Clinical	19.06	–	EDDS	Continuous	OEQ	Published	Weak	0.54
LePage et al. (2012)	73	Both	Non-Clinical	19.08	22.46	EDDS/ EAT-26	Continuous	OEQ	Published	Weak	0.57
Lipsey et al. (2006)	260	Female	Non-Clinical	29.40	23.20	EDE-Q	Continuous	CES	Published	Moderate	0.37
Maraz et al. (2015)	447	Female	Non-Clinical	32.80	22.38	SCOFF	Continuous	EAI	Published	Moderate	0.19
Meulemans et al. (2014)	520	Both	Non-Clinical	19.76	–	EAT-26	Continuous	EDS-R	Published	Strong	0.22
Mond et al. (2004)	169	Female	Non-Clinical	33.40	24.30	EDE-Q	Continuous	CES	Published	Strong	0.11
Mond et al. (2006)	3,472	Female	Non-Clinical	29.91	24.27	EDE-Q	Continuous	CES	Published	Strong	0.25
Mond et al. (2008)	177	Female	Non-Clinical	26.91	27.58	EDE-Q	Continuous	CES	Published	Strong	0.38
Müller et al. (2015)	49	Female	Non-Clinical	26.93	21.49	EDE-Q	Continuous	EDS-R	Published	Moderate	0.40
Müller et al. (2015)	79	Male	Non-Clinical	26.51	23.97	EDE-Q	Continuous	EDS-R	Published	Moderate	0.28
Müller et al. (2018)	216	Both	Non-Clinical	44.00	48.30	EDE-Q	Continuous	EDS-R	Published	Moderate	0.19
Naylor et al. (2011)	64	Female	Clinical	29.98	19.23	EDE-Q	Continuous	CET	Published	Strong	0.40
Naylor et al. (2011)	76	Female	Non-Clinical	20.32	20.86	EDE-Q	Continuous	CET	Published	Strong	0.64
Petty (2010)	208	Both	Non-Clinical	24.45	–	EAT-26	Continuous	EAI	Unpublished	Moderate	−0.33
Pritchard et al. (2011)	332	Female	Non-Clinical	–	–	EDE-Q	Continuous	OEQ	Published	Moderate	0.25
Pritchard et al. (2011)	232	Male	Non-Clinical	–	–	EDE-Q	Continuous	OEQ	Published	Moderate	0.27
Prybock (1999)	253	Female	Non-Clinical	19.90	–	EAT-40	Continuous	OEQ	Unpublished	Strong	−0.36
Rocks et al. (2017)	119	Both	Non-Clinical	27.00	22.46	EAT-26	Categorical	EAI	Published	Moderate	0.27
Serier et al. (2018)	35	Female	Non-Clinical	35.94	23.54	EAT-26	Continuous	OEQ	Published	Strong	0.21
Serier et al. (2018)	35	Female	Non-Clinical	32.51	23.03	EAT-26	Continuous	OEQ	Published	Strong	0.46
Stuart et al. (2015)	86	Female	Non-Clinical	–	23.29	EDE-Q	Continuous	EDS-R	Published	Moderate	0.63
Taranis et al. (2011)	101	Female	Non-Clinical	20.90	21.80	EDE-Q	Continuous	OEQ	Published	Moderate	0.38
Tobar et al. (2017)	39	Female	Non-Clinical	46.00	–	EAT-26	Continuous	EDS-R	Unpublished	Weak	0.02
Wischenka (2018)	949	Both		41.95	–	EAT-26	Continuous	OEQ	Unpublished	Moderate	0.30

(continued)

Table 1. Continued

Study	n	Sex	Sample type	Age	BMI	ED measure	ED assessment	MEB measure	Publication status	Study quality	ES (r)
Young et al. (2017)	78	Both	Non-Clinical	27.12	16.49	EDE-Q	Continuous	CET	Published	Strong	0.64
Young et al. (2018)	78	Both	Non-Clinical	27.38	16.52	EDE-Q	Continuous	CET	Published	Moderate	0.41
Zeulner et al. (2016)	1,093	Both	Non-Clinical	41.20	23.30	SCOFF	Categorical	EAI	Published	Strong	0.15

Note: BMI = Body mass index; ED = Eating disorders; MEB = Morbid exercise behaviour; ES = effect size, EDI = Eating Disorders Inventory; EAT-40 = Eating Attitudes Test-40; EAT-26; Eating Attitudes Test-26; EDE-Q = Eating Disorder Examination-Questionnaire; EDDS = Eating Disorder Diagnostic Scale; EAI = Exercise Addiction Inventory; EDS-R = Exercise Dependence Scale Revised; CET = Compulsive Exercise Test; CES = Commitment to Exercise Scale; OEQ = Obligatory Exercise Questionnaire.

CET (avoidance and rule-driven behaviour, weight control exercise, mood improvement, lack of exercise enjoyment, and exercise rigidity) and the EDS-R (withdrawal, intention effects, lack of control, time, reduction in other activities, and continuance). These analyses (see Appendix C) included five effect-sizes from four studies (N_{total} = 1,798) for each of the specific features proposed in the CET, and six effect-sizes from four studies (N_{total} = 2,428) for each of the specific features proposed in the EDS-R. Findings from the random effects models (see Table 7) showed effect sizes that ranged from (a) small (for features such as mood improvement, lack of exercise enjoyments, and exercise rigidity) to large (for weight control exercise) in the case of the CET, and (b) and from trivial (for time devoted to exercise) to small (for the remaining features) in the case of the EDS-R.

Bulimic symptoms

The analysis examining the relationship between MEB and bulimic symptoms (see Appendix C) included 23 effect-sizes from 20 studies (N_{total} = 6,076). Findings from the random effects model showed a small effect size (r = 0.19, P < 0.001; 95%CI = 0.13 to 0.25). The high heterogeneity observed (I^2 = 80.95) suggested the presence of potential moderators. Regarding categorical variables, findings from analogue to ANOVA analyses employing a mixed-effects model (see Table 6) showed significant differences between groups only in the case of study quality ($Q_{between}$ (2) = 15.33; P < 0.001), the effect size being larger for the group comprising moderate-quality studies (K = 15; r = 0.23; P < 0.001) than for the group comprising strong-quality studies (K = 7; r = 0.08; P = 0.180). In this case, the estimate for the weak-quality studies was not interpreted due to the low number of effect-sizes available (K = 1). Conversely, no significant differences between groups were found according to sex ($Q_{between}$ [1] = 0.08; P = 0.781); sample type ($Q_{between}$ [1] = 0.10; P = 0.752); publication status ($Q_{between}$ [1] = 0.23; P = 0.635); and MEB measure ($Q_{between}$ [3] = 4.16; P = 0.245). Due to unavailability of data, no moderator analysis was conducted according to ED assessment. Regarding continuous variables, after removing effect-sizes for which no mean age and BMI were available (K = 8), findings from the

random model meta-regression analysis did not find either age (K = 23; slope = −0.006; SE = 0.005; P = 0.243) or BMI (K = 15; slope = −0.014; SE = 0.023; P = 0.548) as significant moderators.

Dietary restraint

The analysis examining the relationship between MEB and dietary restraint (see Appendix C) included 22 effect-sizes from 18 studies (N_{total} = 6,736). Findings from the random effects model showed a medium effect size (r = 0.42, P < 0.001; 95%CI = 0.34 to 0.49). The high heterogeneity observed (I^2 = 91.88) suggested the presence of potential moderators. Regarding categorical variables, findings from analogue to ANOVA analyses employing a mixed-effects model (see Table 6) showed significant differences between groups in two cases: (a) sex ($Q_{between}$ [2] = 9.61; P = 0.008), the effect size being larger for women (K = 14; r = 0.47; P < 0.001) than for men (K = 4; r = 0.25; P < 0.001); and (b) MEB measure ($Q_{between}$ [4] = 55.35; P < 0.001), the effect size being larger in studies considering the *Obligatory Exercise Questionnaire* (OEQ) (K = 11; r = 0.46; P < 0.001) than in studies considering the EAI (K = 5; r = 0.20; P < 0.001). In this case, estimates for the remaining measures were not interpreted due to the low number of effect-sizes available (EDS-R, K = 2; CET, K = 3, CES, K = 1). Conversely, no significant differences between groups were found according to publication status ($Q_{between}$ [1] = 0.17; P = 0.681) or study quality ($Q_{between}$ [2] = 4.54; P = 0.103). Due to the unavailability of data, no moderator analysis was conducted according to ED assessment and sample type. Regarding continuous variables, after removing effect-sizes for which no mean age (K = 4) or BMI (K = 8) were available, findings from the random model meta-regression analysis did not find either age (K = 18; slope = −0.009; SE = 0.007; P = 0.170) or BMI (K = 13; slope = −0.012; SE = 0.028; P = 0.676) as significant moderators.

