Heliyon 9 (2023) e22583

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Differential item functioning for the Tendency of Avoiding Physical Activity and Sport Scale across two subculture samples: Taiwanese and mainland Chinese university students

Chia-Wei Fan [a], Po-Ching Huang [b], I-Hua Chen [c], Yu-Ting Huang [d], Jung-Sheng Chen [e], Xavier C.C. Fung [f], Ji-Kang Chen [g], Yung-Ning Yang [h,**], Kerry S. O'Brien [i], Chung-Ying Lin [d,j,k,l,m,*], Mark D. Griffiths [n]

[a] *Department of Occupational Therapy, AdventHealth University, Orlando, FL, USA*
[b] *School of Physical Therapy and Graduate Institute of Rehabilitation Science, College of Medicine, Chang Gung University, Taoyuan, 333, Taiwan*
[c] *Chinese Academy of Education Big Data, Qufu Normal University, Shandong, China*
[d] *Institute of Allied Health Sciences, College of Medicine, National Cheng Kung University, Tainan, 701, Taiwan*
[e] *Department of Medical Research, E-Da Hospital, I-Shou University, Kaohsiung, 824, Taiwan*
[f] *Department of Rehabilitation Sciences, The Hong Kong Polytechnic University, Hong Kong*
[g] *Department of Social Work, Chinese University of Hong Kong, Hong Kong*
[h] *Department of Pediatrics, E-Da Hospital, School of Medicine, I-Shou University, Kaohsiung, Taiwan*
[i] *Behavioural Sciences Research Lab, Monash University, Australia*
[j] *University of Religions and Denominations, Qom, Iran*
[k] *Biostatistics Consulting Center, National Cheng Kung University Hospital, College of Medicine, National Cheng Kung University, Tainan, 701, Taiwan*
[l] *Department of Public Health, College of Medicine, National Cheng Kung University, Tainan, 701, Taiwan*
[m] *Department of Occupational Therapy, College of Medicine, National Cheng Kung University, Tainan, 701, Taiwan*
[n] *Psychology Department, Nottingham Trent University, Nottingham, NG1 4BU, UK*

ARTICLE INFO

Keywords:
Physical activity
Psychometrics
Rasch analysis
Stigma
Weight

ABSTRACT

Aims: The aims of the study were to examine the differential item functioning (DIF) of the Tendency of Avoiding Physical Activity and Sport Scale (TAPAS) among three subgroups (gender, weight status, and region) and to test the construct and concurrent validities of the scale.
Methods: Using an online survey, university students (608 Taiwanese and 2319 mainland Chinese) completed the TAPAS. Rasch analysis examined if all the 10 TAPAS items fitted the same construct and displayed no substantial DIF across three subgroups: gender (male vs. female), weight status (overweight vs. non-overweight), and region (Taiwan vs. China). Concurrent validity was examined using the scores on the Weight Self-Stigma Questionnaire (WSSQ) and Weight Bias Internalization Scale (WBIS).

* Corresponding author. Institute of Allied Health Sciences and Departments of Occupational Therapy and Public Health, National Cheng Kung University Hospital, College of Medicine, National Cheng Kung University, 1 University Rd, Tainan 701, Taiwan.
** Corresponding author. Department of Pediatrics, E-Da Hospital, School of Medicine, I-Shou University, Kaohsiung, Taiwan.
E-mail addresses: chia-wei.fan@ahu.edu (C.-W. Fan), hh780705@hotmail.com (P.-C. Huang), ahole.chen@gmail.com (I.-H. Chen), tnyuting@gmail.com (Y.-T. Huang), ed113118@edah.org.tw (J.-S. Chen), xavierfungzzz@gmail.com (X.C.C. Fung), jkchen@swk.cuhk.edu.hk (J.-K. Chen), ancaly@yahoo.com.tw (Y.-N. Yang), Kerry.O'Brien@monash.edu (K.S. O'Brien), cylin36933@gmail.com (C.-Y. Lin), mark.griffiths@ntu.ac.uk (M.D. Griffiths).

https://doi.org/10.1016/j.heliyon.2023.e22583
Received 23 July 2023; Received in revised form 14 November 2023; Accepted 15 November 2023
Available online 19 November 2023

Results: All TAPAS items, except for Item 10 (*"Prefer to participate in physical activity in a more private setting"*), fitted the same construct. None of the TAPAS items displayed DIF in any of the subgroups except for Item 10 across participants from Taiwan and China (DIF contrast = −1.41). *Conclusion:* The TAPAS can appropriately assess the tendency to avoid physical activity and sport among both Taiwanese and mainland Chinese university students. However, Item 10 may need to be further examined.

1. Introduction

1.1. Background/Rationale

According to the latest World Health Organization (WHO) estimates, more than 1.9 billion adults worldwide were overweight and of these, over 650 million were obese [1]. Prior studies have reported that individuals who are overweight or obese often experience weight-related stigma and discrimination [2–4]. Moreover, body image has been shown to play a critical role in mediating the relationships between sport involvement and the level of activity [2]. Those individuals with concerns about being overweight often feel embarrassed to show their unfit body to others while engaging in physical activity [5] and fear being judged [6]. More specifically, this type of weight-related discrimination is mostly experienced by adolescents [7], university students [8], and those in mid-adulthood [9]. Qualitative interview studies have also reported that some individuals exclude themselves from participating in sports or physical activities due to traumatic weight-related stigma and self-discrimination [10,11].

The Tendency to Avoid Physical Activity and Sport Scale (TAPAS) is a newly developed psychometric instrument that assesses the role of weight stigma and appearance-related concerns in physical activity and sports participation [12]. The initial feasibility study and other psychometric examinations have shown promising results [10,12]. However, to the authors' best knowledge, no previous studies have examined measurement bias of the TAPAS. Measurement bias is critical in relation to outcome validity and test fairness. In modern test theory, measurement invariance is assessed at the item level using differential item functioning (DIF) evaluation.

DIF refers to the presence of bias in test items. More specifically, it occurs when examining its psychometric properties across different subgroups with similar ability levels. In other words, although the participants in the subgroups share the same ability levels on the testing items, they may experience different probabilities of correctly answering items due to the underlying latent variables possessed by the individual subgroup (e.g., males and females have many potentially different inherent characteristics; individuals living in different countries have different cultural backgrounds; and individuals with different weight status have different feelings toward their body image). Given the inconsistency of probabilities, the presence of DIF in measurements can skew and invalidate the outcomes of clinical studies. Although researchers have used DIF since the early 1960s [13], the underlying causes of DIF are varied for each individual assessment. Therefore, every newly developed assessment should undergo DIF analysis to further examine the potential item bias [14,15]. Researchers have also indicated that clinicians and other professionals should be familiar with the potential underlying latent variables that might cause DIF before conducting such evaluation [16].

Over the past few decades, different approaches have been developed to examine DIF [17–19]. A recent study compared accuracy and effectiveness of five DIF detection methods in polytomous items [20]. The results showed that item response theory models, such as the Wald test and the likelihood ratio test outperformed traditional approaches (including the Mantel-Haenszel method and logistic regression) in terms of sensitivity and specificity [20]. Additionally, classical test theory approaches in detecting DIF are sample-specific, and therefore lack evidence regarding measurement invariance [21,22]. Moreover, in classical test theory, equal measurement precision has been assumed for all participants irrespective of their ability levels, which might not be true for all subgroups.

