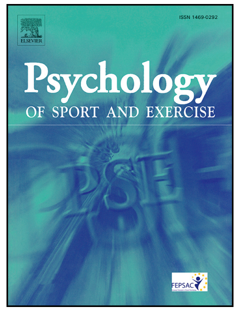
Journal Pre-proof



Gender-related differences in self-reported problematic exercise symptoms: A systematic review and meta-analysis

Manuel Alcaraz-Ibáñez, Adrian Paterna, Mark D. Griffiths, Zsolt Demetrovics, Álvaro Sicilia

PII: S1469-0292(22)00148-0

DOI: https://doi.org/10.1016/j.psychsport.2022.102280

Reference: PSYSPO 102280

To appear in: *Psychology of Sport & Exercise*

Received Date: 9 April 2022

Revised Date: 2 August 2022

Accepted Date: 23 August 2022

Please cite this article as: Alcaraz-Ibáñez, M., Paterna, A., Griffiths, M.D., Demetrovics, Z., Sicilia, Á., Gender-related differences in self-reported problematic exercise symptoms: A systematic review and meta-analysis, *Psychology of Sport & Exercise* (2022), doi: https://doi.org/10.1016/j.psychsport.2022.102280.

Gender-related differences in self-reported symptoms of problematic exercise: A systematic review and meta-analysis

Manuel Alcaraz-Ibáñez[a†], Adrian Paterna[a†], Mark D. Griffiths[b], Zsolt Demetrovics[c, d] Álvaro Sicilia[a]

[a] *Health Research Centre and Department of Education, University of Almería, Spain;*

[b]*Psychology Department, Nottingham Trent University, UK*

[c]*Centre of Excellence in Responsible Gaming, University of Gibraltar, Gibraltar, Gibraltar*

[d]*Institute of Psychology, ELTE Eötvös Loránd University, Budapest, Hungary*

[†]*AP and MAI contributed equally to this work and should be considered co-first authors*

Authors:

Manuel Alcaraz-Ibáñez (m.alcaraz@ual.es); Health Research Centre and Department of Education, University of Almería, Spain [a]; ORCID ID: 0000-0002-1297-9565

Adrian Paterna (a.paterna@ual.es); Health Research Centre and Department of Education, University of Almería, Spain [a]; ORCID ID: 0000-0002-7561-9487

Mark D. Griffiths (mark.griffiths@ntu.ac.uk); Psychology Department, Nottingham Trent University, UK[b]; ORCID ID: 0000-0003-3070-1804

Zsolt Demetrovics (zsolt.demetrovics@unigib.edu.gi); Centre of Excellence in Responsible Gaming, University of Gibraltar, Gibraltar, Gibraltar; Institute of Psychology, ELTE Eötvös Loránd University, Budapest, Hungary; ORCID ID: 0000-0001-5604-7551

Alvaro Sicilia (asicilia@ual.es); Health Research Centre and Department of Education, University of Almería, Spain [a]; ORCID ID: 0000-0001-9436-8743

Author's Note

Correspondence concerning this article should be addressed to Adrian Paterna, Universidad de Almería, Facultad de Ciencias de la Educación, Carretera de Sacramento s/n, 04120, La Cañada de San Urbano (Almería), Spain. Tel: +34+950 015376.

E-mail: a.paterna@ual.es

Contributors

AP and MAI designed the study, performed the systematic search and data extraction, completed all statistical analyses and initial drafts of the manuscript. AS, ZD and MDG contributed to the drafting of the manuscript and revisions. All authors assisted with drafting of the final version of the manuscript, including critical revisions for intellectual content.

Role of Funding Sources

This research is part of the I+D+I project (grant number PID2019-107674RB-I00), funded by Ministerio de Ciencia e Innovación (MCIN), Agencia Estatal de Investigación (AEI/10.13039/501100011033), Spain.

AP (FPU18/01055) is funded by MCIN/AEI/10.13039/501100011033 and Fondo Social Europeo (FSE) “El FSE invierte en tu futuro”.

MAI (UAL RRA202101) is funded by Ministerio de Universidades (Plan de Recuperación, Transformación y Resiliencia, Next Generation EU).

ZD’s contribution was supported by the Hungarian National Research, Development and Innovation Office (KKP126835).

Conflicts of interest

The authors declare no potential conflicts of interest with respect to the research, authorship and/or publication of this article

Gender-related differences in self-reported problematic exercise symptoms: A systematic review and meta-analysis

Abstract

Aims: To provide quantitative summarized evidence on gender-related differences in self-reported problematic exercise symptoms (PE). **Methods:** Eligible studies were searched up to December 31, 2021 in the databases *MEDLINE, Current Contents Connect, PsycINFO, Web of Science, SciELO,* and *Dissertations & Theses Global.* Studies were considered eligible if they included information that allowed the calculation of the differences of interest as expressed by either the aggregate or subscales scores of the main self-reported instruments of PE identified by previous research (i.e., Commitment to Exercise Scale, Compulsive Exercise Test, Exercise Addiction Inventory, Exercise Dependence Questionnaire, Exercise Dependence Scale-Revised, and Obligatory Exercise Questionnaire). Data were analysed using three-level meta-analytic models. Potential moderator variables were examined using meta-regressions. **Results:** A total of 168 effect-sizes from 117 studies (N=65,718) were retrieved. Results showed (i) small overall differences favouring males for the aggregate scores of the instruments (g=0.105), (ii) small-to-moderate differences favouring females for symptoms involving withdrawal (g=0.116 and 0.118), lack of exercise enjoyment (g=0.226), and the employment of exercise as a means to ends such as health improvement (g=0.222), mood management (g=0.158 and 0.226), and body weight control (g=0.453 and 0.465); and (iii) small differences favouring males for symptoms involving spending considerable amount of time in the activity (g=0.250), exercising with greater volume/intensity than planned (g=0.254), a need for increased amounts of exercise to achieve the desired effect (g=0.291), loss of control over the behaviour (g=0.101), reduction or cessation of other activities because of exercise (g=0.323), and continue to exercise despite physical and/or psychological issues being caused or exacerbated by this behaviour (g=0.243). **Conclusions:** Adopting a gender-informed perspective may be needed both in the professional praxis of exercise and health practitioners prescribing and guiding exercise practice and in the design of prevention and treatment efforts aimed at avoiding the occurrence of PE.

Keywords: morbid exercise; exercise dependence; exercise addiction; compulsive exercise; obligatory exercise; problematic exercise; meta-analysis

Gender-related differences in self-reported symptoms of problematic exercise: A systematic review and meta-analysis

Regular exercise is recognized as an effective non-pharmacological strategy in the prevention of a range of pathologies (Bennie et al., 2020; Hu et al., 2020; Thompson et al., 2020). Nevertheless, there is also evidence that some individuals may develop potentially problematic patterns of exercise behaviour (Juwono & Szabo, 2021). This refers to a complex behaviour involving a number of manifestations such as (among others) losing control over exercise habits to the point of experiencing impairment at the physical, psychological, or professional levels (Bamber et al., 2003; Szabo et al., 2015), and/or developing withdrawal symptoms such as anxiety or depression as a result of being prevented from exercising (Weinstein et al., 2017). In view of the potential harmful health implications of this behaviour (which, given the multiplicity of terms used in the literature, is referred to here by using the umbrella term *problematic exercise*; PE) (Sicilia et al., 2021), its potential risk factors need to be identified.

One of the variables that has attracted much research interest when examining the risk factors for PE is gender (Downs et al., 2019). Findings from the only review paper specifically focused on examining gender-related differences in PE concluded than males are more prone to develop this problematic form of exercise than females (Dumitru et al., 2018). However, some of the shortcomings of this review may have compromised the accuracy of its findings. Firstly, it lacked a systematic and reproducible methodology in the process of searching for relevant literature and in the subsequent reporting and discussion of the results (e.g., the one described in consensus guidelines) (Page et al., 2021). Secondly, it did not implement statistical techniques to provide a quantitative summary of the results (Borenstein et al., 2009). Thirdly, it only considered data from two of the many self-report instruments developed for the assessment of PE (Sicilia et al., 2021). This shortcoming is an important limitation in view of evidence suggesting that, due to the complex nature of PE (Szabo et al., 2015), no single psychometric instrument currently available comprehensively assesses the behaviour (Sicilia et al., 2021). Lastly, it did not explore the factors that may account for the contradictory results concerning gender-related differences in PE reported in the literature (Alcaraz-Ibáñez et al., 2019; Costa et al., 2013; de la Vega et al., 2020; Weik & Hale, 2009).

