ORIGINAL ARTICLE

Development and cross-cultural validation of the Goal Content for Weight Maintenance Scale (GCWMS)

Jorge Encantado[1,2] · Marta M. Marques[1,3] · António L. Palmeira[1,4] · Simon J. Sebire[5] · Pedro J. Teixeira[1] · R. James Stubbs[6] · Berit L. Heitmann[7,8] · Maria J. Gouveia[2]

Abstract

Purpose Long-term weight management requires sustained engagement with energy-balance-related behaviours. According to self-determination theory, behaviour goals can support or undermine motivation depending on the quality of their content (i.e., extrinsic and intrinsic motivation). This study aimed to develop and validate the goal content for weight loss maintenance scale (GCWMS).

Methods The GCWMS was administered to 1511 participants who had achieved clinically significant weight losses and were taking part in a large weight loss maintenance study: the NoHoW Trial (ISRCTN88405328). The scale derived from two well-established questionnaires regarding exercise goals. Construct validity was examined for 4 theory-driven domains: *Health Management*, *Challenge*, *Image*, and *Social Recognition*. Split-sample confirmatory factor analysis was conducted to test the factorial validity and multi-group measurement invariance (configural, metric, scalar, and residual invariance). The reliability estimates were also assessed, and discriminant validity was evaluated using 2 conceptually related questionnaires.

Results The first analysis showed a poor fit of the original factorial structure. Subsequent investigation with a new specified model indicated close fit to the data after removal of 3 items $\chi^2(58) = 599.982$; $p < .001$; $\chi^2/df = 10.345$; CFI = 0.940; GFI = 0.941; SRMR = 0.063; RMSEA = 0.079 (LL = 0.073; UL = .084). Good internal consistency was achieved in all subscales ($\alpha > .775$), convergent and divergent validity were verified through associations with other theoretical related constructs. Findings from multi-group invariance test demonstrated that the specified model of GCWMS achieved full measurement invariance for gender but did not support residual invariance across countries.

Conclusion Findings support the hypothesised four-dimension structure of the GCWMS, confirming reliability and multi-group invariance in factor structure. Analysis also supports valid group means comparisons on latent factors at gender and at cross-cultural level. Ways to improve the quality of the scale are discussed.

Level of evidence Level V, cross-sectional descriptive

Keywords Factorial validity · Weight loss maintenance · Motivation · Self-determination theory · Questionnaire

✉ Jorge Encantado
jencantado@fmh.ulisboa.pt

✉ Marta M. Marques
mmoreira@tcd.ie

1 Centro Interdisciplinar para o Estudo da Performance Humana (CIPER), Faculdade de Motricidade Humana, Universidade de Lisboa, Cruz Quebrada, Portugal

2 Applied Psychology Research Center Capabilities & Inclusion (APPsyCI), ISPA – Instituto Universitário, Lisboa, Portugal

3 Trinity College Dublin, ADAPT Centre & Trinity Centre for Practice and Healthcare Innovation, College Green, Dublin, Ireland

4 Centro de Investigação em Desporto, Educação Física, Exercício e Saúde (CIDEFES), Universidade Lusófona, Lisboa, Portugal

5 Centre for Exercise, Nutrition and Health Sciences, School for Policy Studies, University of Bristol, Bristol, UK

6 Faculty of Medicine and Health, School of Psychology, University of Leeds, Leeds, UK

7 Research Unit for Dietary Studies at The Parker Institute Bispebjerg and Frederiksberg Hospital, Part of the Copenhagen University Hospital-The Capital Region, Copenhagen, Denmark

8 Section for General Medicine, The Department of Public Health, University of Copenhagen, Copenhagen, Denmark

Springer

Introduction

Overweight and obesity increase the risk of developing type 2 diabetes [1], cardiovascular disease [2], cancer [3], among other conditions. Despite efforts to lose and maintain weight loss, weight regain after considerable weight loss is often reported [4–6]. Modifiable energy-balance behaviours, such as a healthy diet and physical activity, are crucial for weight loss and weight maintenance [7, 8].

Behavioural interventions focusing on diet and physical activity show beneficial effects for at-risk populations [9, 10] and there is now a reasonably robust evidence-base for the effectiveness of theory-based behavioural change interventions, focusing on individual and social factors [11–13] (for a review: [14]). Recent systematic reviews have reported evidence that increasing levels of autonomous motivation, self-regulation skills and self-efficacy mediated long-term weight management in lifestyle interventions (e.g. [15, 16]) and interventions based on self-determination theory (SDT) are designed to include these components. SDT literature seems to give promising guidance to the explanation of long-lasting health-related behaviours [17–19], including those relevant for weight management [20] and long-term exercise among women [21].

Goal content from the perspective of self-determination theory

Self-determination theory [22] posits that humans have three innate psychological needs that are required for psychological growth: the need for autonomy, or the experience of authorship, volition and integrity of one's actions; the need for competence or the feeling of being effective in responding to challenges; and the need for relatedness, or the sense of belonging and being cared for. To nourish psychological needs, need-supportive social environments are crucial when striving for behaviour change. However, different health goals provide different motivational energy towards health behaviours, such as physical activity or healthy eating.

In goal content theory (a sub-theory in SDT), these behavioural goals are therefore defined as intrinsic or extrinsic based on the nature of their content (e.g., to improve health vs to improve image [23, 24]) and each kind of goal may relate differently to the satisfaction of the basic psychological needs (e.g., [25]). For instance, intrinsic goals are developed from within the individual and are more likely to satisfy psychological needs, for example "to seek out novelty and challenges, to extend and exercise one's capacity to explore and to learn" [22 (pp 70)]. In

contrast, extrinsic goals have an outward instrumental orientation typically focussed on attaining external indicators of worth and are not related to or might undermine psychological need satisfaction (e.g., wealth, possessions, appearance) [26]. The definition of intrinsic and extrinsic goals (the "what" of the behaviour) is conceptually different to the behavioural regulation that motivates people towards that goal which can vary from being highly autonomous to highly controlled (the "why" of the behaviour; for a review see [22, 27]).

The measurement of goal content [intrinsic vs. extrinsic] is critical to understand the relationships between goals, behavioural regulations, health behaviours and long-term psychological well-being [28]. The content of people's goals when pursuing weight loss maintenance is an important but understudied area related to the maintenance of energy-balance behaviours over time (i.e. physical activity and healthy eating). Some advancements have been made in the development of self-reported measures to assess goal content. For example, the Goal Content for Exercise Questionnaire [29] evaluates three intrinsic (i.e., skill development, social affiliation and health management) and two extrinsic (i.e., social recognition and image) exercise goals and previous work reported positive association between intrinsic goals and exercise engagement and psychological well-being (e.g., [27]).