Body/eating concerns

The analysis examining the relationship between MEB and body/eating concerns as operationalised in the drive for

Table 2. Study's characteristics and effect sizes (bulimic symptoms)

Study	n	Sex	Sample type	Age	BMI	ED measure	ED assessment	MEB measure	Publication status	Study quality	ES (r)
Alexander (2013)	22	Both	Clinical	15.36	13.50	EAT-26	Continuous	OEQ	Unpublished	Moderate	0.61
Becker (2000b)	250	Female	Non-Clinical	20.07	–	EDI-2	Continuous	CES	Unpublished	Moderate	0.16
Bell et al. (2016)	388	Female	Non-Clinical	21.46	23.00	EAT-26	Continuous	OEQ	Published	Moderate	0.23
Bratland-Sanda et al. (2011)	43	Female	Non-Clinical	31.30	25.30	EDI-2	Continuous	EDS-R	Published	Moderate	0.07
Bratland-Sanda et al. (2011)	37	Female	Clinical	30.10	20.90	EDI-2	Continuous	EDS-R	Published	Moderate	−0.20
Clark (1995)	111	Female	Non-Clinical	19.04	22.90	EDI-2	Continuous	OEQ	Unpublished	Moderate	0.19
Diehl et al. (1998)	160	Female	Non-Clinical	21.53	22.22	BULIT	Continuous	OEQ	Published	Strong	0.12
Formby et al. (2014)	104	Both	Clinical	14.90	–	EDI-3	Continuous	CET	Published	Moderate	0.32
Goodwin et al. (2011)	1,012	Both	Non-Clinical	13.02	–	EDI-2	Continuous	CET	Published	Moderate	0.21
Lease & Bond (2013)	302	Female	Non-Clinical	22.30	23.20	EAT-26	Continuous	OEQ	Published	Moderate	0.48
Martin & Hausenblas (1998)	286	Female	Non-Clinical	34.11	–	EDI-2	Continuous	CES	Published	Moderate	0.18
Mussap (2007)	130	Female	Non-Clinical	25.10	–	EDI-2	Continuous	OEQ	Published	Weak	0.47
Nieman (1994)	250	Both	Non-Clinical	14.61	22.04	EDI-2	Continuous	OEQ	Unpublished	Moderate	0.20
Pini et al. (2007)	50	Both	Non-Clinical	35.40	–	EDI-2	Continuous	CES	Published	Moderate	0.12
Prybock (1999)	253	Female	Non-Clinical	19.90	–	BULIT	Continuous	OEQ	Unpublished	Strong	−0.19
Sauchelli et al. (2016)	157	Both	Clinical	28.88	22.06	EDI-2	Continuous	CET	Published	Strong	0.11
Taranis et al. (2011)	101	Female	Non-Clinical	20.90	21.80	EDI-2	Continuous	CET	Published	Moderate	0.20
Thome (2004)	599	Female	Non-Clinical	20.12	22.00	EAT-26/ EDI-2	Continuous	CES	Unpublished	Moderate	0.32
Uhlmann et al. (2018)	356	Female	Non-Clinical	20.57	22.79	EAT-26	Continuous	OEQ	Published	Strong	0.21
Wischenka (2018)	1,158	Both	Non-Clinical	41.95	–	EAT-26	Continuous	OEQ	Unpublished	Moderate	0.14
Zeeck et al. (2017)	107	Both	Non-Clinical	20.20	21.70	EDI-2	Continuous	CES	Published	Strong	0.18
Zeeck et al. (2017)	100	Both	Clinical	26.10	19.30	EDI-2	Continuous	CES	Published	Strong	0.04
Zeeck et al. (2017)	100	Both	Non-Clinical	23.30	21.80	EDI-2	Continuous	CES	Published	Strong	0.13

Note: BMI = Body mass index; ED = Eating disorders; MEB = Morbid exercise behaviour; ES = Effect size, EDI = Eating Disorders Inventory; EAT-40 = Eating Attitudes Test-40; EAT-26; Eating Attitudes Test-26; EDE-Q = Eating Disorder Examination-Questionnaire; BULIT = Bulimic Investigatory Test; EDS-R = Exercise Dependence Scale Revised; CET = Compulsive Exercise Test; CES = Commitment to Exercise Scale; OEQ = Obligatory Exercise Questionnaire.

thinness factor included in the EDI (see Appendix C) included 21 effect-sizes from 18 studies (N_{total} = 4,818). Findings from the random effects model showed a medium effect size (r = 0.41, $P < 0.001$; 95%CI = 0.35 to 0.46). The high heterogeneity observed (I^2 = 78.95) suggested the presence of potential moderators. Regarding categorical

variables, findings from analogue to ANOVA analyses employing a mixed-effects model (see Table 6) showed significant differences between groups in two cases: (a) study quality ($Q_{between}$ [2] = 8.38; P = 0.015), the effect size being higher for the group comprising moderate-quality studies (K = 15; r = 0.44; $P < 0.001$) than for the group comprising

Table 3. Study characteristics and effect sizes (dietary restraint)

Study	n	Sex	Sample type	Age	BMI	ED measure	ED assessment	MEB measure	Publication status	Study quality	ES (r)
Adams (2013)	260	Male	Non-Clinical	33.98	–	EDE-Q	Continuous	OEQ	Unpublished	Strong	0.39
Alexander (2013)	22	Both	Clinical	15.36	13.50	EAT-26	Continuous	OEQ	Unpublished	Moderate	0.57
Bratland-Sanda et al. (2011)	43	Female	Non-Clinical	31.30	25.30	EDE-Q	Continuous	EDS-R	Published	Moderate	0.15
Bratland-Sanda et al. (2011)	37	Female	Clinical	30.10	20.90	EDE-Q	Continuous	EDS-R	Published	Moderate	0.49
Bell et al. (2016)	388	Female	Non-Clinical	21.46	23.00	DEBQ	Continuous	OEQ	Published	Moderate	0.46
Cunningham et al. (2016)	885	Female	Non-Clinical	–	–	EDE-Q	Continuous	EAI	Published	Strong	0.19
Cunningham et al. (2016)	608	Male	Non-Clinical	–	–	EDE-Q	Continuous	EAI	Published	Strong	0.16
Godier (2015)	126	Both	Non-Clinical	26.33	23.73	EDE-Q	Continuous	CET	Unpublished	Strong	0.49
Lamarche & Gammage (2012)	51	Female	Non-Clinical	19.06		RRS	Continuous	OEQ	Published	Weak	0.57
Lease & Bond (2013)	302	Female	Non-Clinical	22.30	23.2	EAT-26	Continuous	OEQ	Published	Moderate	0.67
LePage et al. (2012)	51	Female	Non-Clinical	19.06	–	RRS	Continuous	OEQ	Published	Weak	0.57
Noetel et al. (2016)	60	Female	Clinical	15.02	–	EDE-Q	Continuous	CET	Published	Strong	0.69
Pritchard et al. (2011)	331	Female	Non-Clinical	–	–	EDE-Q	Continuous	OEQ	Published	Moderate	0.30
Pritchard et al. (2011)	231	Male	Non-Clinical	–	–	EDE-Q	Continuous	OEQ	Published	Moderate	0.27
Rocks et al. (2017)	119	Both	Non-Clinical	27.00	22.46	TFEQ-R18	Categorical	EAI	Published	Moderate	0.26
Sicilia et al. (2019)	280	Female	Non-Clinical	15.48	21.68	DEAS	Continuous	EAI	Unpublished	Moderate	0.29
Sicilia et al. (2019)	338	Male	Non-Clinical	15.47	21.70	DEAS	Continuous	EAI	Unpublished	Moderate	0.19
Taranis et al. (2011)	101	Female	Non-Clinical	20.90	21.80	EDE-Q	Continuous	CET	Published	Moderate	0.49
Thome (2004)	599	Female	Non-Clinical	20.12	22.00	EAT-26	Continuous	CES	Unpublished	Moderate	0.58
Thome & Espelage (2007)	599	Female	Non-Clinical	20.13	23.61	EAT-26	Continuous	OEQ	Published	Moderate	0.56
Uhlmann et al. (2018)	356	Female	Non-Clinical	20.57	22.79	DEBQ	Continuous	OEQ	Published	Strong	0.44
Wischenka (2018)	949	Both	Non-Clinical	41.95	–	EAT-26	Continuous	OEQ	Unpublished	Moderate	0.31