One candidate factor to be explored within the context of examining the causes underlying the inconsistent results regarding gender-related differences in self-reported PE symptoms may be the existence of assessment-related artefacts. This possibility is plausible given not all of the symptoms proposed as indicative of a problematic pattern of exercise behaviour are always (and to the same extent) present in each of the available instruments (Sicilia et al., 2022). Indeed, specific symptoms included in each of these instruments have found in some cases to be reported in varying degree by males and females (Goodwin et al., 2016; Kotbagi et al., 2017; Weik & Hale, 2009). Another plausible explanation for the observed variability in gender-related differences in PE could be offered according to the influence that several sociodemographic variables may have on the PE scores. This may be the case for age (as older individuals are found to be less prone to develop PE) (Alcaraz-Ibáñez et al., 2018, 2019; Bueno-Antequera et al., 2020; Edmunds et al., 2006), the main exercise modality practised (as individuals practising mainly endurance modalities have been found to be at higher risk of showing problematic patterns of exercise behaviour) (Di Lodovico et al., 2019), and the risk status in terms of eating disorders (as individuals at-risk of eating disorders tend to show higher risk of PE) (Trott et al., 2021). The latter is particularly relevant given the higher prevalence rates of eating disorders found among females compared to males (Smink et al., 2012).

As research examining gender-related differences in PE increases (Alcaraz-Ibáñez et al., 2019; Bueno-Antequera et al., 2020; Karademir, 2020), systematic and comprehensive evaluation of accumulated data that provides deeper insight into this topic is needed. In this vein, meta-analytic techniques provide a reliable method for quantifying differences between groups in a given variable using data gathered from existing literature, as well as for examining the methodological or sociodemographic characteristics that may be affecting these differences (Borenstein et al., 2009). Therefore, a systematic literature review and meta-analysis was conducted with the purpose of addressing two main questions: (i) to what extent do males and females differ in their self-reported PE levels? and (ii) what methodological or sociodemographic variables may amplify, attenuate, inhibit or conceal these differences? Answering these questions may lead to the identification of populations particularly susceptible to developing problematic patterns of exercise behaviour, which may translate into improved professional praxis among exercise and health practitioners. Such findings may also provide evidence on the need to address gender-specific efforts that, aimed at preventing and treating PE, may contribute to maximize the potential health-inducing benefits of exercise.

METHODS

The present study was conducted according to the Preferred Reporting Items for Systematic Reviews and Meta Analyses (PRISMA) statement (Page et al., 2021) (see

Appendix A for the PRISMA checklist) and was pre-registered in PROSPERO (CRD42021237104).

Locating studies

Eligible studies were searched up to December 31, 2021 in the databases *MEDLINE, Current Contents Connect, PsycINFO, Web of Science, SciELO,* and *Dissertations & Theses Global* (see Appendix B for the full search strategy). Reference lists of retrieved studies were hand-searched to identify further potentially eligible studies.

Endnote X9 software was used for managing references at the screening stage. Studies were independently selected by the two first authors in two stages following examination of (a) their titles/abstracts, and (b) their full-texts. Inter-coder reliability in terms of Cohen's Kappa, as computed by 'ReCal' (Freelon, 2013) was .62 (percent agreement 99%) for the abstract/title, and .87 (percent agreement 94%) for the full text. In the presence of suspected duplicate studies (e.g., a dissertation and its derived peer-reviewed publication), only published data were employed. Disagreements were discussed and resolved on a consensual basis with the assistance of a third author, if necessary.

Corresponding authors of the retrieved studies were approached to request unpublished data that may be potentially eligible for inclusion. Missing relevant information for a given retrieved study (e.g., age) was requested from the corresponding authors. The response rate (i.e., the percentage of authors that, after being asked, provided data that were effectively analysed) was 51.4%.

Eligibility criteria

The present review gathered data on gender-related differences in PE symptoms as assessed by self-report instruments. In the interest of minimising publication bias, the literature search aimed to retrieve data from published and unpublished research.

Inclusion criteria. Studies meeting the following criteria were considered eligible: (a) at least one of the following validated self-report instruments (i.e., those whose psychometric properties have been formally tested in a peer-reviewed study) identified in previous meta-analytic research (Alcaraz-Ibáñez et al., 2020; Alcaraz-Ibáñez, Paterna, et al., 2021; Trott et al., 2021) was used for the purpose of assessing PE symptoms: Commitment to Exercise Scale (CES) (Davis et al., 1993), Compulsive Exercise Test (CET) (Taranis et al., 2011), Exercise Addiction Inventory (EAI) (Terry et al., 2004), Exercise Dependence Questionnaire (EDQ) (Ogden et al., 1997), Exercise Dependence Scale-Revised (EDS-R) (Downs et al., 2004), and Obligatory Exercise Questionnaire (OEQ) (Steffen & Brehm, 1999); (b) studies were written in English, Spanish French, or Portuguese, although there was no restrictions in terms of country of origin; and (c) sufficient data were available for calculation of the effect sizes corresponding either to global scores for a given instrument or its sub-scales.

Exclusion criteria. Studies meeting the following criteria were excluded: (a) available PE scores were offered just as composite scores obtained by adding global scores derived from (i) more than one instrument, or (ii) several sub-scales whose aggregate score did not equal to the global score of a given instrument; (b) specific items were excluded when obtaining global PE scores and the sub-scale scores were not available; (c) specific items were excluded from sub-scale PE scores; and (d) available PE scores were obtained using a modified factor structure from the one originally proposed for the instrument.

Coding procedure

A coding frame was developed (and subsequently pilot-tested) according to the common features of the studies retrieved in a preliminary search. The resulting coding sheet (see Appendix C) was independently used by the two authors in charge of extracting the relevant data from the retrieved studies. Inter-coder reliability (Cohen's Kappa) ranged from .72 to .93 (percent agreement 87% to 98%). Disagreements were discussed and resolved on a consensual basis with the assistance of a third author, if needed.

Risk of bias

Assessment of risk of bias was conducted employing the adapted Newcastle-Ottawa Scale (NOS) for evaluating cross-sectional/survey studies (Hillen et al., 2017). The 0-16 range score of the NOS results from the evaluation of: (a) clarity of stated aim; (b) representativeness of the sample; (c) sample size; (d) non-respondents; (e) ascertainment of the exposure; (f) control of confounding factors; (g) comparability of participants in different outcome groups; (h) assessment of the outcome; and (i) statistical tests. Low scores on the NOS suggest higher risk of bias. The risk of bias assessment was independently conducted by the two first authors. Disagreements between reviewers were discussed and resolved on a consensual basis with the assistance of a third author, if needed. As a result of this procedure, the 117 retrieved studies were scored between 7 and 12 in terms of risk of bias.

Statistical analysis

Gender-related differences in PE scores were expressed as the standardized sample-size corrected mean-change (Hedges' g). Prior to the calculation of the effect sizes, the *SD* of the scores derived from studies reporting just the standard error of the mean were obtained by applying the following formula (Higgins et al., 2019):

$$SD = SE \times \sqrt{\mathrm{N}}$$

In cases where it was not possible to obtain the *SD* values, effect sizes were computed from available statistics (i.e., *t*, *d*, or *r*). Eventual attenuations of the effect sizes due to the level of measurement error of the instruments assessing the construct of interest were corrected by using their reliability values according to the procedure described elsewhere (Schmidt & Hunter, 2015). When reliability values were not reported in the retrieved studies, the values provided in a recent reliability generalisation meta-analysis were used (Alcaraz-Ibáñez et al., 2022). A negative effect size implies a lower score in the female group.

Several features present in some of the primary studies included in the present meta-analysis could imply a violation of the principle of independence of effect sizes inherent to this technique (Becker, 2000). The first one concerns the existence of multiple effect size in studies with a longitudinal design (e.g., Goodwin et al., 2014a, 2014b). This was addressed by employing a three-level random effects model that, accounting for the hierarchical structure of the data, allows for a differentiated examination of (a) the sampling variance for the observed effect sizes (level 1); (b) the variance between effect sizes from the same study (level 2); and (c) the variance between studies (level 3) (Cheung, 2014; Van den Noortgate et al., 2013). The adequacy of the described three-level random effect model with respect to its less complex alternative (i.e., a two-level random effect model not assuming that some of the effect sizes are nested within the studies) was checked by means of a likelihood-ratio test. A second source of dependence was the presence of effect sizes from several population subgroups within the same study (e.g., in terms of country of origin) (de la Vega et al., 2020). This was approached by treating each effect size in the three-level random effect model as if it were derived from an independent study. A last feature that could imply that the principle of independence is being violated was the presence of effect sizes corresponding to several instruments coming from the same population (e.g., Alcaraz-Ibáñez et al., 2019). This was addressed by conducting random removal of effect sizes until just one of them remained (Cheung, 2014).