On the other hand, in Rasch analysis, measurement precision varies based on the individual ability levels [23]. A Rasch model can only be utilized when the test items are based on a single underlying trait (i.e., all items are unidimensional). Once this requirement is met, the person and item logit estimates produced from the Rasch analysis are sample-free and item-free. In other words, the item-difficulty calibrations from the test are independent from the participants that were used in the estimation (i.e., 2927 participants in the present study). In addition, the person-ability calibrations of the participants are independent from the used test items of the selected psychometric scale (i.e., TAPAS in the present study) [24]. Given the assumptions of Rasch analysis, the results retrieved from such analysis are sample-free and item-free and can be generalized beyond the enrolled participants [25]. Therefore, the findings will not be restricted to the studied population only. Furthermore, the DIF results using the Rasch approach will provide valuable insight into the potential item bias in TAPAS.

1.2. Objectives

The present study aimed to (i) identify potential bias of TAPAS across three subgroups (i.e., gender, region, and overweight) using DIF, (ii) examine the five-point Likert scale functioning of TAPAS items, (iii) verify the one-dimensionality construct of TAPAS, (iv) evaluate ceiling and floor effects of 10 TAPAS items, and (v) examine the concurrent validity of TAPAS with scores on two measures: Weight Self-Stigma Questionnaire (WSSQ) and Weight Bias Internalization Scale (WBIS). Based on the objectives, it was hypothesized that (i) there would be no DIF items in the TAPAS across the three subgroups; (ii) the five-point Likert scale used for each TAPAS item would have increased sequential monotonical difficulty (e.g., 'strongly agree' would be more difficult than 'agree', etc.); (iii) the ten-

items in the TAPAS would be unidimensional; (iv) there are no severe ceiling and floor effects for the ten TAPAS items; and (v) the TAPAS would have good concurrent validity with the external weight stigma measures (i.e., WSSQ and WBIS).

2. Methods

2.1. Study design

The study was an online cross-sectional survey of Taiwanese and Chinese university students. The study protocol was approved by the National Cheng Kung University Human Research Ethics Committee (Approval No. NCKU HREC-E−110-486-2), the National Cheng Kung University Hospital Institute of Review Board (IRB No. A-ER-111-445), and the Institutional Review Board of the Jiangxi Psychological Consultant Association (IRB ref: JXSXL- 2021- J99). The study was reported following the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) guidelines for reporting cross-sectional studies [26].

2.2. Procedure

An online survey link generated from the *SurveyMonkey* platform was distributed to university students with the assistance of four faculty members in several Taiwan universities located in Northern, Central, and Southern Taiwan. Similarly, the online survey link generated from the *Wenjuanxing* platform was distributed to university students with the assistance of 19 faculty members in several mainland Chinese universities across 13 provinces in mainland China. The target participants who clicked the link or scanned the QR code were presented with the online survey. The first page of the online survey described the information about the study and asked about the student's willingness to participate in the study. Only those participants who pressed the "agree" icon (to confirm informed consent) completed the survey. The survey webpage closed directly if informed consent was not provided. The data were collected from both Taiwanese and Chinese universities between August and December 2022.

2.3. Participants

The same inclusion criteria were applied to both Taiwanese and mainland Chinese participants. Participants were included if they (i) were legal adults under the Civil Law of each region to provide informed consent, (ii) were students in a college or university, (iii) had internet-enabled devices to access the survey, and (iv) provided informed consent.

2.4. Data sources/measurement

2.4.1. Tendency to Avoid Physical Activity and Sport Scale (TAPAS)

The TAPAS is a self-report scale that was developed to assess individuals' tendency to avoid physical activity or sport due to weight-related stigma and appearance concerns [12]. It has two subscales (i.e., physical activity and sport) with five items each. Items were designed for participants to self-report using a five-point scale from 1 (*strongly disagree*) to 5 (*strongly agree*). Higher scores indicate greater avoidance of participating in physical activity or sport. The Mandarin Chinese version of the TAPAS was used in the study, because Mandarin Chinese is the official language used in both Taiwan and China [27,28]. The scale demonstrated good validity among undergraduate students who are overweight [10]. The TAPAS items are shown in Table 2. The scale had high internal consistency in the present study (Cronbach's $\alpha = 0.99$).

Table 1
Demographics (N = 2927).

Variables	Taiwan (n = 608)	China (n = 2319)	t or χ^2
Age [M (S.D.)]	29.10 (6.36)	20.16 (1.75)	34.3**
BMI	23.88 (4.42)	25.30 (9.28)	5.3**
Weight Status			
Overweight	314 (51.64)	780 (33.64)	66.75**
Non-overweight	294 (48.36)	1539 (66.36)	
Gender [n (%)]			8.6*
Male	273 (44.90)	992 (42.77)	
Female	333 (54.77)	1327 (57.22)	
Others	2 (.33)	0 (0)	
Marital Status			324.5**
Single	491 (80.76)	2263 (97.59)	
Married	104 (17.11)	21 (.90)	
Divorced	10 (1.64)	18 (.70)	
Separated	1 (.16)	17 (.70)	
Unmarried but cohabitating	2 (.33)	0 (0)	

Note. N = Number; M = Mean; S.D. = Standard Deviation; BMI = Body Mass Index; $^{**}p < .001$; $^{*}p < .05$.

Table 2
Item average scores from Taiwan and China (N = 2927).

Variables	Taiwan (n = 608)	China (n = 2319)	*t*
TAPAS			
1. I find myself avoiding participating in sport because of my weight.	2.29	2.15	2.88*
2. I avoid participating in sport because of my fear of being judged about my lack of physical ability.	2.43	2.22	4.18**
3. I worry about participating in sport because I don't like how my body looks when playing sport.	2.33	2.23	1.89
4. I am afraid other people will notice my physical flaws when I participate in sport.	2.71	2.38	5.89**
5. I am concerned about what other people think of my appearance when I participate in sport.	2.86	2.45	7.73**
6. I avoid physical activity because I might get teased about my weight.	2.23	2.17	1.22
7. I avoid physical activity because of my fear of being judged about my physical appearance.	2.36	2.23	2.49*
8. I avoid physical activity because I worry that people may make negative comments about my body.	2.34	2.20	2.60*
9. I avoid physical activity because I worry people may be thinking negatively about my physical appearance.	2.36	2.22	2.74*
10. I would prefer to participate in physical activity in a more private setting.	3.45	2.65	14.86**
WBIS			
1. As an overweight person, I feel that I am just as competent as anyone	3.00	2.72	6.56**
2. I am less attractive than most other people because of my weight	2.92	2.40	10.29**
3. I feel anxious about being overweight because of what people might think of me	3.00	2.39	11.95**
4. I wish I could drastically change my weight	3.36	2.70	12.39**
5. Whenever I think a lot about being overweight, I feel depressed	2.82	2.38	8.57**
6. I hate myself for being overweight	2.93	2.37	10.62**
7. My weight is a major way that I judge my value as a person	1.95	1.95	−.07
8. I don't feel that I deserve to have a really fulfilling social life, as long as I'm overweight	1.96	2.00	−.97
9. I am OK being the weight that I am	3.03	3.13	−2.00*
10. Because I'm overweight, I don't feel like my true self	2.33	2.16	3.60**
11. Because of my weight, I don't understand how anyone attractive would want to date me	2.23	2.11	2.52*
WSSQ			
1. I'll always go back to being overweight	2.44	2.17	5.48**
2. I caused my weight problems	3.08	2.34	14.64**
3. I feel guilty because of my weight problems	2.62	2.21	8.11**
4. I became overweight because I'm a weak person	2.18	2.03	3.04*
5. I would never have any problems with weight if I were stronger	2.41	2.22	3.84**
6. I don't have enough self-control to maintain a healthy weight	2.83	2.35	9.09**
7. I feel insecure about others' opinions of me	2.92	2.48	8.41**
8. People discriminate against me because I've had weight problems	2.16	2.05	2.46*
9. It's difficult for people who haven't had weight problems to relate to me	2.28	2.09	3.91**
10. Others will think I lack self-control because of my weight problems	2.64	2.16	9.12**
11. People think that I am to blame for my weight problems	2.35	2.07	5.81**
12. Others are ashamed to be around me because of my weight	2.11	2.07	.83

Note. TAPAS = Tendency of Avoiding Physical Activity and Sport Scale; WBIS = Weight Bias Internalization Scale; WSSB = Weight Self-Stigma Questionnaire. $^{**}p < .001$; $^{*}p < .05$.