Furthermore, we aimed at confirming measurement invariance and cross-cultural generalizability. According to Sue [30], cross-cultural comparisons that allow to test the equivalence of the results of a psychometric instrument is a fundamental approach to test the cross-cultural applicability of theories and models. To test measurement invariance is key when using psychological measures in group comparison, such as countries with different languages and cultural background, as it exposes the degree to which items response patterns are preserved and maintain their meaning across all the tested groups [31, 32]. Simply put, using a non-invariant measure to compare different groups is worthless because the instrument will perform differently across groups. This way, one may misinterpret mean differences [33] that will have a detrimental impact when testing the efficacy of interventional trials with large multi-country samples. When measurement invariance is not met, it may reflect differential functioning of the measure across the different groups rather than the intended differences on the construct that is being measured [31]. Following this argument, because women are generally more concerned with weight loss maintenance-related outcomes than men (e.g., body dissatisfaction) [34], we also investigated model measurement invariance across gender.

Within the context of a large European Trial (the NoHoW H2020 Trial), based on SDT theoretical framework, we assessed the conceptual structure of the motives for weight

loss maintenance to understand if the intervention was effective in helping participants in this transition from extrinsic motives to intrinsic ones. The understanding of the motivational role of peoples' goals for weight loss maintenance may help the development of more effective interventions. From an SDT perspective, the present study developed an adapted instrument to assess the content of people's goals for maintaining weight loss.

Aims

The objectives of this study were to first develop the Goal Content for Weight Maintenance Scale (GCWMS), and second examine its psychometric properties by testing: (1) the factorial validity; (2) the measurement invariance across groups by comparing the stability of the factorial model in a large and diverse population of adults who have lost weight from three European countries—Denmark, Portugal, United Kingdom; (3) the internal consistency and capacity to discriminate the intrinsic and extrinsic goals as defined in SDT; and (4) the convergent and external validity by comparing with validated measures of behavioural regulation for exercise and eating.

Methods

Participants

The data used for this validation procedure were collected in the context of the baseline data from a European Commission funded intervention project: the NoHoW Trial (trial registration number ISRCTN88405328). The NoHoW project is testing a digital toolkit to support weight maintenance in 2 × 2 factorial randomised controlled trial in 1627 European adults (*M* age = 44.01 ± 11.86; 68.7% women) from Denmark, Portugal and the United Kingdom (UK), who achieved 5% weight loss, in the 12 months prior trial enrolment (for more information about this project see [35]).

Following Hair et al. guidance [36], we chose to include only individuals with complete data to use the largest sample size possible without imputing values (73 individuals from 1627 were excluded). Multivariate outliers were detected through Mahalanobis distance and 43 observations reporting a relatively high D_M^2 with a *p* value below 0.001 were removed [37]. After data cleaning, the sample available for this study was 1511 individuals (*M* age = 44; SD = 11.9; 68.3% women; for complete information on descriptive statistics of the final sample after missing data and outlier removal please refer to supplementary file 1 [https://osf.io/vfmgq/]). Approximately 33.7% were from Portugal (*N* = 509; *M* age = 40, SD = 9.7; 44.8% women), 34.1% from the UK (*N* = 516; *M* age = 44.7, SD = 13; 79.3% women), and 32.2%

from Denmark (*N* = 486; *M* age = 47.4, SD = 11.4; 81.3% women). Full details on demographics per country available on supplementary file 1 (https://osf.io/vfmgq/).

Scale development and procedure

The GCWMS' items (see full original scale in supplementary file 2: https://osf.io/5n9h2/) were adapted from two well-validated and commonly used scales to assess goal content (Goal Content for Exercise Questionnaire [GCEQ] [29]) and motives (Exercise Motives Inventory-2 [38]) for exercise. The original GCEQ was adapted to the weight loss maintenance framework by changing the stem (i.e., "I exercise to…" to "I manage my weight to….") and wording of the behaviour expressed in each item (i.e., "exerciser" to "healthy person"). However, the original 4 items tapping Skill Development are exercise specific and could not be adapted to weight loss maintenance. As such, 4 items focusing on challenge and skill development (i.e., "To give me goals to work towards") were used from the Exercise Motives Inventory-2 [38]. The initial item pool comprised 16 items addressing four types of goals that people may have to maintain their weight loss in a non-orthogonal model: *Challenge* and *Health Management* for intrinsic goals, and *Image* and *Social Recognition* for extrinsic goals.

These changes were conducted by a panel of specialists in the fields of psychology, psychometrics, obesity, and behaviour change. All items and instructions were then translated to Portuguese and Danish, and back-translated to English by an external company (Ipsos Mori) and then revised by the panel of specialists to ensure that content and meaning were preserved. The original GCEQ response format (1-to-7-point response scale, from *strongly disagree* to *strongly agree*), was used to allow participants to indicate the extent to which each item was important to them.

Data collection

Responses were taken via the Qualtrics™ online platform as part of the *NoHoW* trial; all participants provided signed, informed consent. Ethical approval was given by institutions involved in the study (Universities of Leeds [17–0082; 27 February 2017], Lisbon [17/2016; 20 February 2017] and Capital Region of Denmark [H-16030495, 8 March 2017]). All data were anonymised, and subjects were given a unique identification code. Information concerning data handling is available elsewhere [35].

Data analysis

We conducted confirmatory factor analysis procedures using the maximum likelihood estimation method to test validity and invariance of factorial structure within and across

country and gender. More specifically, we intended with this study to (1) test for the factorial validity of the GCWMS for an European sample, (2) to cross-validate findings across a second independent European sample, and (3) to test for invariant factorial measurement and structure across gender (women/men) and country (Portugal/Denmark/UK) with a larger sample. To achieve this purpose, we analysed factorial adjustment indices, internal reliability, construct validity, and construct reliability. To test the factorial structure of GCWMS we cross-validated the modified model with split-sample data in a three-stage process. First, a preliminary CFA was done using a randomly selected sub-sample of approximately half of the subjects of the Portuguese data set to test the factorial validity of the four-factor model regarding weight loss maintenance, because the original factorial structure was conceived in a different behavioural domain—exercise [29]. The Portuguese sample was used as it presented the most balanced number of men and women (44.8% women). Second, the second half of the Portuguese sample comprising the remaining participants was used to test independently the hypothesised changes to the model. Third, the fit of the final model was examined in the full sample. Because large samples are susceptible to multivariate outliers, we first inspected the data for distribution of normality (kurtosis; skewness) and checked potential multivariate outliers and Mardia's coefficient for multivariate kurtosis (see supplementary file 3 for AMOS outputs: https://osf.io/hs5aj/). Because the value was superior to the multivariate normality recommended cut-off (> 5.0) [39], we proceeded with Bollen–Stine bootstrap with 2000 samples for all analysis [40] and Spearman's rho as the correlation coefficient to test the strength of association between scale factors.

Confirmatory factor analysis was conducted using AMOS Version 25 [41]. Kline's references [37] were used to analyse univariate normality (skew index ≤ 3.0; kurtosis index ≤ 10.0). Good model fit cut-off values adopted are those proposed by Hair et al. [42] and Schumacker and Lomax [43] (samples with $N > 250$ and number of observed variables between 12 and 30): Chi-square test of model fit (χ^2/df) values less than 5 reflects a good model fit; comparative fit index (CFI) values close to 0.90 or 0.95 reflects a good model fit; goodness-of-fit index (GFI) values close to 0.90 or 0.95 indicates a good model fit; standardized root mean square residual (SRMR) and root mean square error of approximation (RMSEA) values below 0.05 indicate good fit and values between 0.05 to 0.08 with a CFI of 0.92 or higher indicate close model fit. Convergent validity was established by the average variance extracted (AVE ≥ 0.50; [42]) and internal consistency was assessed through Cronbach's alpha ($\alpha \geq 0.70$) and composite reliability (CR ≥ 0.70) [44]. We also evaluated the modification indices to identify potential model specifications (MI < 10) [39].