The presence of statistical heterogeneity at levels two and three was examined and quantified by the respective use of the *Q-test* and the I^2 statistic, with values of 25%, 50%, and 75% of the latter being respectively interpreted as indicative of low, moderate, and high heterogeneity (Higgins et al., 2003). In the presence of heterogeneity, potential sources of variance both in categorical and continuous codified variables were explored using employing mixed-effects three-level meta-regressions models. A binary code was employed to transform categorical variables into *k*-1 dummy variables. Explained variance by the moderators was quantified on a percentage basis and expressed by R^2. The presence of potential outliers and, therefore, the robustness of the results was examined using graphic display of study heterogeneity (GOSH) plot analysis. This procedure allows for fitting not only *K* models but also modelling for all 2^{k-1} possible study combinations. Once the models are calculated by employing three cluster algorithms (i.e., k-means, DBSCAN, and Gaussian mixture models), a plot is obtained in which the pooled effect size and the between-study heterogeneity are respectively displayed on the x- and on y-axis (Olkin et al., 2012). Cook's distance values are subsequently employed for the purpose of assessing whether a given study could be particularly influencing within the context of the emerging clusters (Harrer et al., 2021).

Publication bias was examined using a three-parameter selection model (3PSM) involving a simple model with a single cut-off point (<.05) and no moderators. The resulting unadjusted and adjusted meta-analytic models are compared by means of a likelihood-ratio test. Statistically significant results on this test suggest that the adjusted model should be retained and the likely existence of publication bias (Coburn & Vevea, 2019). The use of 3PSM has been recommended over other available methodological alternatives for the examination of publication bias in the presence of a high degree of heterogeneity (Carter et al., 2019).

Point mean estimates of effect sizes were interpreted as trivial (.00 to .10), small (.10 to .40), moderate (.40 to .70), and large (>.70) (Cohen, 1988). The described statistical analyses were conducted in R (version 3.6.1). The three-level random-effects models were estimated using a method robust to the absence of normal data distributions (i.e., restricted maximum likelihood; REML) (Langan et al., 2019).

RESULTS

Description of studies

A total of 3954 studies were initially identified. As a result of the study selection procedure (see Figure 1), 117 primary studies involving 168 effect sizes (*N*=65,217) published between 1988 and 2021, inclusive, were included in the systematic review and meta-analysis (see Appendix D for the complete list). The main characteristics of the retrieved studies are shown in Table 1. From the studies included in the meta-analyses, 100 were published peer-reviewed papers and 17 were doctoral dissertations or conference proceedings. The retrieved effect sizes were obtained by employing the CES (Likert-scale version, *K*=4; Visual-Analogue-Scale version, *K*=15), the CET (*K*=16), the EAI (*K*=49), the EDQ (*K*=8), the EDS-R (*K*=50), and the OEQ (*K*=26). The retrieved studies were conducted in Asia (*K*=3), Oceania (*K*=6), Europe (*K*=76), Latin America (*K*=21), and North America (*K*=36). From the studies included in the meta-analyses, 115 employed a cross-sectional design (*K*=154) while five employed a longitudinal design (*K*=14). None of the retrieved studies indicated whether the samples included individuals clinically diagnosed with an eating disorder. Mean age of the participants included in the meta-analysis ranged from 12.69 to 51.94 years (M_{age}=26.11 years, SD_{age}=8.46) and in BMI from 19.86 to 26.63 (M_{BMI} =23.12, SD_{BMI} =1.42).

Gender-related differences in PE symptoms

The results from the likelihood-ratio test [$\chi^2(1)$=6.319, *p*=.012] suggest the adequacy of the three-level over the two-level random effect model. Findings from the three-level random effects model showed a small but near to trivial effect size favouring males (Hedge's *g*= -.104, *p*=.004; 95%CI= -.176 to -.033). The results from the *Q*-test indicated significant heterogeneity (*Q*=1527.59, $\tau^2_{(level2)}$=.047, $\tau^2_{(level3)}$ =.139), which was globally estimated in terms of the I^2 statistic to be 92.95% ($I^2_{(level2)}$=25.44, $I^2_{(level3)}$=67.51). Findings from the main univariable meta-regression analysis (see Table 1) demonstrated that PE measure was the only variable explaining significant variance. More specifically, small differences were found that in some cases favoured males (CES-VAS, EDSR-R, and the OEQ) and in other cases favoured females (CET and EDQ). Trivial differences that favoured males where found for the EAI.

In view of the results of the moderator analyses, gender differences in individual symptoms included in each of the multidimensional instruments under consideration (i.e., CET, EDQ, and EDS-R) was also examined using the same three-level approach employed in the previous analyses. The results from these analyses (see Table 2) showed (i) the adequacy of the three-level over the two-level random effect model; and (ii) effect size estimates slightly lower in magnitude but still consistent with those observed in the previous analyses. Findings from the univariable meta-regression analysis (see Table 3) showed that considering the sub-scales of the instruments under consideration explained significant variance in all three cases. For the CET, differences favouring mainly females were found for the different subscales, which ranged from small (for symptoms involving using exercise as a mean of mood improvement and lack of exercise enjoyment) to moderate (for symptoms involving exercise as a mean of weight control). For the EDQ, small differences favouring mainly females were found, which ranged from small (for symptoms involving positive rewards, withdrawal, and exercising for health reasons) to moderate (for symptoms involving exercise as a mean of weight control). In the case of the EDS-R, small differences favouring males were found in all subscales except for the one covering withdrawal symptoms, which favoured females.

Sensitivity analysis and publication bias

After removing five effect sizes from four studies (Costa et al., 2013; Hill et al., 2015; Kotbagi et al., 2017; Zeeck et al., 2017) identified as potential outliers from the results of influence analyses (see Appendix E), the results from the adjusted model (Hedge's *g*= -.100, p<.001; 95%CI= -.149 to -.051; $I^2_{(level2)}$=35.46, $I^2_{(level3)}$=48.44) were found to be consistent with those from the non-adjusted one (Hedge's *g*= -.105, p= .004; 95%CI= -.176 to -.033; $I^2_{(level2)}$=25.44, $I^2_{(level3)}$=67.51). The results of 3PSM suggested publication bias in the non-adjusted model (χ^2[1]=22.612, p<.001) but not in the one where outliers were removed (χ^2[1]=2.550, p=.110).

DISCUSSION

Meta-analytic techniques were used to provide quantitative summarization of gender-related differences in self-report PE symptoms, as well as to identify the circumstances under which these differences may vary. Results from a three-level random-effects model including 168 effect sizes from 117 studies comprising 65,217 participants indicated two main group of findings. Firstly, the existence of overall small and near to trivial differences favouring males for the aggregate scores of the set of psychometric instruments being considered. Secondly, that these differences vary across specific instrument and their subscales, with those derived from the CET/EDQ and subscales covering symptoms referred to exercising to control body weight (female-predominant) and the EDS-R/CES-VAS and subscales covering symptoms involving reductions in daily life because of exercise (male-predominant) showing the more extreme opposite effects. Therefore, it follows that both the direction and the magnitude of gender-related differences on PE may differ according to the specific symptoms involved in this type of problematic behaviour.