2.4.2. Weight Self-Stigma Questionnaire (WSSQ)

The WSSQ was used to assess to participants' weight-related self-stigma [29]. It contains two subscales (i.e., self-devaluation and fear of enacted stigma) with six items in each that are rated on a five-point Likert scale from 1 (*strongly disagree*) to 5 (*strongly agree*). Higher scores indicate a higher level of weight-related self-stigma. The Mandarin Chinese-translated WSSQ, which has robust psychometric properties [30], was used to collect data. The WSSQ items are shown in Table 2. The scale had high internal consistency in the present study (Cronbach's $\alpha = 0.95$).

2.4.3. Weight Bias Internalization Scale (WBIS)

The Mandarin Chinese version of the WBIS was used to assess participants' internalized weight-based stereotypes [31]. It has satisfactory reliability and validity [32]. It contains 11 items which are rated from 1 (*strongly disagree*) to 5 (*strongly agree*) [33]. Higher scores indicate a higher level of internalized weight-related stigma. The WBIS items are shown in Table 2. The scale had high internal consistency in the present study (Cronbach's $\alpha = 0.89$).

2.5. Study sample size justification

There is general consensus that minimum of 200 participants per group are needed to ensure adequate and stable performance (>80 % power) [34]. In the present study, although the sample sizes from each region were different. all three targeted group variables (i.e., gender, region, and overweight status) met this criterion. Therefore, given the sufficient number of participants, it was expected that reliable results would be generated.

2.6. Variables

The present study examined three personal variables for DIF including: gender (male vs. female), region (Taiwan vs. China), and

overweight status (yes vs. no). Gender and region were retrieved from the participants' demographic information. Overweight status was calculated using the participants' body mass index (BMI), which used participants' self-reported weight in kilograms divided by their squared height in meters. A landmark study investigated BMI, waist circumference, and risk factors related to chronic diseases among the Chinese population. The results suggested using cut-off point of 24 on the BMI for being overweight because it had the most optimal sensitivity and specificity for identifying the risk factors for being obese [35]. Two more recent studies also suggested BMI >24 as being overweight [36,37]. Therefore, a BMI over 24 was considered as being overweight in the present study.

2.7. Statistical methods

Rasch analyses was carried out using Winsteps software (version 5.4.0.) to evaluate DIF and other psychometric properties. Other demographic and descriptive statistics were conducted with IBM SPSS version 28.0. Participants' demographics were calculated with means and standard deviations for ratio variables (i.e., age and BMI), and then frequencies and percentages for categorical variables (i.e., gender and marital status).

A previous factor analysis study confirmed the single-factor construct of the TAPAS [12]. Therefore, raw scores of all the ten TAPAS items were combined and analyzed together to be consistent with previous studies [12]. Rasch analysis was used to transform the raw scores and provide an estimate of the participants' level of tendency to avoid physical activity or sport due to appearance concerns. Furthermore, the Rasch analysis estimated the individual ability along with the TAPAS item difficulty calibrations of the latent traits evaluated. Individual ability indicates participants' characteristics regarding their psychological avoidance toward physical activity and exercise, and the item difficulty indicates if the items assess a higher or a lower level of the psychological avoidance toward physical exercise.

Moreover, the estimated item difficulty in logits for each of the subgroups were used to calculate the DIF contrasts. When a DIF contrast between the two comparable groups is > .64, this item should be flagged as exhibiting not only a statistically significant DIF, but also a meaningful effect size (>0.64 DIF contrast indicates moderate to large DIF) [38]. The cut-off value of 0.64 is an absolute value. When researchers identify significant DIF, it indicates that this particular item should not be treated as the same item across two groups of respondents. In the present study, three subgroups were compared: gender, region, and overweight status.

Additionally, the partial credit model [39] was used to analyze the five-point Likert scale to verify if the adjacent rating category in TAPAS advanced monotonically as expected. The goodness-of-fit statistics retrieved from the Rasch analysis were used to determine the item unidimensionality and to confirm the construct validity of the TAPAS. Mean-Square (MnSq) is a type of fit statistic that indicates the size of the randomness in the analysis. In a Rasch measurement model, a MnSq of 1.0 shows there is no distortion of the Rasch calibration system. When MnSq is less than 1.0, it indicates the collected data might be too predictable in the Rasch model, which might lead to the conclusion of item redundancy. On the other hand, when MnSq is greater than 1.0, it could indicate that the collected data might have unpredictable method effects that cause unmodeled noise. The criteria were set that the Infit MnSq statistics should fall between 0.5 and 1.5 associated with a standardized mean square scored (Z_{std}) between −2 and +2 [40]. The Z_{std} provides a *t*-test statistic that evaluates the probability of the MnSq calculation occurring by chance. Therefore, as long as the MnSq values fall within the targeted ranges, the Z_{std} value can be ignored. Moreover, ceiling and floor effects of the TAPAS were examined by evaluating whether more than 15 % of the enrolled participants achieved the highest or the lowest possible scores, respectively [41].

Lastly, the present study also validated the concurrent validity of the TAPAS with two psychometric scales that are highly related to weight-stigma: the Weight Self-Stigma Questionnaire (WSSQ) and the Weight Bias Internalization Scale (WBIS). Given that all the scale items were self-rated with ordinal scales, Spearman correlation coefficients were used. It was expected that there would be at least moderate ($r = 0.4$ to 0.7) or strong ($r > 0.7$) correlation between these assessments.

Table 3
Gender DIF for the 10 TAPAS items.

Item	Calibration		*t*	*p*	DIF Contrast[a]
	Male	Female			
1. I find myself avoiding participating in sport because of my weight.	.46	.52	−.94	.35	−.06
2. I avoid participating in sport because of my fear of being judged about my lack of physical ability.	.21	.24	−.55	.58	−.04
3. I worry about participating in sport because I don't like how my body looks when playing sport.	.30	.21	1.37	.17	.09
4. I am afraid other people will notice my physical flaws when I participate in sport.	−.19	−.43	3.84	<.01	.25
5. I am concerned about what other people think of my appearance when I participate in sport.	−.57	−.57	.04	.97	.00
6. I avoid physical activity because I might get teased about my weight.	.35	.59	−3.61	<.01	−.24
7. I avoid physical activity because of my fear of being judged about my physical appearance.	.17	.28	−1.70	.09	−.11
8. I avoid physical activity because I worry that people may make negative comments about my body.	.28	.34	−.95	.34	−.06
9. I avoid physical activity because I worry people may be thinking negatively about my physical appearance.	.18	.33	−2.22	.03	−.15
10. I would prefer to participate in physical activity in a more private setting.	−1.22	−1.49	4.37	<.01	.27

Note.
DIF = Differential Item Functioning.
[a] Negative values indicate that the probability of the item endorsement is easier for male participants than female participants.