Note: BMI = Body mass index; ED = Eating disorders; MEB = Morbid exercise behaviour; ES = Effect size; EDI = Eating Disorder Inventory; EAT-40 = Eating Attitudes Test-40; EAT-26; Eating Attitudes Test-26; EDE-Q = Eating Disorder Examination-Questionnaire; RRS = Revised Restraint Scale; DEAS = Disordered Eating Attitude Scale; TFEQ = Three Factor Eating Questionnaire; DRES = Dutch Restrained Eating Scale; EDS-R = Exercise Dependence Scale Revised; CET = Compulsive Exercise Test; CES = Commitment to Exercise Scale; OEQ = Obligatory Exercise Questionnaire.

strong-quality studies ($K = 5$; $r = 0.29$; $P < 0.001$). In this case, the estimate for weak-quality studies was not interpreted due to the low number of effect-sizes available ($K = 1$); and (b) MEB measure ($Q_{between}$ [4] = 19.99; $P < 0.001$), with effect sizes ranging from moderate in the case of both the CES ($K = 8$; $r = 0.34$; $P < 0.001$) and OEQ ($K = 6$; $r = 0.45$; $P < 0.001$) to large in the case of CET ($K = 4$; $r = 0.55$; $P < 0.001$). In this case, the estimate for the EDS-R was not interpreted due to the low number of effect-sizes available

($K = 3$). Conversely, no significant differences between groups were found according to sex, ($Q_{between}$ [1] = 0.01; $P = 0.920$), sample type ($Q_{between}$ [1] = 1.19; $P = 0.276$) or publication status ($Q_{between}$ [1] = 0.15; $P = 0.703$). Due to the unavailability of data, no moderator analysis was conducted according to ED assessment. Regarding continuous variables, after removing effect-sizes for which no mean BMIs ($K = 9$) were available, findings from the random model meta-regression analysis did not show either age
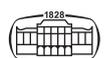
Table 4. Study characteristics and effect sizes (body/eating concerns as operationalised by the EDI)

Study	n	Sex	Sample type	Age	BMI	ED measure	ED assessment	MEB measure	Publication status	Study quality	ES (r)
Alexander (2013)	22	Both	Clinical	15.36	13.50	EDI-3	Continuous	OEQ	Unpublished	Moderate	0.57
Aruguete et al. (2012)	258	Female	Non-clinical	22.45	24.33	EDI	Continuous	CES	Published	Moderate	0.40
Becker (2000b)	250	Female	Non-clinical	20.07	–	EDI-2	Continuous	CES	Unpublished	Moderate	0.40
Bratland-Sanda et al. (2011)	43	Female	Non-clinical	31.30	25.30	EDI-2	Continuous	EDS-R	Published	Moderate	0.11
Bratland-Sanda et al. (2011)	37	Female	Clinical	30.10	20.90	EDI-2	Continuous	EDS-R	Published	Moderate	0.42
Clark (1995)	111	Female	Non-clinical	19.04	22.90	EDI-2	Continuous	OEQ	Unpublished	Moderate	0.42
Cook & Hausenblas (2008)	330	Female	Non-clinical	19.97	–	EDI-2	Continuous	EDS-R	Published	Strong	0.21
Formby et al. (2014)	104	Both	Clinical	14.90	–	EDI-3	Continuous	CET	Published	Moderate	0.70
Goodwin et al. (2011)	1,012	Both	Non-clinical	13.02	–	EDI-2	Continuous	CET	Published	Moderate	0.48
Gulker et al. (2001)	172	Both	Non-clinical	36.00	–	EDI-2	Continuous	OEQ	Published	Moderate	0.46
Martin & Hausenblas (1998)	286	Female	Non-clinical	34.11	–	EDI-2	Continuous	CES	Published	Moderate	0.30
Mussap (2007)	130	Female	Non-clinical	25.10	–	EDI-2	Continuous	OEQ	Published	Weak	0.53
Nieman (1994)	250	Both	Non-clinical	14.61	22.04	EDI-2	Continuous	OEQ	Unpublished	Moderate	0.28
Pini et al. (2007)	50	Both	Non-clinical	35.40	–	EDI-2	Continuous	CES	Published	Moderate	0.20
Sauchelli et al. (2016)	157	Both	Clinical	28.88	22.06	EDI-2	Continuous	CET	Published	Strong	0.49
Taranis et al. (2011)	101	Female	Non-clinical	20.90	21.80	EDI-2	Continuous	CET	Published	Moderate	0.53
Thome (2004)	599	Female	Non-clinical	20.12	22.00	EDI-2	Continuous	CES	Unpublished	Moderate	0.51
Thome & Espelage (2007)	599	Female	Non-clinical	20.13	23.61	EDI-2	Continuous	OEQ	Published	Moderate	0.50
Zeeck et al. (2017)	107	Both	Non-clinical	20.20	21.70	EDI-2	Continuous	CES	Published	Strong	0.23
Zeeck et al. (2017)	100	Both	Clinical	26.10	19.30	EDI-2	Continuous	CES	Published	Strong	0.22
Zeeck et al. (2017)	100	Both	Non-clinical	23.30	21.80	EDI-2	Continuous	CES	Published	Strong	0.26

Note: BMI = Body mass index; ED = Eating disorders; MEB = Morbid exercise behaviour; ES = Effect size; EDI = Eating Disorders Inventory; EDE-Q = Eating Disorder Examination-Questionnaire; EDS-R = Exercise Dependence Scale Revised; CET = Compulsive Exercise Test; OEQ = Obligatory Exercise Questionnaire.

($K = 22$; slope = -0.008; $SE = 0.005$; $P = 0.115$) or BMI ($K = 12$; slope = -0.013; $SE = 0.022$; $P = 0.571$) as significant moderators.

The analysis examining the relationship between MEB and body/eating concerns as conceptualized by the shape, weight, and eating concerns factors of the EDE-Q (see Appendix C) included nine effect-sizes from six studies (N_{total} = 2,424). Findings from the random effects model showed a medium effect size ($r = 0.28$, $P < 0.001$; 95% CI = 0.17–0.38). The high heterogeneity observed ($I^2 = 83.95$) suggested the presence of potential moderators. Due to

unavailability of data, moderator analyses were conducted only for one categorical variable (i.e., study quality). Findings from analogue to ANOVA analyses employing a mixed-effects model (see Table 6) did not show significant differences according to study quality ($Q_{between}$ [1] = 1.46; P =0.228).

Sensitivity analysis and publication bias

Findings from the sensitivity analyses showed that the pooled estimates resulting from the five meta-analyses

Table 5. Study characteristics and effect sizes (body/eating concerns as operationalised by the EDE-Q)

Study	n	Sex	Sample type	Age	BMI	ED measure	ED assessment	MEB measure	Publication status	Study quality	ES (r)
Bratland-Sanda et al. (2011)	43	Female	Non-clinical	31.30	25.30	EDE-Q	Continuous	EDS-R	Published	Moderate	−0.09
Bratland-Sanda et al. (2011)	37	Female	Clinical	30.10	20.90	EDE-Q	Continuous	EDS-R	Published	Moderate	0.43
Cunningham et al. (2016)	885	Female	Non-clinical	–	–	EDE-Q	Continuous	EDS-R	Published	Strong	0.21
Cunningham et al. (2016)	608	Male	Non-clinical	–	–	EDE-Q	Continuous	EDS-R	Published	Strong	0.08
Godier (2015)	126	Both	Non-clinical	26.33	23.73	EDE-Q	Continuous	CET	Unpublished	Strong	0.48
Noetel et al. (2016)	60	Female	Clinical	15.02	–	EDE-Q	Continuous	CET	Published	Strong	0.66
Pritchard et al. (2011)	232	Male	Non-clinical	–	–	EDE-Q	Continuous	OEQ	Published	Moderate	0.21
Pritchard et al. (2011)	329	Female	Non-clinical	–	–	EDE-Q	Continuous	OEQ	Published	Moderate	0.17
Taranis et al. (2011)	101	Female	Non-clinical	20.90	21.80	EDE-Q	Continuous	OEQ	Published	Moderate	0.34

Note: BMI = Body mass index; ED = Eating disorders; MEB = Morbid exercise behaviour; ES = Effect size, EDE-Q = Eating Disorder Examination-Questionnaire; EDS-R = Exercise Dependence Scale Revised; CET = Compulsive Exercise Test; OEQ = Obligatory Exercise Questionnaire.

conducted were not significantly modified when specific data were removed one at a time (see Appendix D). Due to unavailability of data (K < 10) (Page et al., 2019), publication bias analysis was not conducted in the case of body/eating concerns as operationalised in the EDE-Q (shape/weight/ eating concern). Evidence of publication bias was not found by funnel plot symmetry and the results of Egger test for bulimic symptoms (P = 0.842), dietary restraint (P = 0.269), or body/eating concerns in the drive for thinness factor included in the EDI (P = 0.252). Conversely, evidence of publication bias was found by funnel plot asymmetry and the results of Egger test in the case of overall ED symptoms (P = 0.013) (see Appendix E). More specifically, the 'trim and fill' procedure identified 13 potential missing studies in the case of overall ED symptoms, showing differences between the pooled estimate (r = 0.35; P < 0.001; 95% CI = 0.30 to 0.40) and adjusted pooled estimate in terms of publication bias (r = 0.27, P < 0.001; 95% CI = 0.22 to 0.33).