Findings presented here using data derived from six different assessment instruments are in line with those reported in a previous narrative review considering just two instruments (i.e., EDS-R and EAI) (Dumitru et al., 2018). More specifically, both reviews agree that when these two instruments are used, males are more prone to report potentially problematic patterns of exercise behaviour than females. The present study adds to the findings of this previous review by quantifying these differences as trivial in the case of EAI and small in the case of the EDS-R. However, arguably, the main contribution of the present study is that it demonstrates that part of the variability in gender-related differences in self-reported PE levels are due to assessment-related issues. In the case of the aggregate scores, the reason for the change in the general trend of aggregate scores favouring males may be found in the existence of particularly greater differences in symptoms being present just in the two instruments whose scores were found to favour females (i.e., CET and EDQ). A close inspection of the content of these instruments points to the symptom involving exercising as a means of body weight control as the only one being present in both of them but not in the remaining instruments (Sicilia et al., 2022). The occurrence of these particularly large differences for such symptoms is corroborated by the evidence obtained from the analyses examining the contribution of the symptoms present in the different subscales of each instrument to the variability of the differences of interest. This evidence also highlights the fact that males and females may not be equally susceptible to manifesting the full range of PE symptoms. Therefore, females are particularly likely to show PE patterns mainly characterised by experiencing withdrawal symptoms, having a lack of exercise enjoyment, and using exercise as a means to ends such as health improvement, mood management, and particularly body weight control. On the contrary, males tend to report PE patterns to a greater extent characterised by spending considerable amount of time on the activity, exercising with greater volume/intensity than planned, needing increased amounts of exercise to achieve the desired effect, experiencing a loss of control over the behaviour, reducing or ceasing other activities because of exercise, and continuing to exercise despite physical and/or psychological issues being caused or exacerbated by the behaviour.

A first consideration on these differences concerns the non-equivalent nature of all specific symptoms proposed as indicative of PE according to their damage potential (Sicilia et al., 2020). Both the loss of control and the existence of negative consequences in terms of experiencing harm either at the physical, psychological, or social level as a result of exercise behaviour have been proposed as the most distinctive and deleterious features of PE (Szabo et al., 2018). The fact that these kinds of features seem to be mainly present among the subscales whose scores favoured males suggests that individuals of this gender may show a somewhat more problematic PE risk profile. However, the risk pattern shown by females is also not without its dangers. For example, the symptom with the largest magnitude of differences favouring females (i.e., exercising to control body weight) has been proposed as a likely reinforcing factor in the maintenance of thinness-related eating disorders (Alcaraz-Ibáñez et al., 2020; Schaumberg et al., 2022). It could be also argued that engaging in exercise as a way of coping with internal distress does not necessarily imply harm caused by the exercise behaviour. However, the fact that this strategy is perceived as the only effective way to deal with negative mood may lead to exacerbation of these moods in situations where the individual is prevented from exercising, for example, as a result of being injured (Freimuth et al., 2011; Lichtenstein et al., 2018).

A second relevant consideration on the observed differences concerns their plausible causes. One of them may be drawn from pressures towards the thin-body ideal traditionally assigned to females in Western culture, which leads females to experience greater weight concerns and increased propensity to exercise for weight control reasons than males (Pritchard & Beaver, 2012; Sicilia et al., 2020; Wright et al., 2006). Moreover, females are also subjected to health imperatives according to which they must be slim not only to be attractive but also to be healthy, a goal that females are particularly inclined to see as achievable through exercise (Welsh, 2011; Wright et al., 2006). Therefore, it is conceivable to assume that perceiving that exercising contributes towards achieving the body-related goals being pursued may translate into reinforced exercise behaviour. This suggests that these reinforcements may also derive from two types of mood changes resulting from exercise behaviour: (i) an improvement of these moods as a result of engaging in the behaviour, not because it is pleasant, but because it is likely contributing to the intended purpose; and (ii) a worsening of those moods at the prospect of losing the opportunity to contribute to the intended purpose as a result of not being able to engage in the behaviour. Another conceivable cause underlying the gender-related differences found is the greater predisposition shown by males to exercise for motives largely inherent in the activity itself. This would be the case of those referred to as skill development, performance, or competition (Ley, 2020; Rodrigues et al., 2022; Wright et al., 2006). In view of the above, it seems reasonable to assume that this latter category of goals would be the predominant object of the reinforcements leading to an over-committed and ultimately problematic exercise behaviour in the case of males. However, these explanations are presented merely as a possibility whose empirical validity should be subject to further research.

Implications for professional practice and future research

The first group of main implications from the present study concerns the professional practice of health and exercise practitioners. The findings here suggest that prevention efforts aimed at avoiding the occurrence of PE may benefit from adopting a gender-informed perspective when prescribing and guiding exercise practice. For females, this would particularly imply avoiding providing positive (e.g., excessive praise) or negative (e.g., stimulating feelings of shame or guilt) reinforcement for the respective success or failure in achieving weight control and/or health improvement related goals (Alcaraz-Ibáñez, Chiminazzo, et al., 2021; Alcaraz-Ibáñez, Paterna, et al., 2021). For males, it would particularly imply providing specific information on the main deleterious outcomes derived from losing control and/or obsessing with exercise to the point of not integrating this activity harmoniously into the rest of their life activities (e.g., reducing social life or neglecting professional responsibilities due to exercise behaviour) (Juwono & Szabo, 2021) or putting one's own health at risk (e.g., persist on exercising despite an over-use injury) (Lichtenstein et al., 2014).

A second group of notable implications concern future research and treatment efforts in the PE field. Firstly, the findings suggest that efforts should be made to control for the likely confounding effect of gender when testing explanatory models aimed at examining the potential causes underlying PE, even more so when the outcomes of interest are specific PE symptoms. Secondly, the findings suggest that treatment interventions could probably benefit from adopting a gender-specific approach. Cognitive-behavioural interventions used to manage other problematic behaviours have been proposed as worthwhile to be explored within the context of PE (Downs et al., 2019). Therefore, knowledge of the characteristics involved in the specific PE symptoms that males and females are more likely to experience may be employed to identify gender-specific maladaptive thoughts and beliefs being challenged according to the perspective of this therapeutic approach (Kahl et al., 2012).

Limitations and future directions

A first limitation of the present study concerns the lack of data needed for conducting more comprehensive moderator analyses. This originated from at least three different sources. Firstly, the rather incomplete description of the populations in some of the retrieved studies, which prevented detailed exploration of some of the socio-demographic factors that may have accounted for the variability of the differences under consideration. Examples of the latter are the main exercise modality practiced or the risk status in terms of eating disorders (Alcaraz-Ibáñez et al., 2020; Di Lodovico et al., 2019). It is therefore possible that increased data availability for these variables may have allowed additional sources of variability to be identified. Secondly, the under-representation of some potentially relevant populations (e.g., individuals either at-risk or clinically diagnosed with an eating disorder) (Alcaraz-Ibáñez et al., 2020). Thirdly, the traditional binary approach adopted in nearly all the studies included in the present meta-analysis prevented any examination of whether the scores under consideration may vary across the spectrum of gender identities. This limitation highlights the need for future primary research in this area to adhere to guidelines on gender equity in research, so that individuals may also identify themselves as non-binary (Heidari et al., 2016). Fourthly, the fact that mainly aggregated PE scores were provided in the retrieved studies, which implied that evidence on the differences of interest across the different subscales were derived from a lower number of studies. This is a particularly relevant shortcoming considering the markedly multidimensional nature of the PE symptoms (Formby et al., 2014; Sicilia & González-Cutre, 2011). This is even more so given the results showed varying directions and magnitudes between symptoms on the differences of interest. In view of this group of limitations, improved homogeneity in the reporting of primary studies examining PE is warranted. A meaningful move in this direction could be to reach an expert agreement that allows for extending one of the existing reporting guidelines generically proposed for observational studies (e.g., Von Elm et al., 2007) for the purpose of covering the particularities of the specific research field of PE. Similarly, future studies are needed that, by considering individual scores on the range of symptoms potentially involved in PE, examine whether the differences observed in the present study can be extended to a wider range of populations.

A second limitation worth mentioning derives from shortcomings in the reporting of the reliability values of the instruments used to obtain the scores of interest. This consisted either in the complete omission of such values (e.g., Bueno-Antequera et al., 2020; Mayolas-Pi et al., 2017) or the absence of detailed reporting when PE scores were provided for different groups (e.g., Yager & O'Dea, 2010; Yildiz & Senel, 2020). This implies that, even with all the analytical correction efforts made (Schmidt & Hunter, 2015), some degree of bias resulting from not being able to employ the very precise reliability value associated with each effect size cannot be ruled out. In the light of this limitation, researchers are encouraged to follow the recommendations concerning reliability reporting specifically proposed for this field of research (Alcaraz-Ibáñez et al., 2022).

A last noteworthy limitation stems from the limited number of studies that have so far examined the invariant nature in terms of gender among some of the scores of interest (Alcaraz-Ibáñez et al., 2022). In the absence of strong evidence on this matter, it cannot be precluded that differences observed are partly due to differing interpretation by males and females of either individual items or the underlying factor (van de Schoot et al., 2012). In view of this limitation, future primary research on this topic might benefit from gathering preliminary evidence on the gender invariance of the scores derived from the instruments before examining the differences of interest.