To test the degree of measurement invariance of the scale by gender and country, we conducted the commonly used likelihood ratio test (differences in Chi-square between two nested models); however, it is sensitive to sample size [37] and change in CFI was the primary measure of invariance (models are equivalent when ΔCFI ≤ 0.01) [29, 39, 45]. Further, we checked for changes in ΔSRMR (≤ 0.030 for metric invariance; ≤ 0.010 for scalar or residual invariance) and ΔRMSEA (≤ 0.015 for metric invariance, scalar and residual invariance) following Chen's guidelines [46]. We employed a sequential model testing approach where increasingly constrained models were specified and compared (i.e., model estimated freely across all groups simultaneously) to evaluate (1) configural invariance (Model 1; i.e., whether items were associated with the same constructs between groups); (2) metric invariance (Model 2; i.e., to test equivalence of the item loadings on the factors); and (3) scalar invariance (Model 3; i.e., to test equivalence of item intercepts, for metric invariant items); and (4) residual invariance (Model 4; i.e., to test equivalence of item residuals of metric and scalar invariant items) [47].

To examine external convergent validity, two other questionnaires were used to compare the potential correlations of goal content scores with measures of behavioural regulation: Behaviour Regulations for Exercise Questionnaire 3 (BREQ3 [48, 49]) and Regulations for Eating Behaviour Scale (REBS [50]). The BREQ3 was originally developed to assess six different behavioural regulations for the exercise domain, as conceptualized in SDT. The questionnaire contains 24 items using a 1-to-7-point response scale (*strongly disagree* to *strongly agree*) which measures amotivation, external regulation, introjected regulation, identified regulation, integrated regulation, and intrinsic motivation. The same six behavioural regulations for eating behaviour were assessed using the 18-item REBS with the same 1-to-7-point response scale. The BREQ3 and REBS subscale scores were aggregated to form scores for the second order latent factors of autonomous and controlled motivation for exercise and for eating behaviours, as proposed in SDT, that were subsequently used in the correlational analysis with the GCWMS factors. All BREQ3 and REBS latent factors displayed good internal consistency (BREQ3: controlled motivation $a = 0.761$; autonomous motivation $a = 0.948$; REBS: controlled motivation $a = 0.742$; autonomous motivation $a = 0.893$). Based on SDT literature, we anticipated that intrinsic goals (*Challenge* and *Health Management*) would be more positively correlated with autonomous motivation for healthy eating and exercise than with controlled motivation, and extrinsic goals (*Image* and *Social Recognition*) would be more positively correlated with controlled motivation for healthy eating and exercise than with autonomous motivation.

Results

Confirmatory factor analysis

Following the recommended procedures for factorial validation [51], the first factorial analysis was conducted with a small sub-sample randomly extracted from the NoHoW Trial database. To reduce measurement errors, we extracted only participants from one country. The Portuguese sample was chosen because it had the most balanced number of participants from both genders ($N = 509$; M age = 40 years, SD = 9.7 years; 44.8% women). Two samples were extracted: random sample 1 ($N = 260$; M age = 39.9 years, SD years = 9.4; 44.2% women) and random sample 2 ($N = 249$; M age = 40.1 years, SD 10.0 years; 45.4% women) (full details on demographics available on supplementary file 1: https://osf.io/vfmgq/).

The original model (see model in supplementary file 3: https://osf.io/hs5aj/) was first tested in the random sample 1 and presented poor fit to the data: $\chi^2(98) = 455.310$; $p < 0.001$; $\chi^2/df = 4.646$; CFI = 0.857; GFI = 0.822; SRMR = 0.117; RMSEA = 0.119 (LL = 0.108; UL = 0.130). Modification indices for the regression weights revealed that Item 10 ("To be slim so to look attractive to others") cross-loaded on *Social Recognition* (Item10 ← Social R.; MI = 58.439), suggesting that the re-specification of the model by associating Item 10 also to the *Social Recognition* factor would improve model fit. Modification indices also showed evidence of misspecification associated with errors variances of Item 9 ("To improve my overall health") and Item 1 ("To increase my resistance to illness and disease") that could reflect some degree of overlap in item content [39], as both belong to *Health Management* factor (err1 ↔ err9; MI = 27.203).

To proceed with a second analysis, we included two additional parameters: (1) the error covariance parameter between Item 9 and Item 1, and (2) specification of a cross-loading path that allowed Item 10 to load also on *Social Recognition* factor. When accounting for these changes in the model a review of the goodness-of-fit indices revealed better model fit ($\chi^2[96] = 318.359$; $p < 0.001$; $\chi^2/df = 3.316$; CFI = 0.911; GFI = 0.871; SRMR = 0.098; RMSEA = 0.095 [LL = 0.083; UL = 0.106]), still with room for improvement based on modification indices. There was evidence of item content overlap as suggested by the covariation of the error variance from Item 16 ("To measure myself against personal standards"), originally belonging to the *Challenge* intrinsic factor, with the error variance of the extrinsic items 14 (err16 ↔ err14; MI = 22.733) and 15 (err16 ↔ err15; MI = 30.431). There was also cross-loading of Item 16 on two theoretically incongruent extrinsic factors (Item16 ← Social R.;

MI = 22.446; Item16 ← Image; MI = 39.853). Further, this item revealed a low factor loading ($\lambda = 0.47$). Finally, Item 3 ("I manage my weight to be well thought of by others") cross-loaded on *Image* factor (Item3 ← Image MI = 14.146).

A second iteration was specified removing Item 16 and by specifying an additional parameter to allow the Item 3 to load also on *Image* factor. With the suggested changes (see each updated model iteration in supplementary file 3: https://osf.io/hs5aj/), we achieved an acceptable model fit with no further justification for additional specifications: $\chi^2(81) = 194.619$; $p < 0.001$; $\chi^2/df = 2.403$; CFI = 0.952; GFI = 0.917; SRMR = 0.070; RMSEA = 0.074 (LL = 0.060; UL = 0.087).

To confirm model adjustments, a second and independent analysis was conducted with the remaining half of the sample of the selected country ($N = 260$; M age = 39.90; SD = 9.37; 44% women). The analysis in the new sub-sample revealed a marginal fit of the new hypothesized model ($\chi^2(81) = 216.811$; $p < 0.001$; $\chi^2/df = 2.677$; CFI = 0.937; GFI = 0.897; SRMR = 0.065; RMSEA = 0.082 [LL = 0.069; UL = 0.096]). Inspection of the modification indices indicated misspecification associated with error variances related to Items 3 and 10. Despite these items are both extrinsic, they were originally hypothesized as belonging to different factors. Therefore, we decided to delete Item 3 and 10 due to recurrent cross-loading issues. The final model (see updated model iteration in supplementary file 3: https://osf.io/hs5aj/) presented reasonable fit to the data: ($\chi^2(58) = 145.031$; $p < 0.001$; $\chi^2/df = 2.501$; CFI = 0.949; GFI = 0.921; SRMR = 0.069; RMSEA = 0.078 [LL = 0.062; UL = 0.094]). Given the non-threatening modification indices, we saw no rational need for further model specification considering this the most parsimonious model to represent the data.