3. Results

3.1. Participants' characteristics

In the present study, 2927 participants were recruited from higher education population in China (n = 2319) and Taiwan (n = 608). In the Taiwanese sample, 43 % were male, and in the Chinese sample, 45 % were male. The majority of participants from both China (n = 2263, 97.6 %) and Taiwan (n = 491, 80.8 %) were single. Participants from Taiwan were significantly older than those from China ($t = 34.3$, $p < .001$), with the mean age of 29.1 years (SD = 6.4) and 20.2 years (SD = 1.8), respectively. Additionally, Taiwanese and Chinese participants showed significant differences in the proportions relating to gender ($\chi^2 = 8.6$, $p = .01$) and marital status ($\chi^2 = 324.5$, $p < .001$). More detailed demographic information shown in in Table 1.

3.2. Main results

No DIF was found in the gender subgroup (the absolute values of the DIF contrasts ranged from 0 to 0.27 logits) (Table 3). Three TAPAS items (4, 6, 10) showed significant differences ($p < .01$). Similar results were found in overweight status (the absolute values of the DIF contrasts ranged from 0.07 to 0.44 logits) (Table 4). Three items (5, 6, 10) showed significant differences ($p < .01$). Additionally, eight out of ten items for the region DIF showed significant results. However, all the above significant items were still within the DIF <0.64 cut-off point, except for Item 10 (*I would prefer to participate in physical activity in a more private setting*) on the TAPAS. It was not only significant but also meaningful DIF between participants from different countries (the absolute values of the DIF contrast = 1.41) (Table 5). The estimated item difficulty for participants from Taiwan and China were −2.48 and −1.08 logits, respectively. Therefore, the participants from Taiwan found it easier to endorse this item than participants from China (1.41 logits). The line plots of the Rasch logits with the corresponding three subgroups' comparisons are shown in Fig. 1.

Moreover, the five-point Likert scales used by the TAPAS increased difficulty monotonically. The goodness of fit statistics derived from the Rasch analysis also supported its unidimensionality, except for Item 10 (*I would prefer to participate in physical activity in a more private setting*), which did not fit Rasch expectations (Infit MnSq = 1.95; Zstd = 9.90). Further details of the individual item goodness-of-fit statistics and Rasch calibration estimates are shown in Table 6. Lastly, when examining the response patterns of the enrolled participants, it was found that 17 (0.6 %) and 332 (11.3 %) out of the 2927 participants had achieved the highest and lowest possible TAPAS scores, respectively. Therefore, no ceiling or floor effects were found in the TAPAS. Among the 17 participants with the highest possible scores, 16 were from China (94.1 %), 12 were males (70.6 %), and 9 were overweight (52.9 %). Among the 332 participants with the lowest possible scores, 302 were from China (91.0 %), 162 were males (48.8 %), and 108 were overweight (32.5 %).

3.3. Other analyses

Individual item score comparisons of the TAPAS, WBIS, and WSSQ for participants from China and Taiwan are shown in Table 2. In general, independent sample *t*-tests showed significant differences between groups in the majority items in all the three assessments. More specifically, the results indicated that participants from Taiwan were significantly more likely to avoid participating in physical activity or sport ($t = 1.22$ to 14.86; $p < .05$ to $p < .001$, based on TAPAS scores), and also had a significantly higher level of weight-

Table 4
BMI DIF for the 10 TAPAS items.

Item	Calibration		*t*	*p*	DIF Contrast[a]
	Overweight	Non-overweight			
1. I find myself avoiding participating in sport because of my weight.	.35	.54	−2.61	.01	−.19
2. I avoid participating in sport because of my fear of being judged about my lack of physical ability.	.18	.25	−.92	.36	−.07
3. I worry about participating in sport because I don't like how my body looks when playing sport.	.14	.29	−2.05	.04	−.15
4. I am afraid other people will notice my physical flaws when I participate in sport.	−.21	−.37	2.29	.02	.16
5. I am concerned about what other people think of my appearance when I participate in sport.	−.39	−.63	3.43	<.01	.24
6. I avoid physical activity because I might get teased about my weight.	.30	.55	−3.38	<.01	−.25
7. I avoid physical activity because of my fear of being judged about my physical appearance.	.15	.26	−1.45	.15	−.11
8. I avoid physical activity because I worry that people may make negative comments about my body.	.28	.33	−.78	.43	−.05
9. I avoid physical activity because I worry people may be thinking negatively about my physical appearance.	.19	.29	−1.38	.18	−.10
10. I would prefer to participate in physical activity in a more private setting.	−1.05	−1.49	6.37	<.01	.44

Note.
DIF = Differential Item Functioning.
[a] Negative values indicate that the probability of the item endorsement is easier for overweight participants than non-overweight participants.

Table 5
Region DIF for the 10 TAPAS items.

Item	Calibration		*t*	*p*	DIF Contrast[a]
	Taiwan	China			
1. I find myself avoiding participating in sport because of my weight.	.70	.43	3.40	<.01	.27
2. I avoid participating in sport because of my fear of being judged about my lack of physical ability.	.28	.22	.74	.46	.06
3. I worry about participating in sport because I don't like how my body looks when playing sport.	.56	.16	5.16	<.01	.40
4. I am afraid other people will notice my physical flaws when I participate in sport.	−.47	−.29	−2.40	.02	−.18
5. I am concerned about what other people think of my appearance when I participate in sport.	−.89	−.48	−5.51	<.01	−.41
6. I avoid physical activity because I might get teased about my weight.	.88	.38	6.39	<.01	.51
7. I avoid physical activity because of my fear of being judged about my physical appearance.	.48	.16	4.12	<.01	.32
8. I avoid physical activity because I worry that people may make negative comments about my body.	.55	.25	3.86	<.01	.30
9. I avoid physical activity because I worry people may be thinking negatively about my physical appearance.	.49	.20	3.62	<.01	.28
10. I would prefer to participate in physical activity in a more private setting.	−2.48	−1.08	−18.4	<.01	−1.41

Note.
DIF = Differential Item Functioning.
[a] Negative values indicate that the probability of the item endorsement is easier for participants from Taiwan than participants from China.

related self-stigma ($t = -0.07$ to 12.39; $p < .05$ to $p < .001$, based on WBIS scores; $t = 0.84$ to 14.64; $p < .05$ to $p < .001$, based on WSSQ scores). Additionally, the concurrent validity of the TAPAS was investigated with WBIS and WSSQ. The Spearman's rho correlation indicated strong associations between TAPAS and (i) WBIS ($r = 0.74$, $p < .001$) and (ii) WSSQ ($r = 0.71$, $p < .001$). Therefore, convergent validity of the TAPAS was supported.