CONCLUSIONS

The present study contributes to the literature by quantifying the magnitude of gender-related differences in self-reported PE symptoms, as well as by examining the circumstances under which these differences might vary. The findings support the existence of small and near to trivial differences favouring males when PE is expressed according to the aggregate scores derived from the whole set of instruments, as well as the tendency for these differences to become more pronounced and favour either males or females depending on specific symptoms. Therefore, females and males are respectively more likely to report symptoms involving exercising to serve a specific purpose (mainly body weight control and mood modification) or physical/social harms derived from progressive over-involvement. Adopting a gender-informed perspective may be needed both in the professional praxis of exercise and health practitioners prescribing and guiding exercise practice and in the design of prevention and treatment programs aimed at avoiding the occurrence of PE. Further research is warranted that, by considering different populations in terms of their sociodemographic characteristics, may provide additional insight into the causes underlying gender-related differences in specific PE symptoms.

Acknowledgments: The authors wish to express their gratitude to Ornella Corazza, Lorna Jakobson, Andrea Lukács, Zsuzsa Menczel, Halley Pontes, Attila Szabó, Mike Trott and Aviv Weinstein for providing the required raw data.

REFERENCES

Alcaraz-Ibáñez, M., Aguilar-Parra, J. M., & Álvarez-Hernández, J. F. (2018). Exercise addiction: Preliminary evidence on the role of psychological inflexibility. *International Journal of Mental Health and Addiction*, *16*(1), 199–206. https://doi.org/10.1007/s11469-018-9875-y

Alcaraz-Ibáñez, M., Chiminazzo, J. G. C., Sicilia, A., & Teixeira Fernandes, P. (2021). Body and appearance-related self-conscious emotions and exercise addiction in Brazilian adolescents: A person-centred study. *Journal of Sports Sciences*, *39*(13), 1528–1536. https://doi.org/10.1080/02640414.2021.1883290

Alcaraz-Ibáñez, M., Paterna, A., Sicilia, A., & Griffiths, M. D. (2020). Morbid exercise behaviour and eating disorders: A meta-analysis. *Journal of Behavioral Addictions*, *9*(2), 206–224. https://doi.org/10.1556/2006.2020.00027

Alcaraz-Ibáñez, M., Paterna, A., Sicilia, A., & Griffiths, M. D. (2021). A systematic review and meta-analysis on the relationship between body dissatisfaction and morbid exercise behaviour. *International Journal of Environmental Research and Public Health*, *18*, 585. https://doi.org/10.3390/ijerph18020585

Alcaraz-Ibáñez, M., Paterna, A., Sicilia, A., & Griffiths, M. D. (2022). Examining the reliability of the scores of self-report instruments assessing problematic exercise: A systematic review and meta-analysis. *Journal of Behavioral Addictions*, 1–21. https://doi.org/10.1556/2006.2022.00014

Alcaraz-Ibáñez, M., Sicilia, A., Dumitru, D. C., Paterna, A., & Griffiths, M. D. (2019). Examining the relationship between fitness-related self-conscious emotions, disordered eating symptoms, and morbid exercise behavior: An exploratory study. *Journal of Behavioral Addictions*, *8*(3), 603–612. https://doi.org/10.1556/2006.8.2019.43

Bamber, D. J., Cockerill, I. M., Rodgers, S., & Carroll, D. (2003). Diagnostic criteria for exercise dependence in women. *British Journal of Sports Medicine*, *37*(5), 393–400. https://doi.org/10.1136/bjsm.37.5.393

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

Bennie, J. A., Shakespear-Druery, J., & De Cocker, K. (2020). Muscle-strengthening exercise epidemiology: A new frontier in chronic disease prevention. *Sports Medicine - Open*, *6*, 40. https://doi.org/10.1186/s40798-020-00271-w

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

Bueno-Antequera, J., Mayolas-Pi, C., Reverter-Masià, J., López-Laval, I., Oviedo-Caro, M. Á., Munguía-Izquierdo, D., Ruidíaz-Peña, M., & Legaz-Arrese, A. (2020). Exercise addiction and its relationship with health outcomes in indoor cycling practitioners in fitness centers. *International Journal of Environmental Research and Public Health*, *17*, 4159. https://doi.org/10.3390/ijerph17114159

Carter, E. C., Schönbrodt, F. D., Gervais, W. M., & Hilgard, J. (2019). Correcting for bias in psychology: A comparison of meta-analytic methods. *Advances in Methods and Practices in Psychological Science*, *2*(2), 115–144. https://doi.org/10.1177/2515245919847196

Cheung, M. W. L. (2014). Modeling dependent effect sizes with three-level meta-analyses: A structural equation modeling approach. *Psychological Methods*, *19*(2), 211–229. https://doi.org/10.1037/a0032968

Coburn, K. M., & Vevea, J. L. (2019). *Estimating weight-function models for publication bias* (2.0.2).

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

Costa, S., Hausenblas, H. A., Oliva, P., Cuzzocrea, F., & Larcan, R. (2013). The role of age, gender, mood states and exercise frequency on exercise dependence. *Journal of Behavioral Addictions*, *2*(4), 216–223. https://doi.org/10.1556/JBA.2.2013.014

Davis, C., Brewer, H., & Ratusny, D. (1993). Behavioral frequency and psychological commitment: Necessary concepts in the study of excessive exercising. *Journal of Behavioral Medicine*, *16*(6), 611–628. https://doi.org/10.1007/BF00844722

de la Vega, R., Almendros, L. J., Barquín, R. R., Boros, S., Demetrovics, Z., & Szabo, A. (2020). Exercise addiction during the COVID-19 pandemic: An international study confirming the need for considering passion and perfectionism. *International Journal of Mental Health and Addiction*. https://doi.org/10.1007/s11469-020-00433-7

Di Lodovico, L., Poulnais, S., & Gorwood, P. (2019). Which sports are more at risk of physical exercise addiction: A systematic review. *Addictive Behaviors*, *93*, 257–262. https://doi.org/10.1016/j.addbeh.2018.12.030

Downs, D. S., Hausenblas, H. A., & Nigg, C. R. (2004). Factorial validity and psychometric examination of the Exercise Dependence Scale-Revised. *Measurement in Physical Education and Exercise Science*, *8*(4), 183–201. https://doi.org/10.1207/s15327841mpee0804

Downs, D. S., MacIntyre, R. I., & Heron, K. E. (2019). Exercise addiction and dependence. In M. H. Anshel, S. J. Petruzzello, & E. E. Labbé (Eds.), *APA Handbook of Sport and Exercise Psychology, Volume 2: Exercise Psychology* (Vol. 2, pp. 589–604). American Psychological Association. https://doi.org/10.1037/0000124-030

Dumitru, D. C., Dumitru, T., & Maher, A. J. (2018). A systematic review of exercise addiction: Examining gender differences. *Journal of Physical Education and Sport*, *18*(3), 1738–1747. https://doi.org/10.7752/jpes.2018.03253

Edmunds, J., Ntoumanis, N., & Duda, J. L. (2006). Examining exercise dependence symptomatology from a self-determination perspective. *Journal of Health Psychology*, *11*(6), 887–903. https://doi.org/10.1177/1359105306069091

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*(4), 555–557. https://doi.org/10.1016/j.eatbeh.2014.08.013

Freelon, D. (2013). ReCal OIR: Ordinal, interval, and ratio intercoder reliability as a web service. *International Journal of Internet Science*, *8*(1), 10–16.

Freimuth, M., Moniz, S., & Kim, S. R. (2011). Clarifying exercise addiction: Differential diagnosis, co-occurring disorders, and phases of addiction. *International Journal of Environmental Research and Public Health*, *8*(10), 4069–4081. https://doi.org/10.3390/ijerph8104069

Goodwin, H., Haycraft, E., & Meyer, C. (2014a). Emotion regulation styles as longitudinal predictors of compulsive exercise: A twelve month prospective study. *Journal of Adolescence*, *37*(8), 1399–1404. https://doi.org/10.1016/j.adolescence.2014.10.001

Goodwin, H., Haycraft, E., & Meyer, C. (2014b). Psychological risk factors for compulsive exercise: A longitudinal investigation of adolescent boys and girls. *Personality and Individual Differences*, *68*, 83–86. https://doi.org/10.1016/j.paid.2014.03.048

Goodwin, H., Haycraft, E., & Meyer, C. (2016). Disordered eating, compulsive exercise, and sport participation in a UK adolescent sample. *European Eating Disorders Review*, *24*(4), 304–309. https://doi.org/10.1002/erv.2441

Harrer, M., Cuijpers, P., Furukawa, T. A., & Ebert, D. D. (2021). *Doing meta-analysis with R: A hands-on guide* (1st ed.). Chapman & Hall/CRC Press.