For the final step of factorial validity inspection (Table 1), we conducted a CFA on the final model (Fig. 1) with the full sample of 1511 participants.

The new factorial structure presented acceptable model fit to the data: $\chi^2(58) = 599.982$; $p < 0.001$; $\chi^2/df = 10.345$; CFI = 0.940; GFI = 0.941; SRMR = 0.063; RMSEA = 0.079 (LL = 0.073; UL = 0.084). All items present meaningful factor loadings (Table 2; $\lambda \geq 0.5$).

Reliability and validity

Composite reliability and convergent validity indices are described in Table 3. The *Health Management* factor demonstrated convergent factorial validity with values slightly below the recommended level (average variance extracted < 0.50) [36]. The intrinsic goals (*Health Management* and *Challenge*) showed significant moderate positive inter-correlation (rho = 0.37; $p < 0.001$) as did the two

Table 1 Goodness-of-fit indices of the CFA model iteration process across samples

Sample	Model	Modifications	χ^2/df	CFI	GFI	SRMR	RMSEA [LL; UL]
Random 1	Original	–	4.646	0.857	0.822	0.117	0.119 [0.108; 0.130]
Random 1	1° iteration	i10 cross-loading	3.316	0.911	0.871	0.098	0.095 [0.083; 0.106]
Random 1	2° iteration	i16 removed; i3 cross-loading	2.403	0.952	0.917	0.070	0.074 [0.060; 0.087]
Random 2	2° iteration	–	2.677	0.937	0.897	0.065	0.082 [0.069; 0.096]
Random 2	3° iteration	i3 and i10 removed	2.501	0.949	0.921	0.069	0.078 [0.062; 0.094]
Full sample	3° iteration	–	10.345	0.940	0.941	0.063	0.079 [0.073; 0.084]

χ^2/df Chi-square/degrees of freedom, *CFI* comparative fit index, *GFI* goodness-of-fit index, *SRMR* standardized root mean square residual, *RMSEA* root mean square error approximation, *LL* lower limit, *UL* upper limit

Fig. 1 GCWMS refined model—confirmatory factor analysis. *Social R.* social recognition factor, *Health M.* health management factor
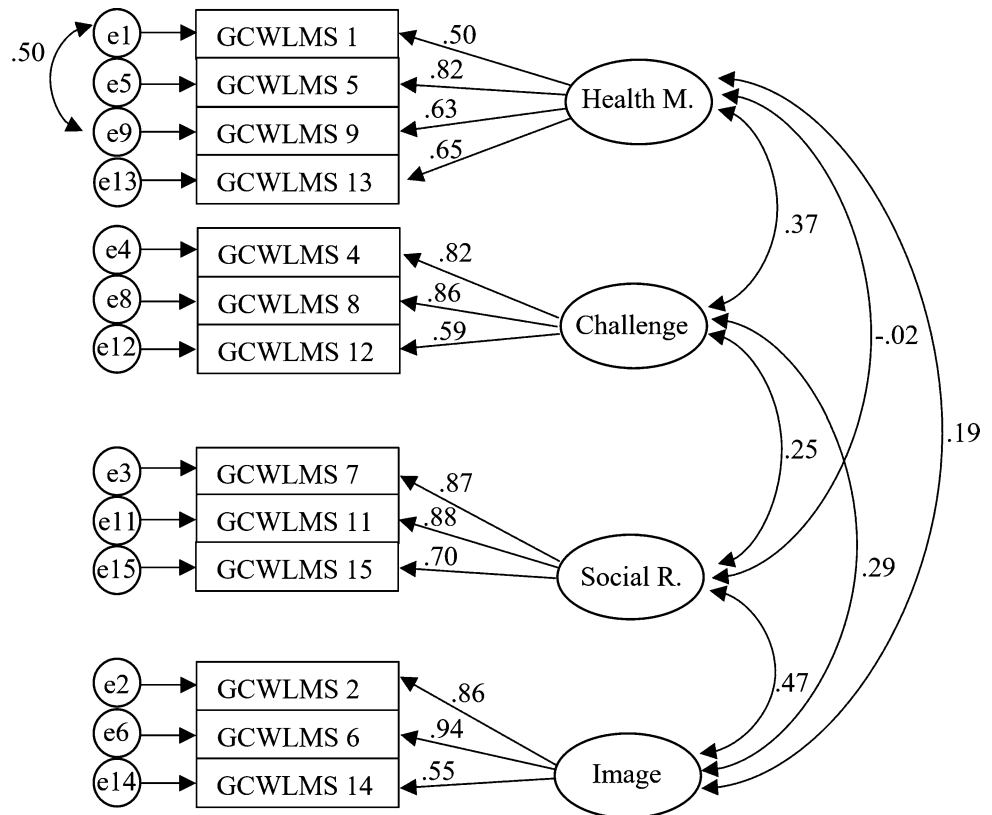
extrinsic goals, (*Image* and *Social Recognition*; rho = 0.47, *p* < 0.001). All other correlations are weak or null in accordance with theoretical predictions.

Discriminant validity was examined through the associations between each goal and the conceptually related constructs from other motivational based scales (Table 4). In line with theoretical assumptions, the *Health Management* factor presented weak to moderate positive correlations with the autonomous motivation (BREQ rho = 0.29; *p* < 0.001; REBS rho = 37; *p* < 0.001) and the two extrinsic goal factors were weakly to moderately positively associated with controlled motivation assessed by BREQ (*Image* rho = 0.22; *p* < 0.001; *Social Recognition* rho = 0.41; *p* < 0.001) and REBS (*Image* rho = 0.27; *p* < 0.001; *Social Recognition*

rho = 0.41; *p* < 0.001). The *Challenge* factor presented weak correlations both with autonomous and controlled motivation, albeit the latter presented higher values than expected (for full information please refer to supplementary file 4: https://osf.io/t5pkb/).

Multi-group invariance—gender

The measurement model invariance analysis results are presented in Table 5. The overall good fit of multi-group model of GCWMS (χ^2/df = 5.363; CFI = 0.944; GFI = 0.940; SRMR = 0.063; RMSEA = 0.054 [LL = 0.050; UL = 0.058]) confirms goodness-of-fit for the configural model across participant gender (i.e., the same number of factors are present

Table 2 Confirmatory factor analysis factor loadings

	Health	Challenge	Social	Image
1. To increase my resistance to illness and disease	0.50			
5. To increase my energy level	0.82			
9. To improve my overall health	0.63			
13. To improve my endurance, stamina	0.65			
4. To give me goals to work towards		0.82		
8. To give me personal challenges to face		0.86		
12. To develop personal skills		0.59		
7. To be socially respected by others			0.87	
11. To gain favourable approval from others			0.88	
15. So that others recognize me as a healthy person			0.70	
2. To improve the look of my overall body shape				0.86
6. To improve my appearance				0.94
14. To change my appearance by altering a specific area of my body				0.55