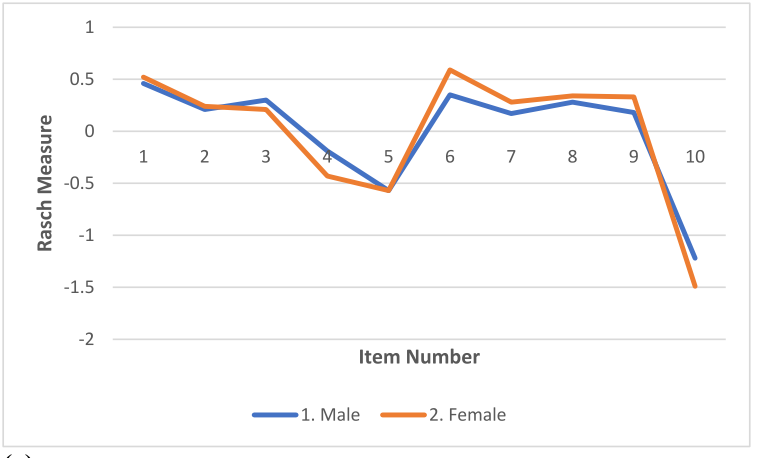
4. Discussion

The present study examined DIF in three subgroups (i.e., gender, region, and overweight status) of the newly developed TAPAS and further investigated its psychometric properties using Rasch analysis. Measurement invariance is extremely important to evaluate in any self-report psychometric instruments to ensure the construct being assessed is the same across different groups [42]. The study confirmed that the TAPAS had no gender or overweight status DIF, which provided evidence of its measurement invariance. However, the answer patterns to Item 10 (*I would prefer to participate in physical activity in a more private setting*) varied for participants from Taiwan and China with a significant DIF.

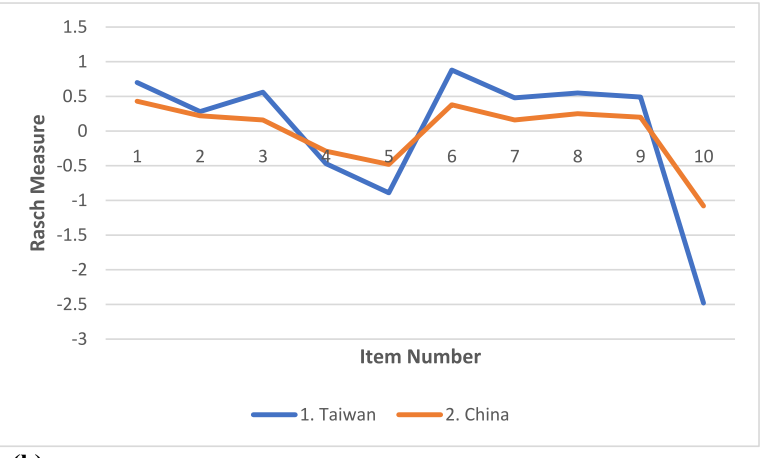
A recent systematic review [43] explored weight discrimination and the level of physical activity. The results confirmed that internalized weight stigma was associated with reduced physical activity for adult and youth populations. However, physical activity participation is important because it would improve an individual's self-perception and acceptance of their body image [44]. Multiple studies have already indicated that compared to males, females are less satisfied with their own body image [45–48]. When individuals do not feel comfortable with their own body shapes and body images, different strategies are used to adapt or to manage thoughts toward negative body image. Cash et al. [49] developed the Body Image Coping Strategies Inventory (an instrument like the TAPAS that assesses the same underlying construct of dealing with weight stigma) and identified that relative to males, females reported using significantly more coping strategies by appearance fixing and avoidance to evade threats to their body-image perceptions. For example, females may engage in unhealthy and compulsive exercise to gain the desired appearance [50], or avoid social exposure and engagement [51]. In contrast, another study reported that male adolescents showed significantly higher avoidance to physical activity and exercise scores compared to female adolescents [52]. Given these varied findings concerning gender and how it may influence individuals' perceived stigma and self-stigma that leads to relevant behaviors in prior studies, the TAPAS can serve as a reliable and valid psychometric instrument. The present study results showed that the TAPAS has robust psychometric properties and that there was no gender or weight status DIF. Therefore, the TAPAS is appropriate to be administered among both male and female populations, as well as overweight and non-overweight populations.

Because both Taiwan and China originate from the same Confucianism culture, people in the two regions may interpret TAPAS items similarly. However, Item 10 on the TAPAS had significant differences in DIF between participants from Taiwan and China (i.e., the item was not treated the same by the participants from the two regions). These DIF results might be explained by the different understandings and perspectives by participants from different regions [53,54]. More specifically, people in Taiwan and those in China may have interpreted "private setting" differently because Taiwan has democratic governance while China has communist governance, although the people in both regions have the same cultural background. People in China may have interpreted "private setting" as facilities that could only be accessed by specific individuals (e.g., celebrities). People in Taiwan may have interpreted "private setting" as being their homes. However, this is speculation, and future qualitative studies are needed to clarify how Taiwanese and Chinese people interpret this item. Moreover, this item did not fit Rasch expectations. Therefore, further study is needed to explore if this particular item statement might be explained differently by the potential users.

Liu and Jane Rogers [55] compared four treatments (deleting, ignoring, multiple-group modeling, and modeling DIF as a secondary dimension) for items that have been identified as showing DIF. The results showed that none of the DIF treatments had a dominant advantage over other treatments. However, they also indicated that ignoring DIF items showed the worst performance of estimating trait values. Additionally, deleting DIF items is not recommended as this would result in the least accurate estimates for item



(a)



(b)

(c)

Fig. 1. Differential Item Function (DIF) in TAPAS. (a) DIF across gender; (b) DIF across regions; (c) DIF across weight status.

Table 6
Rasch analyses of the 10 TAPAS items.

Item	Measure	SE	Infit		Outfit		Average Rating Measure[a]				
			MnSq	*Zstd*	*MnSq*	*Zstd*	1	2	3	4	5
1. I find myself avoiding participating in sport because of my weight.	.58	.03	1.34	9.90	1.41	9.90	−4.74	−1.68	−.11	1.34	3.44
2. I avoid participating in sport because of my fear of being judged about my lack of physical ability.	.22	.03	1.05	1.60	1.07	2.35	−5.03	−1.80	−.15	1.14	3.36
3. I worry about participating in sport because I don't like how my body looks when playing sport.	.19	.03	.83	−6.25	.82	−6.33	−5.15	−1.77	−.03	1.21	3.36
4. I am afraid other people will notice my physical flaws when I participate in sport.	−.29	.03	.89	−4.24	.88	−4.20	−5.57	−2.04	−.36	.64	3.05
5. I am concerned about what other people think of my appearance when I participate in sport.	−.52	.03	1.15	5.28	1.17	5.97	−5.86	−2.13	−.58	.47	2.65
6. I avoid physical activity because I might get teased about my weight.	.44	.03	.78	−8.26	.75	−9.15	−5.03	−1.67	.07	1.56	4.06
7. I avoid physical activity because of my fear of being judged about my physical appearance.	.23	.03	.64	−9.90	.63	−9.90	−5.31	−1.80	−.03	1.38	3.92
8. I avoid physical activity because I worry that people may make negative comments about my body.	.28	.03	.63	−9.90	.62	−9.90	−5.22	−1.78	.03	1.45	4.04
9. I avoid physical activity because I worry people may be thinking negatively about my physical appearance.	.24	.03	.68	−9.90	.67	−9.90	−5.26	−1.80	−.02	1.39	3.95
10. I would prefer to participate in physical activity in a more private setting.[b]	−1.37	.03	1.95	9.90	2.05	9.90	−6.10	−2.29	−1.09	−.27	1.23

Note.
[a] The average measure is expected to increase with category value (Linacre, 2002).
[b] Infit statistics >1.5 associated with Zstd >2: item misfit. MnSq = Mean Square; Zstd = Standardized score.

parameters and calibration [55]. Therefore, future studies should further examine and modify Item 10 in the TAPAS if necessary. Moreover, it is recommended to clinicians that the raw data of Item 10 on the TAPAS should not be used and compared directly. It is suggested that clinicians or researchers who plan to use the TAPAS in an international study that might involve different countries, should split Item 10 into separate sample specific items (e.g., Item 10-Taiwan vs. Item 10-China) to capture the potential item bias, so that it can take account of the DIF issue at the same time retain precision of measurement [56].