Heidari, S., Babor, T. F., De Castro, P., Tort, S., & Curno, M. (2016). Sex and gender equity in research: rationale for the SAGER guidelines and recommended use. *Research Integrity and Peer Review*, *1*(2), 1-9. https://doi.org/10.1186/s41073-016-0007-6

Higgins, J. P. T., Li, T., & Deeks, J. J. (2019). Choosing effect measures and computing estimates of effect. In *Cochrane Handbook for Systematic Reviews of Interventions* (pp. 143–176). Wiley. https://doi.org/10.1002/9781119536604.ch6

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

Hill, A. P., Robson, S. J., & Stamp, G. M. (2015). The predictive ability of perfectionistic traits and self-presentational styles in relation to exercise dependence. *Personality and Individual Differences*, *86*, 176–183. https://doi.org/10.1016/j.paid.2015.06.015

Hillen, M. A., Medendorp, N. M., Daams, J. G., & Smets, E. M. A. (2017). Patient-driven second opinions in oncology: A systematic review. *The Oncologist*, *22*(10), 1197–1211. https://doi.org/10.1634/theoncologist.2016-0429

Hu, M. X., Turner, D., Generaal, E., Bos, D., Ikram, M. K., Ikram, M. A., Cuijpers, P., & Penninx, B. W. J. H. (2020). Exercise interventions for the prevention of depression: A systematic review of meta-analyses. *BMC Public Health*, *20*, 1255. https://doi.org/10.1186/s12889-020-09323-y

Juwono, I. D., & Szabo, A. (2021). 100 cases of exercise addiction: More evidence for a widely researched but rarely identified dysfunction. *International Journal of Mental Health and Addiction*, *19*, 1799–1811. https://doi.org/10.1007/s11469-020-00264-6

Kahl, K. G., Winter, L., & Schweiger, U. (2012). The third wave of cognitive behavioural therapies: What is new and what is effective? *Current Opinion in Psychiatry*, *25*(6), 522–528. https://doi.org/10.1097/YCO.0b013e328358e531

Karademir, T. (2020). The effects of regular sports activities on exercise dependence. *International Journal of Applied Exercise Physiology*, *9*(9), 190–197.

Kotbagi, G., Morvan, Y., Romo, L., & Kern, L. (2017). Which dimensions of impulsivity are related to problematic practice of physical exercise? *Journal of Behavioral Addictions*, *6*(2), 221–228. https://doi.org/10.1556/2006.6.2017.024

Langan, D., Higgins, J. P. T., Jackson, D., Bowden, J., Veroniki, A. A., Kontopantelis, E., Viechtbauer, W., & Simmonds, M. (2019). A comparison of heterogeneity variance estimators in simulated random-effects meta-analyses. *Research Synthesis Methods*, *10*(1), 83–98. https://doi.org/10.1002/jrsm.1316

Ley, C. (2020). Participation motives of sport and exercise maintainers: Influences of age and gender. *International Journal of Environmental Research and Public Health*, *17*(21), 1–13. https://doi.org/10.3390/ijerph17217830

Lichtenstein, M. B., Christiansen, E., Elklit, A., Bilenberg, N., & Støving, R. K. (2014).

Exercise addiction: A study of eating disorder symptoms, quality of life, personality traits and attachment styles. *Psychiatry Research*, *215*(2), 410–416. https://doi.org/10.1016/j.psychres.2013.11.010

Lichtenstein, M. B., Nielsen, R. O., Gudex, C., Hinze, C. J., & Jørgensen, U. (2018). Exercise addiction is associated with emotional distress in injured and non-injured regular exercisers. *Addictive Behaviors Reports*, *8*, 33–39. https://doi.org/10.1016/j.abrep.2018.06.001

Mayolas-Pi, C., Simón-Grima, J., Peñarrubia-Lozano, C., Munguía-Izquierdo, D., Moliner-Urdiales, D., & Legaz-Arrese, A. (2017). Exercise addiction risk and health in male and female amateur endurance cyclists. *Journal of Behavioral Addictions*, *6*(1), 74–83. https://doi.org/10.1556/2006.6.2017.018

Ogden, J., Veale, D., & Summers, Z. (1997). The development and validation of the Exercise Dependence Questionnaire. *Addiction Research*, *5*(4), 343–356. https://doi.org/10.3109/16066359709004348

Olkin, I., Dahabreh, I. J., & Trikalinos, T. A. (2012). GOSH - a graphical display of study heterogeneity. *Research Synthesis Methods*, *3*(3), 214–223. https://doi.org/10.1002/jrsm.1053

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., … Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *International Journal of Surgery*, *88*, 105906. https://doi.org/10.1016/j.ijsu.2021.105906

Pritchard, M. E., & Beaver, J. L. (2012). Do exercise motives predict obligatory exercise? *Eating Behaviors*, *13*(2), 139–141. https://doi.org/10.1016/j.eatbeh.2011.11.012

Rodrigues, F., Moutão, J., Teixeira, D., Cid, L., & Monteiro, D. (2022). Examining exercise motives between gender, age and activity: A first-order scale analysis and measurement invariance. *Current Psychology*, *41*(1), 112–125. https://doi.org/10.1007/s12144-019-00560-y

Schaumberg, K. E., Robinson, L., Hochman, A., & Micali, N. (2022). Prospective associations between driven exercise and other eating disorder behaviors in adolescence: A longitudinal cohort study. *Journal of Adolescent Health*, 1–7. https://doi.org/10.1016/j.jadohealth.2021.11.022

Schmidt, F. L., & Hunter, J. E. (2015). *Methods of meta-analyisis: Correcting error and bias* *in research findings* (3rd ed.). Sage Publications. https://doi.org/10.4135/9781483398105

Sicilia, A., Alcaraz-Ibáñez, M., Chiminazzo, J. G. C., & Fernandes, P. T. (2020). Latent profile analysis of exercise addiction symptoms in Brazilian adolescents: Association with health-related variables. *Journal of Affective Disorders, 273,* 223–230. https://doi.org/10.1016/j.jad.2020.04.019

Sicilia, A., Alcaraz-Ibáñez, M., Paterna, A., & Griffiths, M. D. (2022). A review of the components of problematic exercise in psychometric assessment instruments. *Frontiers in Public Health, 10,* 839902. https://doi.org/10.3389/fpubh.2022.839902

Sicilia, A., & González-Cutre, D. (2011). Dependence and physical exercise: Spanish validation of the Exercise Dependence Scale-Revised (EDS-R). *The Spanish Journal of Psychology, 14*(1), 421–431. https://doi.org/10.5209/rev_SJOP.2011.v14.n1.38

Sicilia, A., Paterna, A., Alcaraz-Ibáñez, M., & Griffiths, M. D. (2021). Theoretical conceptualisations of problematic exercise in psychometric assement instruments: A systematic review. *Journal of Behavioral Addictions, 10*(1), 4–20. https://doi.org/10.1556/2006.2021.00019

Smink, F. R. E., Van Hoeken, D., & Hoek, H. W. (2012). Epidemiology of eating disorders: Incidence, prevalence and mortality rates. *Current Psychiatry Reports, 14*(4), 406–414. https://doi.org/10.1007/s11920-012-0282-y

Steffen, J. J., & Brehm, B. J. (1999). The dimensions of obligatory exercise. *Eating Disorders, 7*(3), 219–226. https://doi.org/10.1080/10640269908249287

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). Routledge.

Szabo, A., Griffiths, M. D., de La Vega Marcos, R., Mervó, B., & Demetrovics, Z. (2015). Methodological and conceptual limitations in exercise addiction research. *Yale Journal of Biology and Medicine, 88,* 303–308.