DISCUSSION

This systematic review and meta-analysis provides a synthesis of the evidence linking MEB and ED. Meta-analysis summary effects for 66 studies showed small-sized associations between MEB and bulimic symptoms, these being medium-sized in the case of dietary restraint and overall ED symptoms. Additionally, small- to medium sized associations were found between MEB and diagnostic features such as body/eating concerns depending upon the different operationalisations of this latter construct. The magnitude of the observed relationships was slightly lower than those reported by previous meta-analysis examining other

correlates of ED such as weight teasing (Menzel et al., 2010), body checking and body image avoidance (Walker, White, & Srinivasan, 2018), or self-objectification (Schaefer & Thompson, 2018). On balance, these findings support the positive association between MEB and ED reported by previous reviews (Fietz et al., 2014; Meyer et al., 2011; Starcevic & Khazaal, 2017; Trott et al., 2020), while also pointing to dietary restraint and body/eating concerns as operationalised in the EDI because the ED outcomes more strongly related to MEB.

Moderators of the relationship between MEB and ED

A first finding concerning moderators of the relationship under consideration is that their strength was significantly greater when overall ED symptoms were treated as a continuous rather than a categorical variable. Additionally, the relationship between MEB and overall ED symptoms was also found to be greater in clinical than in non-clinical samples. These findings suggest that individuals clinically diagnosed with an ED may feature increased MEB symptoms. However, the possibility that these findings may be due to methodological artefacts should not be discounted. On the one hand, it is possible that some of the participants belonging to the non-clinical subsamples included in the present analyses may qualify for a clinical diagnosis (Di Lodovico, Dubertret, & Ameller, 2018; Lease & Bond, 2013; Maraz, Urbán, Griffiths, & Demetrovics, 2015). On the other hand, the relationship of interest was examined within clinical samples in only six studies (Alexander, 2013; Bratland-Sanda et al., 2011; Formby, Watson, Hilyard, Martin, & Egan, 2014; Godier, 2015; Naylor, Mountford, & Brown, 2011; Sauchelli et al., 2016), of which four considered the CET for the assessment of MEB (Formby et al., 2014;

Table 6. Results of moderators analyses

| | Overall ED symptoms | | | Bulimic symptoms | | | Dietary restraint | | | Body/Eating concerns | | | | | |
| | | | | | | | | | | Drive for thinness (EDI) | | | Shape, weight and eating concerns (EDE-Q | | |
	K	ES (r)	SE	K	ES (r)	SE	K	ES (r)	SE	K	ES (r)	SE	K	ES (r)	SE
ED assessment															
Categorical	4	**0.21**	0.04	–	–	–	–	–	–	–	–	–	–	–	–
Continuous	56	**0.36**	0.03	–	–	–	–	–	–	–	–	–	–	–	–
Sex															
Male	11	**0.28**	0.04	–	–	–	4	**0.25**	0.06	–	–	–	–	–	–
Female	35	**0.36**	0.04	13	**0.19**	0.06	14	**0.47**	0.06	11	**0.41**	0.05	–	–	–
Both	14	**0.38**	0.07	10	**0.18**	0.03	4	**0.37**	0.07	10	**0.40**	0.06	–	–	–
Sample type															
Clinical	5	**0.60**	0.08	5	0.17	0.10	–	–	–	5	**0.50**	0.13	–	–	–
Non-Clinical	55	**0.33**	0.03	18	**0.20**	0.04	–	–	–	16	**0.39**	0.04	–	–	–
Publication status															
Published	49	**0.36**	0.03	16	**0.21**	0.04	15	**0.43**	0.06	16	**0.40**	0.04	–	–	–
Unpublished	11	**0.30**	0.09	7	**0.17**	0.07	7	**0.40**	0.07	5	**0.42**	0.07	–	–	–
Study quality															
Strong	19	**0.29**	0.05	7	0.08	0.06	6	**0.39**	0.08	5	**0.29**	0.07	4	**0.24**	0.06
Moderate	34	**0.37**	0.04	15	**0.23**	0.03	14	**0.41**	0.06	15	**0.44**	0.04	5	**0.35**	0.11
Weak	7	**0.47**	0.09	1	**0.47**	0.09	2	**0.57**	0.10	1	**0.53**	0.09	–	–	–
MEB measure															
EAI	11	**0.15**	0.05	–	–	–	5	**0.20**	0.02	–	–	–	–	–	–
EDS-R	19	**0.37**	0.04	2	−0.06	0.14	2	0.33	0.19	3	**0.22**	0.06	–	–	–
CET	8	**0.56**	0.05	4	**0.21**	0.03	3	**0.55**	0.09	4	**0.55**	0.07	–	–	–
OEQ	16	**0.39**	0.08	10	**0.23**	0.06	11	**0.46**	0.06	6	**0.45**	0.06	–	–	–
CES	6	**0.28**	0.05	7	**0.19**	0.04	1	**0.58**	0.04	8	**0.34**	0.05	–	–	–

Note: Bolded values indicate statistically significant effects (i.e., $P < 0.05$); K = Number effect sizes; ES = effect sizes; SE = Standard error; ED = Eating disorders; EAI = Exercise Addiction Inventory; EDS-R = Exercise Dependence Scale Revised; CET = Compulsive Exercise Test; OEQ = Obligatory Exercise Questionnaire; CES = Commitment to Exercise Scale.

Table 7. Results summarizing the relationships between specific features of MEB and overall ED symptoms

| | | | | | 95% CI | | | |
Instrument	Factor	K	ES (r)	SE	Lo	Up	P	I^2
CET	Avoidance and rule-driven behaviour	5	0.39	0.08	0.26	0.51	<0.001	87.00
CET	Weight control exercise	5	0.57	0.08	0.44	0.66	<0.001	88.57
CET	Mood improvement	5	0.13	0.03	0.06	0.19	<0.001	35.20
CET	Lack of exercise enjoyment	5	0.18	0.03	0.13	0.23	<0.001	10.88
CET	Exercise rigidity	5	0.20	0.06	0.08	0.06	0.001	79.00
EDS-R	Tolerance	6	0.13	0.05	0.04	0.22	0.004	76.62
EDS-R	Withdrawal	6	0.22	0.04	0.13	0.30	<0.001	73.79
EDS-R	Intention effects	6	0.10	0.04	0.02	0.18	0.020	73.46
EDS-R	Lack of control	6	0.24	0.06	0.24	0.34	<0.001	84.29
EDS-R	Time	6	0.08	0.08	0.01	0.16	0.042	71.06
EDS-R	Reduction in other activities	6	0.23	0.05	0.13	0.34	<0.001	83.11
EDS-R	Continuance	6	0.22	0.03	0.16	0.28	<0.001	53.20

Note: ES = Effect size; SE = Standard Error; Lo = Lower; Up = Upper; EDS-R = Exercise Dependence Scale Revised; CET = Compulsive Exercise Test.

Godier, 2015; Naylor et al., 2011; Sauchelli et al., 2016). This is relevant because, as discussed below, moderator analyses showed the scores derived from the CET as the most strongly correlated with ED outcomes. Consequently, further research on this topic may benefit from employing more accurate screening procedures that allow for a reliable classification of participants in terms of their clinical condition, as well as a wider range of instruments for the assessment of MEB in clinical populations.