4.1. Limitations and Generalizability

The study's findings were reliant on participants' self-report data. Multiple studies have indicated that socials desirability can exist when individuals report their height and weight, as well as health-related information [57,58]. Therefore, in the present study, the calculation of the BMI might not be the most accurate information because height and weight were self-reported by the participants. Therefore, future studies should try to obtain more objective measurements to validate height, weight, and health-related information. Additionally, the present study evaluated an individual's tendency to avoid physical activity and sports. However, it did not assess the actual level of participants' physical activity and sports participation. Therefore, the perceptions of high avoidance tendency may not be necessarily connected with less physical activity and sport engagement. Future studies should collect actual participation levels and further examine its association with TAPAS. Moreover, due to the nature of the present study, it only collected limited demographic data. Other potential socio-economic variables that might affect DIF were not examined. For example, educational level has been considered as influential in relation to perceptions and body image [59]. Therefore, educational level is a key foundation for participants to possess different levels of health knowledge and it might also influence their interpretation of self-stigma. Future study should consider examining DIF across different educational levels. In this case, if DIF exists, clinicians and researchers should be mindful when conducting a research study that involves populations with differing educational levels because the results might not be comparable across educational level. On the other hand, if no educational DIF exists in the TAPAS, then researchers could be confident that the TAPAS can be used across different educational levels with measurement invariance.

The data were collected using convenience sampling and may not be generalizable to either the whole Taiwanese or Chinese populations. Additionally, the mean ages of the two regions' samples were admittedly different (i.e., participants from Taiwan were almost nine years older than the participants from China). Moreover, although participants in both Taiwan and China can both read and speak Mandarin Chinese, there could possibly have been potential cultural variations between these two samples. Future study should include more diverse participants with different age ranges and also investigate if cultural adaptation might be needed in further examination of the potential DIF in the TAPAS.

A European University Consortium surveyed multi-disciplinary healthcare professionals (including nurses, doctors, and allied health professionals) regarding understanding of patients' appearance concerns and confidence in addressing concerns/needs. The results showed that participants self-reported inadequate knowledge regarding appearance-specific care and wanted access to an accredited course [60]. Additionally, multiple studies indicate that high degree of body image dissatisfaction and inaccurate perceptions exists among students within health-related fields, such as medical students [61,62], nursing students [63], and pharmacy

students [64]. Therefore, future research should not only further explore whether DIF exists between educational level but also between participants' study majors (e.g., health-related vs. non-health-related).

5. Conclusion

The present study showed that the TAPAS can assess the tendency to avoid physical activity and sport for both Taiwanese and Chinese university students under the same latent concept. No DIF items were identified for the TAPAS across gender (male vs. female) or weight status (overweight vs. non-overweight). However, Item 10 (i.e., *Prefer to participate in physical activity in a more private setting*) was found to display DIF across Taiwanese and Chinese university students. Moreover, the same item was found not to fit with the concept of tendency to avoid physical activity and sport. Therefore, future studies should further examine this item from the TAPAS.

Data availability statement

The data associated with the study have not been deposited into a publicly available repository. The data and code that support the present findings will be made available from the corresponding author upon reasonable request.

Ethics approval and patient consent statement

All participants from Taiwan and mainland China obtained the detailed information regarding the study purpose and informed consent was obtained. The study protocol was approved by the National Cheng Kung University Human Research Ethics Committee (Approval No. NCKU HREC-E−110-486-2), the National Cheng Kung University Hospital Institute of Review Board (IRB No. A-ER-111-445), and the Institutional Review Board of the Jiangxi Psychological Consultant Association (IRB ref: JXSXL-2021-J99).

Permission to reproduce material

Not applicable.

CRediT authorship contribution statement

Chia-Wei Fan: Writing – original draft, Visualization, Software, Resources, Formal analysis, Conceptualization. **Po-Ching Huang:** Writing – review & editing, Validation, Methodology, Investigation, Data curation, Conceptualization. **I-Hua Chen:** Writing – review & editing, Validation, Resources, Methodology, Investigation, Data curation, Conceptualization. **Yu-Ting Huang:** Writing – review & editing, Validation, Methodology, Conceptualization. **Jung-Sheng Chen:** Writing – review & editing, Validation, Methodology, Conceptualization. **Xavier C.C. Fung:** Writing – review & editing, Validation, Methodology, Conceptualization. **Ji-Kang Chen:** Writing – review & editing, Validation, Methodology, Conceptualization. **Yung-Ning Yang:** Writing – review & editing, Validation, Resources, Methodology, Investigation, Conceptualization. **Kerry S. O'Brien:** Writing – review & editing, Supervision, Methodology, Conceptualization. **Chung-Ying Lin:** Writing – review & editing, Supervision, Software, Resources, Project administration, Investigation, Funding acquisition, Data curation, Conceptualization. **Mark D. Griffiths:** Writing – review & editing, Supervision, Methodology, Conceptualization.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Chung-Ying Lin reports financial support was provided by Ministry of Science and Technology, Taiwan. Chung-Ying Lin reports financial support was provided by Ministry of Education, Taiwan. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This research was supported in part by (received funding from) the Ministry of Science and Technology, Taiwan (MOST 110-2410-H-006-115; MOST 111-2410-H-006-100), the Higher Education Sprout Project, the Ministry of Education at the Headquarters of University Advancement at the National Cheng Kung University (NCKU), the 2021 Southeast and South Asia and Taiwan Universities Joint Research Scheme (NCKU 31), and the International Research Collaboration Fund granted by the Department of Social Work, The Chinese University of Hong Kong (Grant number: 19231106).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.heliyon.2023.e22583.