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*(3), 256–268. https://doi.org/10.1002/erv.1108

Terry, A., Szabo, A., & Griffiths, M. D. (2004). The Exercise Addiction Inventory: A new brief screening tool. *Addiction Research and Theory, 12*(5), 489–499. https://doi.org/10.1080/16066350310001637363

Thompson, W. R., Sallis, R., Joy, E., Jaworski, C. A., Stuhr, R. M., & Trilk, J. L. (2020). Exercise is medicine. *American Journal of Lifestyle Medicine, 14*(5), 511–523.

https://doi.org/10.1177/1559827620912192
Trott, M., Jackson, S. E., Firth, J., Jacob, L., Grabovac, I., Mistry, A., Stubbs, B., & Smith, L. (2021). A comparative meta-analysis of the prevalence of exercise addiction in adults with and without indicated eating disorders. *Eating and Weight Disorders*, *26*(1), 37–46. https://doi.org/10.1007/s40519-019-00842-1
van de Schoot, R., Lugtig, P., & Hox, J. (2012). A checklist for testing measurement invariance. *European Journal of Developmental Psychology*, *9*(4), 486–492. https://doi.org/10.1080/17405629.2012.686740
Van den Noortgate, W., López-López, J. A., Marín-Martínez, F., & Sánchez-Meca, J. (2013). Three-level meta-analysis of dependent effect sizes. *Behavior Research Methods*, *45*(2), 576–594. https://doi.org/10.3758/s13428-012-0261-6
Von Elm, E., Altman, D. G., Egger, M., Pocock, S. J., Gøtzsche, P. C., & Vandenbroucke, J. P. (2007). The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: Guidelines for reporting observational studies. *PloS Medicine*, *4*(10), 1623–1627. https://doi.org/10.1371/journal.pmed.0040296
Weik, M., & Hale, B. D. (2009). Contrasting gender differences on two measures of exercise dependence. *British Journal of Sports Medicine*, *43*(3), 204–207. https://doi.org/10.1136/bjsm.2007.045138
Weinstein, A. A., Koehmstedt, C., & Kop, W. J. (2017). Mental health consequences of exercise withdrawal: A systematic review. *General Hospital Psychiatry*, *49*, 11–18. https://doi.org/10.1016/j.genhosppsych.2017.06.001
Welsh, T. (2011). Healthism and the bodies of women: Pleasure and discipline in the war against obesity. *Journal of Feminist Scholarship*, *1*, 33–48.
Wright, J., O'Flynn, G., & MacDonald, D. (2006). Being fit and looking healthy: Young women's and men's constructions of health and fitness. *Sex Roles*, *54*, 707–716. https://doi.org/10.1007/s11199-006-9036-9
Yager, Z., & O'Dea, J. (2010). A controlled intervention to promote a healthy body image, reduce eating disorder risk and prevent excessive exercise among trainee health education and physical education teachers. *Health Education Research*, *25*(5), 841–852. https://doi.org/10.1093/her/cyq036
Yildiz, M., & Senel, E. (2020). The relationship between sport motivation and exercise dependence: Comparing Turks living in different countries. *International Journal of Applied Exercise Physiology*, *9*(7), 173–181.
Zeeck, A., Schlegel, S., Giel, K. E., Junne, F., Kopp, C., Joos, A., Davis, C., & Hartmann, A.

(2017). Validation of the German version of the Commitment to Exercise Scale. *Psychopathology*, *50*(2), 146–156. https://doi.org/10.1159/000455929

Table 1

Results of overall univariable meta-regression analyses

Moderators	s	k	β_0	95% CI		β_1	95% CI		χ^2	p	R^2_{level2}	R^2_{level3}	τ^2_{Level2}	τ^2_{Level3}
				Lower	Upper		Lower	Upper						
Eating disorders	161	168							0.351	.950	.000	.004	.051	**.140**
Unknown (RC)	127	134	**-0.102**	**-0.183**	**-0.022**									
At-risk	3	3	-0.192	-0.707	0.323	-0.089	-0.611	0.431						
Not at-risk	10	10	-0.046	-0.331	0.240	0.056	-0.239	0.354						
Mixed	21	21	-0.131	-0.325	0.063	-0.029	-0.239	0.181						
Report of LTE	161	168							0.006	.940	.000	.002	.051	**.138**
No (RC)	61	66	-0.101	-0.216	0.014									
Yes	100	102	**-0.106**	**-0.198**	**-0.015**	-0.007	-0.153	0.141						
Regular exercisers	161	168							0.609	.435	.029	.000	.049	**.139**
Unknown (RC)	78	83	**-0.133**	**-0.234**	**-0.032**									
Yes	83	85	-0.076	-0.177	0.025	-0.057	-0.086	0.200						
EP measure	161	168							13.709	.033	.000	.191	.054	**.120**
CET (RC)	13	16	0.210	-0.013	0.433									
CES	4	4	-0.085	-0.506	0.337	-0.295	-0.771	0.182						
CES-VAS	15	15	**-0.249**	**-0.475**	**-0.024**	**-0.459**	**-0.776**	**-0.142**						
EAI	47	49	-0.073	-0.199	0.053	**-0.283**	**-0.539**	**0.027**						
EDQ	8	8	0.103	-0.201	0.406	-0.107	-0.483	0.269						
EDS-R	50	50	**-0.218**	**-0.338**	**-0.097**	**-0.427**	**-0.681**	**-0.174**						
OEQ	24	26	-0.111	-0.292	0.070	**-0.321**	**-0.608**	**-0.034**						
Region	161	168							2.311	.889	.005	.015	.051	**.143**
Unknown (RC)	22	24	-0.081	-0.274	0.112									
Latin America	21	21	-0.022	-0.222	0.178	0.059	-0.219	0.337						
Oceania	6	6	-0.123	-0.503	0.257	-0.042	-0.468	0.384						
North America	36	36	-0.111	-0.261	0.039	-0.030	-0.275	0.214						
Mixed	2	2	-0.329	-0.947	0.289	-0.248	-0.896	0.399						
Europe	71	76	**-0.133**	**-0.240**	**-0.027**	-0.052	-0.273	0.168						
Asia	3	3	0.140	-0.396	0.676	0.221	-0.349	0.791						
Type of survey	161	168							3.200	.362	.000	.031	.052	**.137**
Unknown (RC)	54	58	**-0.170**	**-0.292**	**-0.049**									
Paper-pencil	54	57	-0.029	-0.152	0.095	0.142	-0.031	0.315						
Online	52	52	-0.120	-0.245	0.005	0.051	-0.123	0.225						
Both	1	1	0.224	-0.627	1.076	0.395	-0.466	1.255						
Publication status	161	168							.554	.457	.000	.007	.047	**.139**
Published (RC)	144	151	**-0.095**	**-0.170**	**-0.019**									
Unpublished	17	17	-0.186	-0.404	0.031	-0.092	-0.322	0.139						
Study design	161	168							1.413	.234	.000	.061	.054	**.132**
Longitudinal (RC)	7	14	**-0.114**	**-0.187**	**-0.042**									
Cross-sectional	154	154	0.086	-0.233	0.406	0.201	-0.127	0.528						
Continuous moderators														
Age	124	127	**-0.135**	**-0.211**	**-0.059**	0.006	-0.003	0.015	80.221	<.001	.096	.000	.074	.086
BMI	61	61	**-0.186**	**-0.320**	**-0.051**	0.041	-0.054	0.135	131.356	.063	.012	.012	.128	.128
Year of publication	161	168	**-0.104**	**-0.176**	**-0.033**	0.001	-0.006	0.010	0.052	.818	.000	.001	.051	**.137**
Quality	161	168	**-0.104**	**-0.176**	**-0.033**	0.015	-0.059	0.088	0.156	.693	.003	.001	.051	**.138**

Note: s = Number of studies; k = Number of effect sizes; β_0 = Intercept/mean effect size; β_1 = Estimated regression coefficient; CI= Confidence interval; RC= Reference category; R^2 = Explained variance; τ^2_{Level2} = Variance between the effect sizes from the same study; τ^2_{Level3} = Variance between studies; LTE = Leisure time exercise; CES = Commitment to Exercise Scale (Likert scale); CES-VAS = Commitment to Exercise Scale (Visual Analogue Scale); CET = Compulsive Exercise Test; EAI = Exercise Addiction Inventory; EDQ = Exercise Dependence Questionnaire; EDS-R = Exercise Dependence Scale-Revised; OEQ = Obligatory Exercise Questionnaire.

Statistically-significant effects ($p < .05$) appear highlighted in bold.