A second finding in terms of the potential moderators of the relationship under investigation concerns the instrument employed for assessing MEB. Overall, the smallest and largest effect-sizes tended to be respectively associated with the use of the EAI and the CET, respectively. These findings suggest that the relationship between ED and MEB may be respectively weaker and stronger when this latter construct is conceptualized as either a behavioural addiction or according to the expected maintenance factors for excessive exercise within the ED domain. Additionally, results pointed to exercising for weight and shape reasons to be the single feature of MEB more strongly related to overall ED symptoms. Altogether, these findings suggest that a considerable portion of shared variance between overall ED symptoms and composite scores of MEB (i.e., those derived from full instruments) may not be due to the features reflecting the morbid nature of this behaviour (i.e., the inability to control it or the derived physical or psychological harm) but, conversely, to an exercise reason such as weight control. In support of this possibility, it may be argued that among the different instruments assessing MEB in the present study, exercising for weight control reasons was only considered as a specific feature of MEB in the CET. However, it should be noted that exercise reasons have been proposed as substantive psychological constructs that may eventually play a relevant role on the onset and maintenance of MEB (Egorov & Szabo, 2013). In view of these findings, further research examining the association between ED and MEB could clearly benefit from a previous consensual definition of features involved in this potential disorder and, subsequently, from a unified assessment approach.

A third group of noteworthy findings in terms of moderator effects concerns variables such as sex, age, and BMI. Regarding sex, it should be noted that the relationship between MEB and dietary restraint was found to be greater in women than in men. This finding suggests that, as pointed by previous research (Tiggemann & Williamson, 2000), women may turn to exercise as a mean of weight control to a greater extent than men. However, the possibility also exists that this finding is due to an assessment bias. More specifically, most of the instruments assessing dietary restraint were originally developed considering almost exclusively female populations (Murray et al., 2017). Regarding age and BMI, it should be noted that the relationship between MEB and overall ED symptoms was strengthened in samples comprising younger and thinner individuals (i.e., having lower BMI). These findings suggest that younger and thinner individuals may be at greater risk of (a) adopting exercise as a mean of coping with the range

of generic symptoms underlying ED, and/or (b) developing an ED as a result of the desire to attain the leanness and muscular toned ideal body frequently endorsed by young people overinvolved in exercise for such reasons (Holland & Tiggemann, 2017; Uhlmann, Donovan, Zimmer-Gembeck, Bell, & Ramme, 2018).

A fourth notable finding in terms of moderator effects concerns study quality which showed that the weak relationship found in moderate-quality studies between MEB and bulimic symptoms became negligible in high-quality studies. Relatedly, MEB was more strongly related to diagnostic features of anorexia nervosa such as dietary restraint and body/eating concerns than to bulimic symptoms. Taken together, these latter two findings somewhat question the universal validity of "excessive exercise" (at least, as operationalised in the MEB instruments included in the present review) as a diagnostic feature of bulimia nervosa (American Psychiatric Association, 2013). Another important finding in this sense was that the relationship between MEB and body/eating concerns as assessed by the EDI was weaker in moderate-quality than in strong-quality studies. This finding suggests that the greater relationship found between MEB and body/eating concerns when assessed with the EDI compared to the EDE-Q may be partially explained by study quality. Despite these findings suggesting a somewhat similar pattern of relationships irrespective of the instrument employed for the assessment of body/eating concerns, further research aimed at delineating this relationship by considering alternative assessment approaches of body/eating concerns is warranted. In particular, part of the content included in these instruments may not be entirely aligned with the proposed diagnostic feature (i.e., experiencing an intense fear of gaining weight or of becoming fat). This could be the case, for instance, of content reflecting a behavioural component (e.g., eating in secret) or other emotions beyond fear (e.g., guilt).

Strengths and limitations

The present study has several strengths. Firstly, for the first time, a review has summarized evidence concerning the relationship between symptoms underlying MEB and ED employing meta-analytic techniques. Secondly, an attempt was made to make the present review as comprehensive as possible by including data from a variety of populations from both published and unpublished studies. Thirdly, the review examined a wide range of methodological and demographic moderators, consequently providing insight into factors that may strengthen the MEB-ED relationship. Finally, the review addressed not just generic symptoms underlying ED but, additionally, addressed symptoms of specific pathologies and single diagnostic features included in the DSM-5 (American Psychiatric Association, 2013).

Notwithstanding these strengths, several limitations must be acknowledged. Firstly, while efforts were made to explore a wide range of sources of heterogeneity, limited data were available for some of the examined moderators (e.g., sample type, sex, BMI, and MEB measure). This
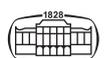
limitation is particularly relevant due to the fact that these variables emerged in the present study as significant moderators of some of the relationships examined. Similarly, demographic information of potential relevance in terms of the correlates of ED such as exercise characteristics (e.g., participating in sports where low body weight or leanness confers a competitive advantage) (Joy, Kussman, & Nattiv, 2016) were rarely reported, thereby preventing this variable from being examined as a possible source of heterogeneity. Secondly, aggregated scores instead of scores on individual features of MEB were mostly provided in the retrieved studies. Therefore, evidence in support of the relationship between specific features of MEB and ED outcomes were derived from a low number of studies. This limitation is particularly important because evidence exists suggesting that instruments assessing MEB may not reflect a unidimensional construct but may be reflecting a multidimensional construct (Formby et al., 2014; Sicilia & González-Cutre, 2011). This becomes more relevant in view of the results suggesting a differentiated pattern of relationships depending upon specific components of MEB. In the absence of a consensual definition on both the features involved in MEB (Berczik et al., 2012; Egorov & Szabo, 2013), further research on this topic should consider reporting not just aggregated scores of MEB but also scores for specific components of MEB. Thirdly, despite the results of the publication bias analyses showing that the findings are largely robust to publication bias, the slightly lower adjusted estimate found in the case of overall ED symptoms suggests that some studies may have been missed. This could be due to the existence of unpublished negative results but also due to the eligibility criteria restricting inclusion of studies written in the languages other than English or Spanish (the languages spoken by the research team). Finally, despite the fact there were no *a priori* restrictions considered in terms of research design, only one of the 67 studies included in the systematic review featured a longitudinal design. This fact precludes from inferring any temporal or causal relationship between MEB and ED based solely on findings from the present meta-analysis. Apart from advising a cautious interpretation of the results, this limitation emphasizes the need for further research that, by employing longitudinal designs, may clarify the direction of the relationships identified in the present study.

Clinical implications

The results from the present study underscore the importance of assessing MEB symptoms in individuals at-risk or clinically diagnosed with an ED, as well as symptoms underlying ED among individuals at greater risk of MEB. The relationships found between both dietary restrain and body/eating concerns and MEB (more specifically, as assessed by the CET) suggest that exercising for thinness-oriented weight and shape reasons may be a key factor in the onset and maintenance of ED. For clinicians and exercise practitioners, this means being cognisant to individuals featuring such reasons for exercising as a means of preventing the

onset and maintenance of ED. Additionally, clinicians and exercise practitioners should provide psychoeducational guidance about either the potential role of MEB in maintaining ED or the dangers derived from exercising as a compensatory behaviour for weight control. These findings also raised the need of exercise programs targeted to clinical populations in terms of ED of being effectively supervised in order to avoid the onset of a morbid form of exercise that may eventually exacerbate the health-related outcomes derived from the primary eating pathology.

CONCLUSIONS AND FUTURE DIRECTIONS

In the present meta-analysis, we reviewed and quantitatively summarised the past three decades of scientific inquiry examining the relationship between MEB and ED. Despite the number of potential demographic and methodological characteristics that remain unexplored, our findings provide preliminary support for the positive associations between MEB and (a) overall ED symptoms, (b) dietary restraint, and (c) body/eating concerns, as being small to moderate sized. The high heterogeneity observed suggests that these relationships may be strengthened depending upon the measure used and/or the population considered. This is particularly relevant in two cases: (a) for overall ED symptoms, when employing a continuously-scored instrument for its assessment or the CET for assessing MEB, as well as in clinical population and younger and thinner individuals; and (b) in the case of dietary restraint, among female population. Since the summarized evidence is entirely derived from cross-sectional data, further longitudinal research should clarify the true nature of the relationship between specific components of MEB and symptoms underlying specific ED over time – more specifically, by considering a range of non-clinical populations (e.g., older age-groups or exercisers practicing different modalities) and clinical populations (e.g., segmented according to the diagnosed disorder). As more data become available, future meta-analyses may benefit from examining the relationship between MEB and ED outcomes not included in the present study such as symptoms of pica, rumination disorder, anorexia nervosa, or binge-eating disorder.

draughting of the final version of the manuscript, including critical revisions for intellectual content.

SUPPLEMENTARY MATERIAL

The online version of this article offers supplementary material (https://doi.org/10.1556/2006.2020.00027).

REFERENCES

Adams, M. (2013). Through the rainbow looking glass: Exploration of the impact of gay community and media exposure on body image, depression and controlling physical appearance in gay men. (Doctoral dissertation). University of Tasmania, Australia.