Table 3 Cronbach's alpha [diagonal], spearman's correlations [below diagonal], and factorial validity outcomes

GCWMS Factors		Health management	Challenge	Social recognition	Image	CR	AVE
Health M.	Spearman's rho	[0.775]	–	–	–	0.75	0.44
Challenge	Spearman's rho	0.374*	[0.790]	–	–	0.81	0.59
Social R.	Spearman's rho	− 0.024	0.249*	[0.854]	–	0.86	0.67
Image	Spearman's rho	0.186*	0.288*	0.470*	[0.783]	0.84	0.64

CR composite reliability, *AVE* average variance extracted

*$p < 0.001$, two-tailed

Table 4 Correlations between weight maintenance goal content and autonomous and controlled motivation

GCWMS Factors		Autonomous motivation		Controlled motivation	
		Exercise	Eating behaviour	Exercise	Eating behaviour
Health M.	Spearman's rho	0.287*	0.367*	0.031	0.011
Challenge	Spearman's rho	0.267*	0.240*	0.196*	0.172*
Social R.	Spearman's rho	− 0.034	0.041	0.413*	0.407*
Image	Spearman's rho	0.074*	0.102*	0.217*	0.269*

*$p < 0.001$, two-tailed

in each group and are explained by the same set of items). For metric invariance, goodness-of-fit results also showed evidence of well-fitting model with factor loadings specified to be invariant ($\chi^2/df = 5.184$; CFI = 0.942; SRMR = 0.062; RMSEA = 0.053). The difference in the Chi-square was significant but changes in fit indices were all below the suggested cut-off values (ΔCFI = − 0.002; ΔSRMR = − 0.001; ΔRMSEA = 0.001), supporting the conclusion that factor loadings were operating similarly across men and women.

Next, we tested scalar invariance with variable intercepts and factorial loadings constrained to be equivalent across groups (Model 3). This model presented good fit to the data ($\chi^2/df = 4.914$; CFI = 0.942; SRMR = 0.062;

RMSEA = 0.051). The test of Chi-square-difference between models was significant but the fit indices differences met the cut-off criteria for suggested fit indices (ΔCFI = 0.000; ΔSRMR = 0.000; ΔRMSEA = − 0.002). Using the delta fit indices model comparison criteria, we are confident in asserting that GCWMS achieved scalar invariance for gender comparison, confirming that there is no differential scoring on each factor's items between men and women.

Final step was to test for residual invariance by restraining the item residuals to be equivalent between groups. The overall fit indices suggested acceptable fit to the data ($\chi^2/df = 4.985$; CFI = 0.935; SRMR = 0.062; RMSEA = 0.051) and differences when comparing to the

Table 5 Scale invariance analysis showing fit statistics for the unconstrained model versus the constrained models [country and gender]

Model	χ^2 (df)	χ^2/df	CFI	ΔCFI	SRMR	ΔSRMR	RMSEA	ΔRMSEA
Women	186.999 (58)	3.224	0.955	–	0.064	–	0.068	–
Male	435.063 (58)	7.501	0.939	–	0.063	–	0.079	–
M1: configural invariance	622.051 (116)	5.363	0.944	–	0.063	–	0.054	–
M2: metric invariance	648.038 (125)	5.184	0.942	− 0.002	0.062	− 0.001	0.053	0.001
M3: scalar invariance	663.392 (135)	4.914	0.942	0.000	0.062	0.000	0.051	− 0.002
M4: residual invariance	742.751 (149)	4.985	0.935	− 0.007	0.062	0.000	0.051	0.000
UK	293.295 (58)	5.057	0.922	–	0.080	–	0.089	–
Portugal	238.376 (58)	4.110	0.950	–	0.067	–	0.078	–
Denmark	257.160 (58)	4.434	0.930	–	0.075	–	0.084	–
Portugal–UK								
M1: configural invariance	531.671 (116)	4.583	0.937	–	0.067	–	0.059	–
M2: metric invariance	569.916 (125)	4.559	0.933	− 0.004	0.071	0.004	0.059	0.000
M3: scalar invariance	591.282 (135)	4.380	0.931	− 0.002	0.076	0.005	0.057	− 0.002
M4: residual invariance	787.660 (149)	5.286	0.903	− 0.028	0.081	0.005	0.065	0.008
Portugal–Denmark								
M1: configural invariance	495.538 (116)	4.272	0.941	–	0.075	–	0.057	–
M2: metric invariance	535.256 (125)	4.282	0.936	− 0.005	0.078	0.003	0.057	0.000
M3: scalar invariance	561.222 (135)	4.157	0.934	− 0.002	0.083	0.005	0.056	− 0.001
M4: residual invariance	861.428 (149)	5.781	0.889	− 0.045	0.079	− 0.004	0.069	0.013
UK–Denmark								
M1: configural invariance	550.454 (116)	4.745	0.926	–	0.075	–	0.061	–
M2: metric invariance	576.705 (125)	4.614	0.923	− 0.003	0.078	0.003	0.060	− 0.001
M3: scalar invariance	610.533 (135)	4.522	0.919	− 0.004	0.087	0.009	0.059	− 0.001
M4: residual invariance	660.382 (149)	4.432	0.913	− 0.006	0.088	0.001	0.059	0.000

χ^2 Chi-square, *df* degrees of freedom, *CFI* comparative fit index, *SRMR* standardized root mean square residual, *RMSEA* root mean square error approximation

scalar invariance model (Model 3) fit indices were all below the proposed cut-off for residual invariance (ΔCFI = − 0.007; ΔSRMR = 0.000; ΔRMSEA = 0.000). In line with proposed procedures and guidelines [32, 45, 47, 52], we have statistical evidence to assume that the GCWMS final model achieved full factorial invariance regarding gender.

Multi-group invariance—country

The individual country model fit indices and cross-cultural measurement invariance models' results are also compared in Table 5. Despite model fit for Denmark and for UK presented marginal fit, configural invariance (Model 1) presented good fit to the data for all country's comparisons (Portugal–UK: χ^2/df = 4.583; CFI = 0.937; SRMR = 0.067; RMSEA = 0.059; Portugal–Denmark: χ^2/df = 4.272; CFI = 0.941; SRMR = 0.075; RMSEA = 0.057; UK—Denmark: χ^2/df = 4.745; CFI = 0.926; SRMR = 0.075; RMSEA = 0.061) supporting evidence for the same organization of the constructs for all three countries. Metric invariance indices presented acceptable model fit according to goodness-of-fit criteria (Portugal–UK: χ^2/df = 4.559;

CFI = 0.933; SRMR = 0.071; RMSEA = 0.059; Portugal–Denmark: χ^2/df = 4.282; CFI = 0.936; SRMR = 0.078; RMSEA = 0.057; UK–Denmark: χ^2/df = 4.614; CFI = 0.923; SRMR = 0.078; RMSEA = 0.060). Albeit, the differences in all Chi-square tests were significant, changes in CFI, SRMR, and RMSEA are below the assumed cut-off values (Portugal–UK: ΔCFI = − 0.004; ΔSRMR = 0.004; ΔRMSEA = 0.000; Portugal–Denmark: ΔCFI = − 0.005; ΔSRMR = 0.003; ΔRMSEA = 0.000; UK–Denmark: ΔCFI = − 0.003; ΔSRMR = 0.003; ΔRMSEA = − 0.001) and following Cheung's and Rensvold's [45, 52] guidelines, there was statistical support for the assumption that GCWMS factors have the same meaning across the evaluated countries.