References

[1] World Health Organization, Obesity and overweight [Internet], World Health Organization, 2021. URL: https://www.who.int/news-room/fact-sheets/detail/obesity-and-overweight. (Accessed 14 June 2023).
[2] C. Greenleaf, E.M. Boyer, T.A. Petrie, High school sport participation and subsequent psychological well-being and physical activity: the mediating influences of body image, physical competence, and instrumentality, Sex. Roles 61 (2009) 714–726, https://doi.org/10.1007/s11199-009-9671-z.
[3] A. Hilbert, W. Rief, E. Braehler, Stigmatizing attitudes toward obesity in a representative population-based sample, Obesity 16 (2008) 1529–1534, https://doi.org/10.1038/oby.2008.263.
[4] C. Sikorski, M. Luppa, M. Kaiser, H. Glaesmer, G. Schomerus, H.H. König, et al., The stigma of obesity in the general public and its implications for public health - a systematic review, BMC Publ. Health 11 (1) (2011) 661, https://doi.org/10.1186/1471-2458-11-661.
[5] S. Allender, G. Cowburn, C. Foster, Understanding participation in sport and physical activity among children and adults: a review of qualitative studies, Health Educ. Res. 21 (2006) 826–835, https://doi.org/10.1093/her/cyl063.
[6] S. Somerset, D.J. Hoare, Barriers to voluntary participation in sport for children: a systematic review, BMC Pediatr. 18 (1) (2018) 47, https://doi.org/10.1186/s12887-018-1014-1.
[7] V.P.N. Miranda, N. Sousa de Morais, E. Rodrigues de Faria, P. Roberto dos Santos Amorim, J.C.B. Marins, S. do Carmo Castro Franceschini, et al., Body dissatisfaction, physical activity, and sedentary behavior in female adolescents, Rev. Paul. de Pediatr. 36 (2018) 482–490, https://doi.org/10.1590/1984-0462/, 2018;36;4;00005.
[8] M. Hao, Y. Fang, W. Yan, J. Gu, Y. Hao, C. Wu, Relationship between body dissatisfaction, insufficient physical activity, and disordered eating behaviors among university students in southern China, BMC Publ. Health 22 (2022), https://doi.org/10.1186/s12889-022-14515-9, 2054.
[9] A. Thiel, J.M. John, J. Carl, H.K. Thedinga, Weight stigma experiences and physical (in)activity: a biographical analysis, Obes. Facts 13 (3) (2020) 386–402, https://doi.org/10.1159/000507936.
[10] N. Bevan, K.S. O'Brien, C.Y. Lin, J.D. Latner, B. Vandenberg, R. Jeanes, et al., The relationship between weight stigma, physical appearance concerns, and enjoyment and tendency to avoid physical activity and sport, Int. J. Environ. Res. Publ. Health 18 (2021) 9957, https://doi.org/10.3390/ijerph18199957.
[11] H.K. Thedinga, R. Zehl, A. Thiel, Weight stigma experiences and self-exclusion from sport and exercise settings among people with obesity, BMC Publ. Health 21 (1) (2021) 565, https://doi.org/10.1186/s12889-021-10565-7.
[12] N. Bevan, K.S. O'Brien, J.D. Latner, C.Y. Lin, B. Vandenberg, R. Jeanes, et al., Weight stigma and avoidance of physical activity and sport: development of a scale and establishment of correlates, Int. J. Environ. Res. Publ. Health 19 (2022), 16370, https://doi.org/10.3390/ijerph192316370.
[13] W.H. Angoff, Perspectives on differential item functioning methodology, in: P.W. Holland, H. Wainer (Eds.), Differential Item Functioning, Lawrence Erlbaum Associates, Inc, 1993, pp. 3–23.
[14] R.N. Jones, Differential item functioning and its relevance to epidemiology, Curr. Epidemiol. Rep. 6 (2019) 174–183, https://doi.org/10.1007/s40471-019-00194-5.
[15] P. Martinková, A. Drabinová, Y.L. Liaw, E.A. Sanders, J.L. McFarland, R.M. Price, Checking equity: why differential item functioning analysis should be a routine part of developing conceptual assessments, CBE-Life Sci. Educ. 16 (2017) rm2, https://doi.org/10.1187/cbe.16-10-0307.
[16] W.R. Nugent, Understanding DIF and DTF: description, methods, and implications for social work research, J. Soc. Soc. Work. Res. 8 (2017) 305–334, https://doi.org/10.1086/691525.
[17] E. Díaz, G. Brooks, G. Johanson, Detecting differential item functioning: item response theory methods versus the Mantel-Haenszel procedure, Int. J. Assess. Tools Educ. 8 (2021) 376–393, https://doi.org/10.21449/ijate.730141.
[18] J.S. Lai, J. Teresi, R. Gershon, Procedures for the analysis of Differential item functioning (DIF) for small sample sizes, Eval. Health Prof. 28 (2005) 283–294, https://doi.org/10.1177/0163278705278276.
[19] S.J. Osterlind, H.T. Everson, Differential Item Functioning. Quantitative Applications in the Social Sciences, second ed., SAGE Publications, Inc, 2009 https://doi.org/10.4135/9781412993913.
[20] D. Magis, S. Beland, F. Tuerlinckx, A comparison of five methods to detect differential item functioning in polytomous items, Educ. Psychol. Meas. 79 (2019) 310–336.
[21] G.R. Budgell, N.S. Raju, D.A. Quartetti, Analysis of differential item functioning in translated assessment instruments, Appl. Psychol. Meas. 19 (1995) 309–321, https://doi.org/10.1177/014662169501900401.
[22] C.L. Hulin, F. Drasgrow, C.K. Parsons, Item Response Theory: Application to Psychological Measurement, Dow Jones-Irwin, 1983.
[23] R. Jabrayilov, W.H. Emons, K. Sijtsma, Comparison of classical test theory and item response theory in individual change assessment, Appl. Psychol. Meas. 40 (2016) 559–572, https://doi.org/10.1177/0146621616664046.
[24] W.J. Van Der Linden, R.K. Hambleton (Eds.), Handbook of modern item response theory, Springer, New York, NY, 2014.
[25] B.D. Wright, N. Panchapakesan, A procedure for sample-free item analysis, Educ. Psychol. Meas. 29 (1969) 23–48.
[26] J.P. Vandenbroucke, E. von Elm, D.G. Altman, P.C. Gøtzche, C.D. Mulrow, S.J. Pocock, et al., Strengthening the reporting of observational studies in epidemiology (STROBE): explanation and elaboration, PLoS Med. 4 (2007) e297, https://doi.org/10.1371/journal.pmed.0040297.
[27] C.W. Fan, Y.L. Chang, P.C. Huang, X.C.C. Fung, J.K. Chen, N. Bevan, et al., The Tendency to Avoid Physical Activity and Sport Scale (TAPAS): Rasch analysis with differential item functioning testing among a Chinese sample, BMC Psychol 11 (2023) 369, https://doi.org/10.1186/s40359-023-01377-y.
[28] Y. C. Lin, J. S. Chen, N. Bevan, K. S. O'Brien, C. Strong, M.C. Tsai, et al., Chinese Version of the Tendency to Avoid Physical Activity and Sport (TAPAS) Scale: Testing unidimensionality, measurement invariance, concurrent validity, and known-group validity, Under revision for publication in Heliyon.
[29] J. Lillis, J.B. Luoma, M.E. Levin, S.C. Hayes, Measuring weight self-stigma: the weight self-stigma questionnaire, Obesity 18 (2010) 971–976, https://doi.org/10.1038/oby.2009.353.
[30] K.P. Lin, M.L. Lee, Validating a Chinese version of the Weight Self-Stigma Questionnaire for use with obese adults, Int. J. Nurs. Pract. 23 (2017), e12537, https://doi.org/10.1111/ijn.12537.
[31] C.Y. Lin, V. Imani, P. Cheung, A.H. Pakpour, Psychometric testing on two weight stigma instruments in Iran: weight self-stigma questionnaire and weight bias internalized scale, Eat, Weight Disord 25 (2020) 889–901, https://doi.org/10.1007/s40519-019-00699-4.
[32] H. Chen, Y. Ye, Validation of the Weight Bias Internalization Scale for mainland Chinese children and adolescents, Front. Psychol. 11 (2021), 594949, https://doi.org/10.3389/fpsyg.2020.594949.
[33] L.E. Durso, J.D. Latner, Understanding self-directed stigma: development of the Weight Bias Internalization Scale, Obesity 16 (2008) S80–S86, https://doi.org/10.1038/oby.2008.448.
[34] N.W. Scott, P.M. Fayers, N.K. Aaronson, A. Bottomley, Alexander de Graeff, M. Groenvold, et al., A simulation study provided sample size guidance for differential item functioning (DIF) studies using short scales, J. Clin. Epidemiol. 62 (2009) 288–295, https://doi.org/10.1016/j.jclinepi.2008.06.003.
[35] B.F. Zhou, Cooperative Meta-Analysis Group of the Working Group on Obesity in China, Predictive values of body mass index and waist circumference for risk factors of certain related diseases in Chinese adults–study on optimal cut-off points of body mass index and waist circumference in Chinese adults, Biomed. Environ. Sci. 15 (1) (2002) 83–96.
[36] R. Bai, W. Wu, W. Dong, J. Liu, L. Yang, J. Lyu, Forecasting the populations of overweight and obese Chinese adults, Diabetes Metab. Syndr. Obes. 13 (2020) 4849–4857, https://doi.org/10.2147/dmso.s274110.
[37] M. Gao, Y.X. Wei, J. Lyu, C.Q. Yu, Y. Guo, Z. Bian, et al., The cut-off points of body mass index and waist circumference for predicting metabolic risk factors in Chinese adults, Zhonghua Liuxingbingxue Zazhi 40 (12) (2019) 1533–1540, https://doi.org/10.3760/cma.j.issn.0254-6450.2019.12.006.
[38] R. Zwick, D.T. Thayer, C. Lewis, An empirical Bayes approach to Mantel-Haenszel DIF analysis, J. Educ. Meas. 36 (1) (1999) 1–28, https://doi.org/10.1111/j.1745-3984.1999.tb00543.x.
[39] G.N. Maters, Partial credit model, in: W.J. van der Linden (Ed.), Handbook of item response theory: three volume set, Chapman and Hall, 2019.