Alcaraz-Ibáñez, M., Sicilia, Á., Dumitru, D. C., Paterna, A., & Griffiths, M. D. (2019). Examining the relationship between finess-related self-conscious emotions, disordered eating symptoms, and morbid exercise behavior: An exploratory study. *Journal of Behavioral Addictions*, 8, 603–611. https://doi.org/10.1556/2006.8.2019.43.

Alexander, D. W. (2013). Diagnostic and psychosocial predictors of excessive exercise use among adolescents with symptoms of anorexia or bulimia. (Doctoral dissertation). University of North Carolina, USA.

American Psychiatric Association. (2013). Diagnostic and statistical manual of mental disorders. Arlington, VA: American Psychiatric Association.

Aruguete, M. S., Edman, J. L., & Yates, A. (2012). The relationship between anger and other correlates of eating disorders in women. *North American Journal of Psychology*, 14, 139–148.

Becker, B. J. (2000a). Multivariate meta-analysis. In H. E. A. Tinsley & S. Brown (Eds.), Handbook of applied multivariate statistics and mathematical modeling (pp. 499–525). San Diego: Academic Press.

Becker, D. (2000b). An examination of exercise dependence and its relation to eating pathology. (Doctoral dissertation). University of Windsor, Canada.

Béland, M., Lavoie, K. L., Briand, S., White, U. J., Gemme, C., & Bacon, S. L. (2019). Aerobic exercise alleviates depressive symptoms in patients with a major non-communicable chronic disease: A systematic review and meta-analysis. *British Journal of Sports Medicine*, 54, 272–278. https://doi.org/10.1136/bjsports-2018-099360.

Bell, H. S., Donovan, C. L., & Ramme, R. (2016). Is athletic really ideal? An examination of the mediating role of body dissatisfaction in

predicting disordered eating and compulsive exercise. *Eating Behaviors*, 21, 24–29. https://doi.org/10.1016/j.eatbeh.2015.12.012.

Berczik, K., Szabo, A., Griffiths, M. D., Kurimay, T., Kun, B., Urbán, R., et al. (2012). Exercise addiction: Symptoms, diagnosis, epidemiology, and etiology. *Substance Use & Misuse*, 47, 403–417. https://doi.org/10.3109/10826084.2011.639120.

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2005). Comprehensive meta analysis version 2.0. Engelwood, NJ: Biostat.

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). Introduction to meta-analysis. Chichester: John Wiley & Sons. https://doi.org/10.1002/9780470743386.

Bratland-Sanda, S., Martinsen, E. W., Rosenvinge, J. H., Rø, Ø., Hoffart, A., & Sundgot-Borgen, J. (2011). Exercise dependence score in patients with longstanding eating disorders and controls: The importance of affect regulation and physical activity intensity. *European Eating Disorders Review*, 19, 249–255. https://doi.org/10.1002/erv.971.

Braun, S. J. (2009). Exercise can be bad for your health? Models of obligatory exercise in males and females. (Doctoral dissertation). Southern Illinois University- Carbondale, USA.

Bureau, A. T., Razon, S., Saville, B. K., Tokac, U., & Judge, L. W. (2017). Passion for academics and problematic health behaviors. *International Journal of Exercise Science*, 10, 417–433. https://doi.org/10.4171/IFB/94.

Cheung, M. W. L. (2014). Modeling dependent effect sizes with three-level meta-analyses: A structural equation modeling approach. *Psychological Methods*, 19, 211–229. https://doi.org/10.1037/a0032968.

Clark, J. (1995). Obligatory exercise and its clinical correlates. (Doctoral dissertation). University of Florida, USA.

Cohen, J. (1988). Statistical power analysis for the behavioral sciences (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum.

Compte, E. J., Murray, S. B., Sepúlveda, A. R., Schweiger, S., Bressan, M., & Torrente, F. (2018). What position do you play? Eating disorder pathology among rugby players, and the understudied role of player position. *International Journal of Eating Disorders*, 51, 1015–1019. https://doi.org/10.1002/eat.22933.

Cook, B. J., Engel, S., Crosby, R., Hausenblas, H. A., Wonderlich, S. A., & Mitchell, J. (2014a). Pathological motivations for exercise and eating disorder specific health-related quality of life. *International Journal of Eating Disorders*, 47, 268–272. https://doi.org/10.1002/eat.22198.

Cook, B. J., & Hausenblas, H. A. (2008). The role of exercise dependence for the relationship between exercise behavior and eating pathology: Mediator or moderator? *Journal of Health Psychology*, 13, 495–502. https://doi.org/10.1177/1359105308088520.

Cook, B. J., & Hausenblas, H. A. (2011). Eating disorder-specific health-related quality of life and exercise in college females. *Quality of Life Research*, 20, 1385–1390. https://doi.org/10.1007/s11136-011-9879-6.

Cook, B. J., Hausenblas, H. A., Crosby, R. D., Cao, L., & Wonderlich, S. A. (2015). Exercise dependence as a mediator of the exercise and eating disorders relationship: A pilot study. *Eating Behaviors*, 16, 9–12. https://doi.org/10.1016/j.eatbeh.2014.10.012.

Cook, B. J., Hausenblas, H. A., & Freimuth, M. (2014b). Exercise addiction and compulsive exercising: Relationship to eating

disorders, substance use disorders, and addictive disorders. In T. Brewerton & A. Baker Dennis (Eds.), Eating disorders, addictions and substance use disorders: Research, clinical and treatment perspectives (pp. 127–144). Berlin: Springer-Verlag Berlin Heidelberg. https://doi.org/10.1007/978-3-642-45378-6.

Costa, S., Hausenblas, H. A., Oliva, P., Cuzzocrea, F., & Larcan, R. (2016). Maladaptive perfectionism as mediator among psychological control, eating disorders, and exercise dependence symptoms in habitual exerciser. Journal of Behavioral Addictions, 5, 77–89. https://doi.org/10.1556/2006.5.2016.004.

Cunningham, H. E., Pearman, S., & Brewerton, T. D. (2016). Conceptualizing primary and secondary pathological exercise using available measures of excessive exercise. International Journal of Eating Disorders, 49, 778–792. https://doi.org/10.1002/eat.22551.

Di Lodovico, L., Dubertret, C., & Ameller, A. (2018). Vulnerability to exercise addiction, socio-demographic, behavioral and psychological characteristics of runners at risk for eating disorders. Comprehensive Psychiatry, 81, 48–52. https://doi.org/10.1016/j.comppsych.2017.11.006.

Diehl, N. S., Johnson, C. E., Rogers, R. L., & Petrie, T. A. (1998). Social physique anxiety and disordered eating: What's the connection? Addictive Behaviors, 23, 1–6. https://doi.org/10.1016/S0306-4603(97)00003-8.

Duval, S., & Tweedie, R. (2000a). A nonparametric "trim and fill" method of accounting for publication bias in meta-analysis. Journal of the American Statistical Association, 95, 89–98. https://doi.org/10.1080/01621459.2000.10473905.

Duval, S., & Tweedie, R. (2000b). Trim and fill: A simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. Biometrics, 56, 455–463. https://doi.org/10.1111/j.0006-341X.2000.00455.x.

Egorov, A. Y., & Szabo, A. (2013). The exercise paradox: An interactional model for a clearer conceptualization of exercise addiction. Journal of Behavioral Addictions, 2, 199–208. https://doi.org/10.1556/JBA.2.2013.4.2.

Fietz, M., Touyz, S., & Hay, P. J. (2014). A risk profile of compulsive exercise in adolescents with an eating disorder: A systematic review. Advances in Eating Disorders, 2, 241–263. https://doi.org/10.1080/21662630.2014.894470.

Formby, P., Watson, H. J., Hilyard, A., Martin, K., & Egan, S. J. (2014). Psychometric properties of the compulsive exercise test in an adolescent eating disorder population. Eating Behaviors, 15, 555–557. https://doi.org/10.1016/j.eatbeh.2014.08.013.

Fortes, L. de S., De Sousa Almeida, S., & Ferreira, M. E. C. (2014). Influence of psychological, anthropometric and sociodemographic factors on the symptoms of eating disorders in young athletes. Paideia, 24, 21–28. https://doi.org/10.1590/1982-43272457201404.

Freelon, D. (2013) ReCal OIR: Ordinal, interval, and ratio intercoder reliability as a web service. International Journal of Internet Science, 8, 10-16.

Fu, R., Gartlehner, G., Grant, M., Shamliyan, T., Sedrakyan, A., Wilt, T. J., et al. (2011). Conducting quantitative synthesis when comparing medical interventions: AHRQ and the effective health care program. Journal of Clinical Epidemiology, 64, 1187–1197. https://doi.org/10.1016/j.jclinepi.2010.08.010.