After restricting variable intercepts and factorial loadings concurrently to test scalar invariance, the multi-group models presented marginal acceptable fit to the data for all country's comparisons (Portugal–UK: χ^2/df = 4.380; CFI = 0.931; SRMR = 0.076; RMSEA = 0.057; Portugal–Denmark: χ^2/df = 4.157; CFI = 0.934; SRMR = 0.083; RMSEA = 0.056; UK–Denmark: χ^2/df = 4.522; CFI = 0.919; SRMR = 0.087; RMSEA = 0.059). The tests of Chi-square-difference

between models were all significant but the fit indices differences between Model 3 (scalar invariance) and Model 2 (metric invariance) were all below the suggested cut-off values (Portugal–UK: $\Delta CFI = -0.002$; $\Delta SRMR = 0.005$; $\Delta RMSEA = -0.002$; Portugal–Denmark: $\Delta CFI = -0.002$; $\Delta SRMR = 0.005$; $\Delta RMSEA = -0.001$; UK–Denmark: $\Delta CFI = -0.006$; $\Delta SRMR = 0.001$; $\Delta RMSEA = 0.000$). Following recommended procedures and guidelines [32, 45, 47, 52], we assume that constraining the item intercepts among groups did not significantly affect the model fit and thus supporting evidence for scalar invariance (i.e., mean differences in the hypothesized constructs encapsulate all mean differences in shared variance of the items).

Lastly, residual invariance was tested (Model 4) and presented a poorer fit to the data for all country's comparisons (Portugal–UK: $\chi^2/df = 5.286$; CFI = 0.903; SRMR = 0.081; RMSEA = 0.065; Portugal–Denmark: $\chi^2/df = 5.781$; CFI = 0.889; SRMR = 0.079; RMSEA = 0.069; UK–Denmark: $\chi^2/df = 4.432$; CFI = 0.913; SRMR = 0.088; RMSEA = 0.059). Also, chi-square test differences were significant and CFI difference to the previous model (Model 3) surpassed the cut-off value proposed (Portugal–UK: $\Delta CFI = -0.028$; Portugal–Denmark: $\Delta CFI = -0.045$). We have no robust statistical argument to defend residual invariance of the GCWMS.

Discussion

This study reported the development and factorial validation of the Goal Content for Weight Maintenance Scale. After model re-specifications following the rationale supporting the development of this scale and following Byrne's [39] and Brown's [51] recommendations, we decided to go further in our endeavour to test the validity of the measurement model. Therefore, using a systematic theory-informed methodology, we observed a good fit of the factorial structure to the model and good internal consistency of all subscales. According to Self-Determination Theory, the essence of goal content is a core feature for sustained health behaviour and weight loss maintenance motivation [26]. Examining the extrinsic or intrinsic nature of goals to engage in the targeted behaviour for sustained weight-loss maintenance may help practitioners and researchers understand the quality, sustainability, and likely effectiveness of people's motivation. However, advancement in this field of knowledge has been hampered by the lack of a validated psychometric instruments to measure the content of people's weight loss maintenance goals.

Sixteen items derived from measures of goal content for exercise were specified and reviewed by an expert panel to assess four goals: *Health Management*; *Challenge*; *Image*; and *Social Recognition*. Initial CFA analysis suggested theoretically consistent and pragmatic scale modifications. Based on modification indices, *Image* factor's item 10 ("To be slim so to look attractive to others") cross-loaded on *Social Recognition*, which is plausible given the item's reference to other people's perceptions. Item's 1 ("To increase my resistance to illness and disease") and item's 9 ("To improve my overall health") error variances presented a possible covariance between each other indicating content overlap.

Furthering the analysis, modification indices associated with item 16 ("To measure myself against personal standards") provided evidence of cross-loading issues with extrinsic factors (*Image* and *Social Recognition*) that were incongruent with the SDT framework. Indeed, the underlying competitive nature of this item's content may be misleading and therefore induce social comparison of one's abilities (extrinsic). Due to this incongruency, we deleted item 16. Additionally, item 3 ("To be well thought of by others") cross-loaded on *Image* factor and presented a misspecification with item 10 error variance. Item content may be misrepresented as a notion of social image instead of the social identification of one's identity. Subjecting items 1, 3, 9, 10 and 16 to further refinement by a panel of specialists with face validity analysis among people trying to maintain their weight loss is recommended. Furthermore, we suggest investigation of the potential validity and reliability of shorter version of GCWMS with 12 or even 8 items, which would be suitable for increasingly used digital platforms that allow faster and simpler assessments (e.g., in apps or web-based platforms instead of paper and pencil).

In line with previous research using the GCEQ (the scale on which the GCWMS was based), our analysis confirmed that goal content subscales associations were aligned with the proposed motivational model, that is, extrinsic goals were inter-correlated as were intrinsic goals. Also as expected, extrinsic goals (*Image* and *Social Recognition*) were positively correlated with controlled motivation for exercise. The *Health Management* goal, a hypothesised intrinsic goal, correlated positively with autonomous motivation. In contrast, the challenge-oriented goals correlated positively with both autonomous and controlled motivation. One reason for this might be different interpretations of some items in this subscale (e.g., "To give me goals to work towards"; "To give me personal challenges to face") as either personal skill development or as an outcome-oriented pursuit. It would be of interest to verify this correlation pattern in future studies. Future research could also refine this factor's items to capture a sense of self-development by focusing on the process rather that a goal-oriented mindset, e.g.: "To feel competent and in control; "To overcome my current difficulties".

Measurement invariance tests confirmed full factorial invariance of the GCWMS across gender. Residual invariance was not achieved across countries. Although important for full factorial invariance, residual invariance

is not a prerequisite for testing mean differences [53]. Therefore, based on these results, the scale scores can be compared across gender and the three analysed countries (Portugal, UK, and Denmark) as the underlying latent variables were interpreted in the same way by all individuals. We further demonstrated that the factor structure of the GCWMS is consistent across the studied groups and we may assume an unbiased scoring of the scale's items. We are confident that further refinements may improve scale performance to achieve full factorial validity also across countries.

Initial poor model fit indices of the GCWMS may be explained in terms of its derivation from two other scales that were originally designed to assess goal contents for exercise. Future item re-specification to address weight loss maintenance-related goals might improve the scale (possibly based in qualitative research considering people's experiences). However, identifying latent constructs that focus on more distal outcomes (weight) rather than tangible behaviours (such as exercising) could be challenging.

Strengths and limitations

The findings are supported by advanced multivariate analysis allowing the model to be tested in a large sample of individuals engaged in a weight loss maintenance intervention. The scale was also thoroughly developed based on previous original work [29, 38]. The use of a large data set derived from three European countries allowed testing of gender and cultural measurement invariance. The data were collected as part of baseline measures of a large controlled trial, providing confidence in the quality of the recruitment and data collection procedures.