Table 2

Results of analyses for aggregate scores of multidimensional instruments

Outcome	*s*	*k*	ES (*g*)	95% CI		χ^2	*p*	*Q*	τ^2_{Level2}	τ^2_{Level3}	I^2_{Level2}	I^2_{Level3}
				Lower	Upper							
CET	4	20	0.180	-0.061	0.421	4.761	.029	**102.496**	**.042**	.068	.382	.435
EDQ	6	48	0.088	-0.081	0.256	4.121	.042	**150.212**	**.056**	.039	.464	.242
EDS-R	25	175	**-0.181**	**-0.282**	**-0.079**	38.321	<.001	**1689.816**	**.059**	**.056**	.459	.411

Note: s = Number of studies; *k* = Number of effect sizes; ES = Pooled effect size; *g* = Corrected Hedges' g; CI= Confidence interval; τ^2_{Level2} = Variance between the effect sizes from the same study; τ^2_{Level3} = Variance between studies; CET = Compulsive Exercise Test; EDQ= Exercise Dependence Questionnaire; EDS-R = Exercise Dependence Scale-Revised.

Statistically-significant effects ($p < .05$) appear highlighted in bold.

Table 3

Results of univariable meta-regression analyses for specific symptoms included in multidimensional instruments

Moderators	*s*	*k*	β_0	95% CI		β_1	95% CI		χ^2	*p*	R^2_{Level2}	R^2_{Level3}	τ^2_{Level2}	τ^2_{Level3}
				Lower	Upper		Lower	Upper						
CET	4	20							10.354	.034	.820	.000	.018	.754
Weight control exercise (RC)	4	4	**0.465**	**0.173**	**0.757**									
Avoidance and rule-driven behaviour	4	4	-0.024	-0.313	0.265	**-0.489**	**-0.730**	**-0.247**						
Mood improvement	4	4	0.158	-0.133	0.449	**-0.306**	**-0.548**	**-0.064**						
Lack of exercise enjoyment	4	4	0.226	-0.056	0.507	**-0.239**	**-0.470**	**-0.007**						
Exercise rigidity	4	4	0.053	-0.242	0.348	**-0.412**	**-0.660**	**-0.163**						
EDQ	6	48							24.613	.001	.788	.000	.023	.044
Weight control (RC)	6	6	**0.453**	**0.213**	**0.693**									
Interference	6	6	-0.017	-0.254	0.221	**-0.470**	**-0.726**	**-0.214**						
Positive reward	6	6	0.226	-0.008	0.461	-0.227	-0.480	0.027						
Withdrawal	6	6	0.118	-0.120	0.357	**-0.335**	**-0.592**	**-0.078**						
Insight into problem	6	6	-0.176	-0.418	0.065	**-0.629**	**-0.889**	**-0.370**						
Social reasons	6	6	-0.013	-0.259	0.233	**-0.466**	**-0.729**	**-0.203**						
Health reasons	6	6	0.222	-0.014	0.457	-0.232	-0.485	0.022						
Stereotyped behaviour	6	6	-0.141	-0.384	0.102	**-0.594**	**-0.856**	**-0.333**						
EDS-R	25	175							63.313	<.001	.672	.000	**.022**	**.059**
Withdrawal (RC)	25	25	**0.169**	**0.044**	**0.295**									
Tolerance	25	25	**-0.291**	**-0.416**	**-0.166**	**-0.460**	**-0.577**	**-0.344**						
Intention effects	25	25	**-0.254**	**-0.379**	**-0.130**	**-0.424**	**-.0540**	**-.0307**						
Lack of control	25	25	**-0.101**	**-0.227**	**0.025**	**-0.270**	**-0.389**	**-0.151**						
Time	25	25	**-0.250**	**-0.375**	**-0.125**	**-0.419**	**-0.537**	**-0.302**						
Reduction in other activities	25	25	**-0.323**	**-0.452**	**-0.194**	**-0.492**	**-0.614**	**-0.371**						
Continuance	25	25	**-0.243**	**-0.368**	**-0.118**	**-0.412**	**-0.530**	**-0.295**						

Note: *s* = Number of studies; *k* = Number of effect sizes; β_0 = intercept/mean effect size; β_1= estimated regression coefficient*;* CI = Confidence interval; Lo = RC = Reference category; R^2 = Explained variance; τ^2_{Level2} = Variance between the effect sizes from the same study; τ^2_{Level3} = Variance between studies; CET = Compulsive Exercise Test; EDQ = Exercise Dependence Questionnaire; EDS-R = Exercise Dependence Scale-Revised.

Statistically-significant effects ($p < .05$) appear highlighted in bold.

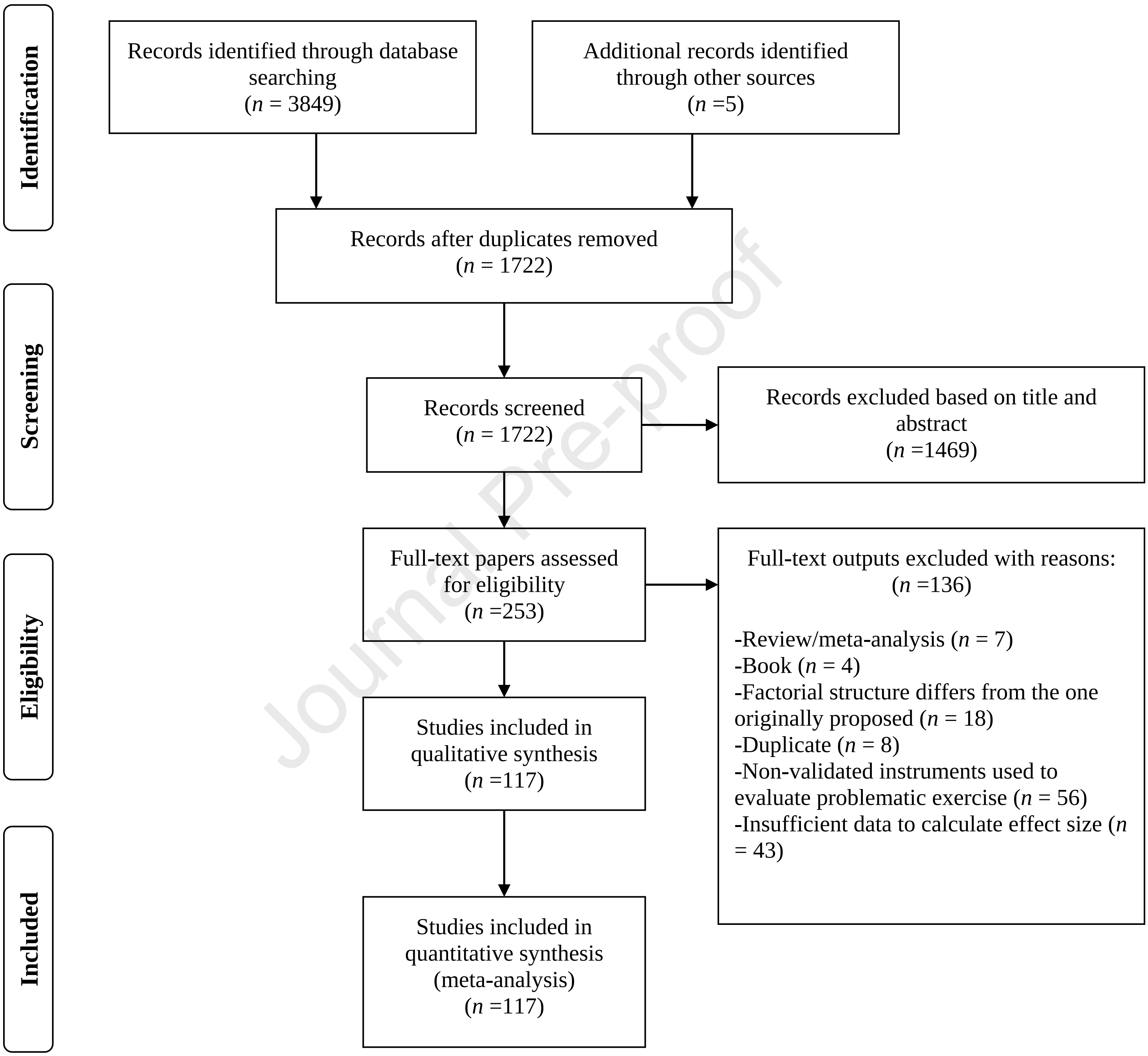


Figure 1. PRISMA flow diagram of study

- Inconsistent results on gender-related differences in problematic exercise symptoms (PES) have been reported
- For the first time, meta-analytic techniques are employed to examine gender-related differences in PES
- Females are more likely to report PES involving exercise for weight control and mood modification purposes
- Males are more likely to report PES involving harms derived from exercise over-involvement

Conflicts of interest

The authors declare no potential conflicts of interest with respect to the research, authorship and/or publication of this article