Gianini, L. M., Klein, D. A., Call, C., Mayer, L., Foltin, R. W., Walsh, B. T., et al. (2016). The reinforcing effect of exercise in anorexia nervosa: Clinical correlates and relationship to outcome. Eating Disorders, 24, 412–423. https://doi.org/10.1080/10640266.2016.1198204.

Giardino, J. C., & Procidano, M. E. (2012). Muscle dysmorphia symptomatology: A cross-cultural study in Mexico and the United States. International Journal of Men's Health, 11, 83–103. https://doi.org/10.3149/jmh.1101.83.

Godier, L. R. (2015). Compulsivity in anorexia nervosa. (Doctoral dissertation). University of Oxford, UK.

Goodwin, H., Haycraft, E., Taranis, L., & Meyer, C. (2011). Psychometric evaluation of the Compulsive Exercise Test (CET) in an adolescent population: Links with eating psychopathology. European Eating Disorders Review, 19, 269–279. https://doi.org/10.1002/erv.1109.

Gulker, M. G., Laskis, T. A., & Kuba, S. A. (2001). Do excessive exercisers have a higher rate of obsessive-compulsive symptomatology? Psychology Health & Medicine, 6, 387–398. https://doi.org/10.1080/13548500126535.

Hedges, L. V. (2009). Statistical considerations. In H. Cooper, L. V. Hedges, & J. C. Valentine (Eds.), The handbook research of synthesis and meta-analysis (pp. 37–47). New York: Russell Sage Foundation.

Hefner, V., Dorros, S. M., Jourdain, N., Liu, C., Tortomasi, A., Greene, M. P., et al. (2016). Mobile exercising and tweeting the pounds away: The use of digital applications and microblogging and their association with disordered eating and compulsive exercise. Cogent Social Sciences, 2, 1176304. https://doi.org/10.1080/23311886.2016.1176304.

Higgins, J. P. T., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring inconsistency in meta-analyses testing for heterogeneity. BMJ, 327, 557–560. https://doi.org/10.1136/bmj.327.7414.557.

Holland, G., & Tiggemann, M. (2017). "Strong beats skinny every time": Disordered eating and compulsive exercise in women who post fitspiration on Instagram. International Journal of Eating Disorders, 50, 76–79. https://doi.org/10.1002/eat.22559.

Joy, E., Kussman, A., & Nattiv, A. (2016). 2016 update on eating disorders in athletes: A comprehensive narrative review with a focus on clinical assessment and management. British Journal of Sports Medicine, 50, 154–162. https://doi.org/10.1136/bjsports-2015-095735.

Kessler, K. L. (2010). Self-objectification, body image, eating behaviors, and exercise dependence among college females. (Doctoral dissertation). University of North Texas, USA.

Lamarche, L., & Gammage, K. L. (2012). Predicting exercise and eating behaviors from appearance evaluation and two types of investment. Sport, Exercise, and Performance Psychology, 1, 145–157. https://doi.org/10.1109/RADIOELEK.2007.371672.

Lease, H. J., & Bond, M. J. (2013). Correspondence between alternate measures of maladaptive exercise, and their associations with disordered eating symptomatology. Journal of Behavioral Addictions, 2, 153–159. https://doi.org/10.1556/JBA.2.2013.012.

Lease, H. J., Doley, J. R., & Bond, M. J. (2016). My mother told me: The roles of maternal messages, body image, and disordered eating in maladaptive exercise. Eating and Weight Disorders, 21, 469–476. https://doi.org/10.1007/s40519-015-0238-4.

LePage, M. L., Price, M., O'Neil, P., & Crowther, J. H. (2012). The effect of exercise absence on affect and body dissatisfaction as moderated by obligatory exercise beliefs and eating disordered beliefs and behaviors. *Psychology of Sport and Exercise*, *13*, 500–508. https://doi.org/10.1016/j.psychsport.2012.02.008.

Lin, X., Zhang, X., Guo, J., Roberts, C. K., McKenzie, S., Wu, W. C., et al. (2015). Effects of exercise training on cardiorespiratory fitness and biomarkers of cardiometabolic health: A systematic review and meta-analysis of randomized controlled trials. *Journal of the American Heart Association*, *4*(7), e002014. https://doi.org/10.1161/JAHA.115.002014.

Lipsey, Z., Barton, S. B., Hulley, A., & Hill, A. J. (2006). "After a workout." Beliefs about exercise, eating and appearance in female exercisers with and without eating disorder features. *Psychology of Sport and Exercise*, *7*, 425–436. https://doi.org/10.1016/j.psychsport.2006.01.005.

Maraz, A., Urbán, R., Griffiths, M. D., & Demetrovics, Z. (2015). An empirical investigation of dance addiction. *PLoS One*, *10*(5), e0125988. https://doi.org/10.1371/journal.pone.0125988.

Martin, K. A., & Hausenblas, H. A. (1998). Psychological commitment to exercise and eating disorder symptomatology among female aerobic instructors. *The Sport Psychologist*, *12*, 180–190. https://doi.org/10.1080/09593330802505029.

Menzel, J. E., Schaefer, L. M., Burke, N. L., Mayhew, L. L., Brannick, M. T., & Thompson, J. K. (2010). Appearance-related teasing, body dissatisfaction, and disordered eating: A meta-analysis. *Body Image*, *7*, 261–270. https://doi.org/10.1016/j.bodyim.2010.05.004.

Meulemans, S., Pribis, P., Grajales, T., & Krivak, G. (2014). Gender differences in exercise dependence and eating disorders in young adults: A path analysis of a conceptual model. *Nutrients*, *6*, 4895–4905. https://doi.org/10.3390/nu6114895.

Meyer, C., Taranis, L., Goodwin, H., & Haycraft, E. (2011). Compulsive exercise and eating disorders. *European Eating Disorders Review*, *19*, 174–189. https://doi.org/10.1002/erv.1122.

Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS Medicine*, *6*(7), e1000097. https://doi.org/10.1371/journal.pmed.1000097.

Mond, J. M., Cook Myers, T., Crosby, R., Hay, P. J., & Mitchell, J. (2008). 'Excessive Exercise' and eating-disordered behaviour in young adult women: Further evidence from a primary care sample. *European Eating Disorders Review*, *16*, 215–221. https://doi.org/10.1002/erv.

Mond, J. M., Hay, P. J., Rodgers, B., & Owen, C. (2006). An update on the definition of "Excessive Exercise" in eating disorders research. *International Journal of Eating Disorders*, *39*, 147–153. https://doi.org/10.1002/eat.20214.

Mond, J. M., Hay, P. J., Rodgers, B., Owen, C., & Beumont, P. J. V. (2004). Relationships between exercise behaviour, eating-disordered behaviour and quality of life in a community sample of women: When is exercise "excessive"? *European Eating Disorders Review*, *12*, 265–272. https://doi.org/10.1002/erv.579.

Mueller, M., D'Addario, M., Egger, M., Cevallos, M., Dekkers, O. M., Mugglin, C., et al. (2018). Methods to systematically review and meta-analyse observational studies: A systematic scoping review of recommendations. *BMC Medical Research Methodology*, *18*, 44. https://doi.org/10.1186/s12874-018-0495-9.

Müller, A., Leukefeld, C., Hase, C., Gruner-Labitzke, K., Mall, J. W., Köhler, H., et al. (2018). Food addiction and other addictive behaviours in bariatric surgery candidates. *European Eating Disorders Review*, *26*, 585–596. https://doi.org/10.1002/erv.2629.

Müller, A., Loeber, S., Söchtig, J., Te Wildt, B., & De Zwaan, M. (2015). Risk for exercise dependence, eating disorder pathology, alcohol use disorder and addictive behaviors among clients of fitness centers. *Journal of Behavioral Addictions*, *4*, 273–280. https://doi.org/10.1556/2006.4.2015.044.

Murray, S. B., Nagata, J. M., Griffiths, S., Calzo, J. P., Brown, T. A., Mitchison, D., et al. (2017). The enigma of male eating disorders: A critical review and synthesis. *Clinical Psychology Review*, *57*, 1–11. https://doi.org/10.1016/j.cpr.2017.08.001.

Mussap, A. J. (2007). Motivational processes associated with unhealthy body change attitudes and behaviours. *Eating Behaviors*, *8*, 423–428. https://doi.org/10.1016/j.eatbeh.2006.12.001.

Naylor, H., Mountford, V., & Brown, G. (2011). Beliefs about excessive exercise in eating disorders: The role of obsessions and compulsions. *European Eating Disorders Review*, *19*, 226–236. https://doi.org/10.1002/erv.1110.