Despite the advantages of this study being nested within a controlled trial, this also may limit the generalisability of the developed scale to the overall population as participants in the NoHoW trial had successfully lost ≥ 5% weight in the last 12 months and may be more motivated in their weight management efforts than the general population. Additionally, the absence of residual invariance across countries may be a result of measurement errors related to the complex process of translation to different languages and cultural backgrounds [32].

Future studies should refine GCWMS items that were removed due to cross-loading and items with error variances that were found to co-vary suggesting content overlap. Indeed, new more specific weight management items may be formulated to improve scale's performance. A qualitative method approach with interviews may be of substantial usefulness for item refinement.

What is already known on this subject?

The literature that aims to provide a motivational viewpoint on regulation of eating behaviour is limited and does not pertain adequate attention to the motivational dynamics involved in the weight maintenance behavioural processes. To the best of our knowledge, there are no scales available to assess the nature of people's goals for weight loss maintenance.

What your study adds?

This study provides initial evidence for the validity and reliability of scores derived from Goal Content for Weight Maintenance Scale and confirms measurement invariance across gender and three European countries. Findings support that GCWMS can be used to measure accurately the content of weight loss maintenance goals, establishing its usefulness in the endeavour of advancing measurement of theory-based aspects of motivation for weight loss maintenance that will enable a better understanding of what contributes, in motivational terms, to weight loss maintenance and why.

Compliance with ethical standards

Conflict of interest On behalf of all authors, the corresponding author states that there is no conflict of interest. R. James Stubbs consults for Slimming World through the University of Leeds. Marta Marques also conducted consultancy work with Slimming World. Slimming World were a partner of the project to aid recruitment within the UK but will not be involved in the analysis and interpretation of the trial.

Ethical approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the Institutional Ethics Committees at the Universities of Leeds (17-0082; 27 February 2017), Lisbon (17/2016; 20 February 2017) and Capital Region of Denmark (H-16030495, 8 March 2017) and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

References

1. Stein CJ, Colditz GA (2004) The epidemic of obesity. J Clin Endocrinol Metab 89(6):2522–2525. https://doi.org/10.1210/jc.2004-0288

2. Field AE, Coakley EH, Must A, Spadano JL, Laird N, Dietz WH et al (2001) Impact of overweight on the risk of developing common chronic diseases during a 10-year period. Arch Intern Med 161(13):1581–1586. https://doi.org/10.1001/archinte.161.13.1581

3. Calle EE, Rodriguez C (2003) Overweight, obesity, and mortality from cancer in a prospectively studied cohort of US adults. N Engl J Med 348:1625–1638. https://doi.org/10.1056/NEJMoa021423

4. Ulen CG, Huizinga MM, Beech B, Elasy TA (2008) Weight regain prevention. Clin Diabetes 26(3):100–113. https://doi.org/10.2337/diaclin.26.3.100

5. Wing RR, Phelan S (2005) Long-term weight loss maintenance. Am J Clin Nutr 82(1):222S-225S. https://doi.org/10.1093/ajcn/82.1.222S

6. Curioni CC, Lourenço PM (2005) Long-term weight loss after diet and exercise: a systematic review. Int J Obes 29(10):1168–1174. https://doi.org/10.1038/sj.ijo.0803015

7. Santos I, Vieira PN, Silva MN, Sardinha LB, Teixeira PJ (2017) Weight control behaviors of highly successful weight loss maintainers: the Portuguese Weight Control Registry. J Behav Med 40(2):366–371. https://doi.org/10.1007/s10865-016-9786-y

8. Stubbs RJ, Lavin JH (2013) The challenges of implementing behaviour changes that lead to sustained weight management. Nutr Bull 38(1):5–22. https://doi.org/10.1111/nbu.12002

9. Samdal GB, Eide GE, Barth T, Williams G, Meland E (2017) Effective behaviour change techniques for physical activity and healthy eating in overweight and obese adults; systematic review and meta-regression analyses. Int J Behav Nutr Phys Act 14(1):42. https://doi.org/10.1186/s12966-017-0494-y

10. Dombrowski SU, Knittle K, Avenell A, Araújo-Soares V, Sniehotta FF (2014) Long term maintenance of weight loss with non-surgical interventions in obese adults: systematic review and meta-analyses of randomised controlled trials. BMJ 348:g2646. https://doi.org/10.1136/bmj.g2646

11. Olander EK, Fletcher H, Williams S, Atkinson L, Turner A, French DP (2013) What are the most effective techniques in changing obese individuals' physical activity self-efficacy and behaviour: a systematic review and meta-analysis. Int J Behav Nutr Phys Act 10:29. https://doi.org/10.1186/1479-5868-10-29

12. Prestwich A, Sniehotta FF, Whittington C, Dombrowski SU, Rogers L, Michie S (2014) Does theory influence the effectiveness of health behavior interventions? Meta-analysis. Health Psychol 33(5):465–474. https://doi.org/10.1037/a0032853

13. Gourlan M, Bernard P, Bortolon C, Romain AJ, Lareyre O, Carayol M et al (2016) Efficacy of theory-based interventions to promote physical activity. A meta-analysis of randomised controlled trials. Health Psychol Rev 10(1):50–66. https://doi.org/10.1080/17437199.2014.981777

14. Teixeira PJ, Marques MM (2017) Health behavior change for obesity management. Obes Facts 10(6):666–673. https://doi.org/10.1159/000484933

15. Varkevisser RDM, van Stralen MM, Kroeze W, Ket JCF, Steenhuis IHM (2019) Determinants of weight loss maintenance: a systematic review. Obes Rev 20(2):171–211. https://doi.org/10.1111/obr.12772

16. Teixeira PJ, Carraça EV, Marques MM, Rutter H, Oppert J-M, De Bourdeaudhuij I et al (2015) Successful behavior change in obesity interventions in adults: a systematic review of self-regulation mediators. BMC Med 13:84. https://doi.org/10.1186/s12916-015-0323-6

17. Teixeira PJ, Silva MN, Mata J, Palmeira AL, Markland D (2012) Motivation, self-determination, and long-term weight control. Int J Behav Nutr Phys Act 9(1):22. https://doi.org/10.1186/1479-5868-9-22

18. Teixeira PJ, Carraça EV, Markland D, Silva MN, Ryan RM (2012) Exercise, physical activity, and self-determination theory: a systematic review. Int J Behav Nutr Phys Act 9:78. https://doi.org/10.1186/1479-5868-9-78

19. Teixeira PJ, Palmeira AL, Vansteenkiste M (2012) The role of self-determination theory and motivational interviewing in behavioral nutrition, physical activity, and health: an introduction to the IJBNPA special series. Int J Behav Nutr Phys Act 9(1):17. https://doi.org/10.1186/1479-5868-9-17

20. Silva MN, Vieira PN, Coutinho SR, Minderico CS, Matos MG, Sardinha LB et al (2010) Using self-determination theory to promote physical activity and weight control: a randomized controlled trial in women. J Behav Med 33(2):110–122. https://doi.org/10.1007/s10865-009-9239-y

21. Silva MN, Markland D, Carraça EV, Vieira PN, Coutinho SR, Minderico CS et al (2011) Exercise autonomous motivation predicts 3-yr weight loss in women. Med Sci Sports Exerc 43(4):728–737. https://doi.org/10.1249/MSS.0b013e3181f3818f

22. Deci EL, Ryan RM (2000) The "what" and "why" of goal pursuits: Human needs and the self-determination of behavior. Psychol Inq 11(4):227–268. https://doi.org/10.1207/S15327965PLI1104_01

23. Sheldon KM, Krieger LS (2004) Does legal education have undermining effects on law students? Evaluating changes in motivation, values, and well-being. Behav Sci Law 22(2):261–286. https://doi.org/10.1002/bsl.582

24. Vansteenkiste M, Lens W, Deci EL (2006) Intrinsic versus extrinsic goal contents in self-determination theory: another look at the quality of academic motivation. Educ Psychol 41(1):19–31. https://doi.org/10.1207/s15326985ep4101_4

25. Vansteenkiste M, Neyrinck B, Niemiec CP, Soenens B, De Witte H, Van den Broeck A (2007) On the relations among work value orientations, psychological need satisfaction and job outcomes: a self-determination theory approach. J Occup Organ Psychol 80(2):251–277. https://doi.org/10.1348/096317906X111024

26. Ryan R, Deci E (2017) Self-determination theory: basic psychological needs in motivation, development, and wellness. Guilford Publications, New York (**ISBN: 9781462528783**)

27. Lindwall M, Weman-Josefsson K, Sebire SJ, Standage M (2016) Viewing exercise goal content through a person-oriented lens: a self-determination perspective. Psychol Sport Exerc 27:85–92. https://doi.org/10.1016/j.psychsport.2016.06.011

28. Sheldon KM, Kasser T (1995) Coherence and congruence: two aspects of personality integration. J Pers Soc Psychol 68(3):531–543. https://doi.org/10.1037/0022-3514.68.3.531

29. Sebire SJ, Standage M, Vansteenkiste M (2008) Development and validation of the Goal Content for Exercise Questionnaire. J Sport Exerc Psychol 30:353–377. https://doi.org/10.1123/jsep.30.4.353

30. Sue S (1999) Science, ethnicity, and bias: where have we gone wrong? Am Psychol 54(12):1070–1077. https://doi.org/10.1037/0003-066x.54.12.1070

31. Millsap RE, Kwok O-M (2004) Evaluating the impact of partial factorial invariance on selection in two populations. Psychol Methods 9(1):93–115. https://doi.org/10.1037/1082-989X.9.1.93

32. Fischer R, Karl JA (2019) A primer to (cross-cultural) multi-group invariance testing possibilities in R. Front Psychol 10:1507. https://doi.org/10.3389/fpsyg.2019.01507

33. Hoyle RH, Smith GT (1994) Formulating clinical research hypotheses as structural equation models: a conceptual overview. J Consult Clin Psychol 62(3):429–440. https://doi.org/10.1037//0022-006x.62.3.429

34. Weinberger N-A, Kersting A, Riedel-Heller SG, Luck-Sikorski C (2016) Body dissatisfaction in individuals with obesity compared to normal-weight individuals: a systematic review and meta-analysis. Obes Facts 9(6):424–441. https://doi.org/10.1159/000454837

35. Scott SE, Duarte C, Encantado J, Evans EH, Harjumaa M, Heitmann BL et al (2019) The NoHoW protocol: a multicentre 2×2 factorial randomised controlled trial investigating an evidence-based digital toolkit for weight loss maintenance in European adults. BMJ Open 9(9):e029425. https://doi.org/10.1136/bmjopen-2019-029425

36. Hair JF, Anderson RE, Babin BJ, Black WC (2010) Multivariate data analysis: a global perspective, 7th edn. Prentice Hall, Upper Saddle River (**ISBN: 9780138132637**)

37. Kline RB (2015) Principles and practice of structural equation modeling, 4th edn. Guilford Publications, New York (**ISBN: 9781462523344**)

38. Markland D, Ingledew DK (1997) The measurement of exercise motives: factorial validity and invariance across gender of a revised Exercise Motivations Inventory. Br J Health Psychol 2(4):361–376. https://doi.org/10.1111/j.2044-8287.1997.tb00549.x

39. Byrne BM (2016) Structural equation modeling with AMOS: Basic concepts, applications, and programming, 3rd edn. Routledge, London (**ISBN: 9781138797031**)

40. Nevitt J, Hancock GR (2001) Performance of bootstrapping approaches to model test statistics and parameter standard error estimation in structural equation modeling. Struct Equ Model 8(3):353–377. https://doi.org/10.1207/S15328007SEM0803_2

41. Arbuckle J (2010) IBM SPSS Amos 19 user's guide. Amos Development Corporation, Crawfordville

42. Hair J, Anderson R, Black B, Babin B (2016) Multivariate data analysis. Pearson Education, London (**ISBN: 9780133792683**)

43. Schumacker RE, Lomax RG (2015) A beginner's guide to structural equation modeling, 4th edn. Routledge, London (**ISBN: 9781317608097**)

44. Hair JF, Hult GTM, Ringle CM, Sarstedt M (2014) A primer on partial least squares structural equation modeling. Sage, Thousand Oaks (**ISBN: 9781452217444**)

45. Cheung GW, Rensvold RB (2002) Evaluating goodness-of-fit indexes for testing measurement invariance. Struct Equ Model 9(2):233–255. https://doi.org/10.1207/S15328007SEM0902_5

46. Chen FF (2007) Sensitivity of goodness of fit indexes to lack of measurement invariance. Struct Equ Model 14(3):464–504. https://doi.org/10.1080/10705510701301834

47. Putnick DL, Bornstein MH (2016) Measurement invariance conventions and reporting: the state of the art and future directions for psychological research. Dev Rev 41:71–90. https://doi.org/10.1016/j.dr.2016.06.004

48. Markland D, Tobin V (2004) A modification to the behavioural regulation in exercise questionnaire to include an assessment of amotivation. J Appl Sport Psychol 26(2):191–196. https://doi.org/10.1123/jsep.26.2.191

49. Wilson PM, Rodgers WM, Loitz CC, Scime G (2007) "It's Who I Am … Really!'' The importance of integrated regulation in exercise contexts1." J Appl Biobehav Res 11:79–104. https://doi.org/10.1111/j.1751-9861.2006.tb00021.x

50. Pelletier LG, Dion SC, Slovinec-D'Angelo M, Reid R (2004) Why do you regulate what you eat? Relationships between forms of regulation, eating behaviors, sustained dietary behavior change, and psychological adjustment. Motiv Emot 28(3):245. https://doi.org/10.1023/B:MOEM.0000040154.40922.14

51. Brown TA (2015) Confirmatory factor analysis for applied research, 2nd edn. Guilford Publications, New York (**ISBN: 9781462515363**)

52. Cheung GW, Rensvold RB (2000) Assessing extreme and acquiescence response sets in cross-cultural research using structural equations modeling. J Cross Cult Psychol 31(2):187–212. https://doi.org/10.1177/0022022100031002003

53. Vandenberg RJ, Lance CE (2000) A review and synthesis of the measurement invariance literature: suggestions, practices, and recommendations for organizational research. Organ Res Methods 3:4–70. https://doi.org/10.1177/109442810031002