Nieman, B. (1994). Relation between obligatory exercise and eating disorders. (Doctoral dissertation). University of Cincinnati, USA.

Noetel, M., Miskovic-Wheatley, J., Crosby, R. D., Hay, P. J., Madden, S., & Touyz, S. (2016). A clinical profile of compulsive exercise in adolescent inpatients with anorexia nervosa. *Journal of Eating Disorders*, *4*, 1. https://doi.org/10.1186/s40337-016-0090-6.

Page, M. J., Higgins, J. P. T., & Sterne, J. A. C. (2019). Assessing risk of bias due to missing results in a synthesis. In J. Higgins, J. Thomas, J. Chandler, M. Cumpston, T. Li, M. Page, et al. (Eds.), *Cochrane handbook for systematic reviews of interventions*. London: Cochrane.

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

Petty, K. N. (2010). Implicit theories, dissatisfaction, and risk behavior: A new theoretical approach to body image. (Doctoral dissertation). University of Houston, USA.

Pigott, T. D. (2012). *Advances in meta-analysis*. New York: Springer. https://doi.org/10.1007/978-1-4614-2278-5.

Pini, M., Calamari, E., Puleggio, A., & Pullerà, M. (2007). Exercise commitment and proneness to eating disorders in a group of physical education teachers. *Perceptual & Motor Skills*, *105*, 639–645.

Pritchard, M. E., Parker, C., & Nielsen, A. (2011). What predicts drive for muscularity in college students? *Eating Behaviors*, *12*, 228–231. https://doi.org/10.1016/j.eatbeh.2011.04.002.

Prybock, D. J. (1999). Body dissatisfaction as a predictor of eating, exercise, and anabolic-androgenic steroid use among males. (Doctoral dissertation). Arizona State University, USA.

Rocks, T., Pelly, F., Slater, G., & Martin, L. A. (2017). Prevalence of exercise addiction symptomology and disordered eating in australian students studying Nutrition and Dietetics. *Journal of the Academy of Nutrition and Dietetics*, *117*, 1628–1636. https://doi.org/10.1016/j.jand.2017.04.001.

Rosenthal, R. (1995). Writing meta-analytic reviews. *Psychological Bulletin*, 118, 183–192. https://doi.org/10.1037/0033-2909.118.2.183.

Sauchelli, S., Arcelus, J., Granero, R., Jiménez-Murcia, S., Agüera, Z., Del Pino-Gutiérrez, A., et al. (2016). Dimensions of compulsive exercise across eating disorder diagnostic subtypes and the validation of the Spanish Version of the Compulsive Exercise Test. *Frontiers in Psychology*, 7, 1852. https://doi.org/10.3389/fpsyg.2016.01852.

Schaefer, L. M., & Thompson, J. K. (2018). Self-objectification and disordered eating: A meta-analysis. *International Journal of Eating Disorders*, 51, 483–502. https://doi.org/10.1002/eat.22854.

Serier, K. N., Smith, J. E., Lash, D. N., Gianini, L. M., Harriger, J. A., Sarafin, R. E., et al. (2018). Obligatory exercise and coping in treatment-seeking women with poor body image. *Eating and Weight Disorders*, 23, 331–338. https://doi.org/10.1007/s40519-018-0504-3.

Sicilia, Á., Alcaraz-Ibáñez, M., Cren Chiminazzo, J. G., & Teixeira Fernandes, P. (2019). Latent profile analysis of exercise addiction symptoms in Brazilian adolescents: Association with health-related variables. *Journal of Affective Disorders*. Submitted for publication.

Sicilia, Á., & González-Cutre, D. (2011). Dependence and physical exercise: Spanish validation of the exercise dependence scale-revised (EDS-R). *Spanish Journal of Psychology*, 14, 421–431. https://doi.org/10.5209/rev_SJOP.2011.v14.n1.38.

Starcevic, V. (2016). Behavioural addictions: A challenge for psychopathology and psychiatric nosology. *Australian and New Zealand Journal of Psychiatry*, 50, 721–725. https://doi.org/10.1177/0004867416654009.

Starcevic, V., & Khazaal, Y. (2017). Relationships between behavioural addictions and psychiatric disorders: What is known and what is yet to be learned? *Frontiers in Psychiatry*, 8, 53. https://doi.org/10.3389/fpsyt.2017.00053.

Stuart, A. R., Trepp, B. F., Kuhle, R., Slaugh, B., & Murtaugh, M. A. (2015). Eating disorders, exercise dependence and body image dissatisfaction in female age group ironman triathletes. *Journal of Food & Nutritional Disorders*, 4, 4. https://doi.org/10.4172/2324-9323.1000176.

Szabo, A., Demetrovics, Z., & Griffiths, M. D. (2018). Morbid exercise behavior: Addiction or psychological escape? In H. Budde & M. Wegner (Eds.), The exercise effect on mental health: Neurobiological mechanisms (pp. 277–311). New York: Routledge.

Taranis, L., Touyz, S., & Meyer, C. (2011). Disordered eating and exercise: Development and preliminary validation of the compulsive exercise test (CET). *European Eating Disorders Review*, 19, 256–268. https://doi.org/10.1002/erv.1108.

Thomas, B. H., Ciliska, D., Dobbins, M., & Micucci, S. (2004). A process for systematically reviewing the literature: Providing the research evidence for public health nursing interventions. *Worldviews on Evidence-Based Nursing*, 1, 176–184. https://doi.org/10.1111/j.1524-475X.2004.04006.x.

Thome, J. L. (2004). Obligatory exercise and eating pathology in college females: Replication and development of a structural model. (Doctoral dissertation). University of Illinois at Urbana-Champaign, USA.

Thome, J. L., & Espelage, D. L. (2007). Obligatory exercise and eating pathology in college females: Replication and development of a structural model. *Eating Behaviors*, 8, 334–349. https://doi.org/10.1016/j.eatbeh.2006.11.009.

Tiggemann, M., & Williamson, S. (2000). The effect of exercise on body satisfaction and self-esteem as a function of gender and age. *Sex Roles*, 43, 119–127.

Tobar, D. A., Holcomb, B. C., & Berger, B. G. (2017). Exercise dependence, affect, and preferred weight of female exercisers. *Medicine & Science in Sports & Exercise*, 49(5S), 552. https://doi.org/10.1249/01.mss.0000518431.72438.7e.

Trott, M., Jackson, S. E., Firth, J., Jacob, L., Grabovac, I., Mistry, A., et al. (2020). A comparative meta-analysis of the prevalence of exercise addiction in adults with and without indicated eating disorders. Eating and Weight Disorders-Studies on Anorexia, Bulimia and Obesity. Advance online publication. https://doi.org/10.1007/s40519-019-00842-1.

Uhlmann, L. R., Donovan, C. L., Zimmer-Gembeck, M. J., Bell, H. S., & Ramme, R. A. (2018). The fit beauty ideal: A healthy alternative to thinness or a wolf in sheep's clothing? *Body Image*, 25, 23–30. https://doi.org/10.1016/j.bodyim.2018.01.005.

Walker, C. D., White, E. K., & Srinivasan, V. J. (2018). A meta-analysis of the relationships between body checking, body image avoidance, body image dissatisfaction, mood, and disordered eating. *International Journal of Eating Disorders*, 51, 745–770. https://doi.org/10.1002/eat.22867.

Wischenka, D. M. (2018). Experiential avoidance and disordered eating pathology. (Doctoral dissertation). Yeshiva University, USA.

Young, S., Touyz, S., Meyer, C., Arcelus, J., Rhodes, P., Madden, S., et al. (2017). Validity of exercise measures in adults with anorexia nervosa: The EDE, Compulsive Exercise Test and other self-report scales. *International Journal of Eating Disorders*, 50, 533–541. https://doi.org/10.1002/eat.22633.

Young, S., Touyz, S., Meyer, C., Arcelus, J., Rhodes, P., Madden, S., et al. (2018). Relationships between compulsive exercise, quality of life, psychological distress and motivation to change in adults with anorexia nervosa. *Journal of Eating Disorders*, 6(2), 1–8. https://doi.org/10.1186/s40337-018-0188-0.

Zeeck, A., Schlegel, S., Giel, K. E., Junne, F., Kopp, C., Joos, A., et al. (2017). Validation of the German version of the commitment to exercise scale. *Psychopathology*, 50, 146–156. https://doi.org/10.1159/000455929.

Zeulner, B., Ziemainz, H., Beyer, C., Hammon, M., & Janka, R. (2016). Disordered eating and exercise dependence in endurance athletes. *Advances in Physical Education*, 6, 76–87. https://doi.org/10.4236/ape.2016.62009.