[40] J.M. Linacre, Fit diagnosis: Infit outfit mean-square standardized [Internet], URL: https://www.winsteps.com/winman/misfitdiagnosis.htm. (Accessed 25 May 2023).
[41] C.M. Gulledge, V.A. Lizzio, D.G. Smith, E. Guo, E.C. Makhni, What are the floor and ceiling effects of patient-reported outcomes measurement information system computer adaptive test domains in orthopaedic patients? A systematic review, Arthroscopy 36 (2020) 901–912.e7, https://doi.org/10.1016/j.arthro.2019.09.022.
[42] W. Meredith, Measurement invariance, factor analysis and factorial invariance, Psychometrika 58 (1993) 525–543, https://doi.org/10.1007/bf02294825.
[43] R.L. Pearl, T.A. Wadden, J.M. Jakicic, Is weight stigma associated with physical activity? A systematic review, Obesity 29 (2021) 1994–2012, https://doi.org/10.1002/oby.23274.
[44] E. Gualdi-Russo, N. Rinaldo, S. Masotti, B. Bramanti, L. Zaccagni, Sex differences in body image perception and ideals: analysis of possible determinants, Int. J. Environ. Res. Publ. Health 19 (2022) 2745, https://doi.org/10.3390/ijerph19052745.
[45] G.X. Carvalho, A.P. Nunes, C.L. Moraes, G.V. Veiga, Body image dissatisfaction and associated factors in adolescents, Ciência Saúde Coletiva 25 (2020) 2769–2782, https://doi.org/10.1590/1413-81232020257.27452018.
[46] C. Mahon, D. Hevey, Processing body image on social media: gender differences in adolescent boys' and girls' agency and active coping, Front. Psychol. 12 (2021), 626763, https://doi.org/10.3389/fpsyg.2021.626763.
[47] S. Toselli, A. Grigoletto, L. Zaccagni, N. Rinaldo, G. Badicu, W.R. Grosz, et al., Body image perception and body composition in early adolescents: a longitudinal study of an Italian cohort, BMC Publ. Health 21 (1) (2021) 1381, https://doi.org/10.1186/s12889-021-11458-5.
[48] M.M. Voges, C.M. Giabbiconi, B. Schöne, M. Waldorf, A.S. Hartmann, S. Vocks, Gender differences in body evaluation: do men show more self-serving double standards than women? Front. Psychol. 10 (2019) 544, https://doi.org/10.3389/fpsyg.2019.00544.
[49] T.F. Cash, M.T. Santos, E.F. Williams, Coping with body-image threats and challenges: validation of the body image coping strategies inventory, J. Psychosom. Res. 58 (2005) 190–199, https://doi.org/10.1016/j.jpsychores.2004.07.008.
[50] J. Ruiz-Turrero, K. Massar, D. Kwasnicka, G.A. Ten Hoor, The relationship between compulsive exercise, self-esteem, body image and body satisfaction in women: a cross-sectional study, Int. J. Environ. Res. Publ. Health 19 (2022) 1857, https://doi.org/10.3390/ijerph19031857.
[51] N.J. Sabik, Is social engagement linked to body image and depression among aging women? J. Women Aging 29 (2017) 405–416, https://doi.org/10.1080/08952841.2016.1213106.
[52] M. Avci, C. Akliman, Gender differences in perception of body, expressions of body image and body image coping strategies among Turkish adolescents, Int. J. Curric. Instr. 10 (2018) 30–44.
[53] C.-C. Chang, C.-Y. Lin, P.C. Gronholm, T.-H. Wu, Cross-validation of two commonly used self-stigma measures, Taiwan versions of the Internalized Stigma Mental Illness scale and Self-Stigma Scale-Short, for people with mental illness, Assessment 25 (2018) 777–792, https://doi.org/10.1177/1073191116658547.
[54] I.-H. Chen, D.K. Ahorsu, A.H. Pakpour, M.D. Griffiths, C.-Y. Lin, C.-Y. Chen, Psychometric properties of three simplified Chinese online-related addictive behavior instruments among mainland Chinese primary school students, Front Psychiatry 11 (2020) 875, https://doi.org/10.3389/fpsyt.2020.00875.
[55] X. Liu, H. Jane Rogers, Treatments of differential item functioning: a comparison of four methods, Educ. Psychol. Meas. 82 (2022) 225–253, https://doi.org/10.1177/00131644211012050.
[56] C. Hagquist, D. Andrich, Recent advances in analysis of differential item functioning in health research using the Rasch model, Health Qual. Life Outcomes 15 (2017) 181, https://doi.org/10.1186/s12955-017-0755-0.
[57] M.A. Burke, K.G. Carman, You can be too thin (but not too tall): social desirability bias in self-reports of weight and height, Econ. Hum. Biol. 27 (Pt A) (2017) 198–222, https://doi.org/10.1016/j.ehb.2017.06.002.
[58] M.R. Larson, Social desirability and self-reported weight and height, Int. J. Obes. Relat. Metab. Disord. 24 (5) (2000) 663–665, https://doi.org/10.1038/sj.ijo.0801233.
[59] H. Shoraka, A. Amirkafi, B. Garrusi, Review of body image and some of contributing factors in Iranian population, Int. J. Prev. Med. 10 (2019) 19, https://doi.org/10.4103/ijpvm.ijpvm_293_18.
[60] H. Williamson, P. Antonelli, Å. Bringsén, G. Davies, D. Dèttore, D. Harcourt, et al., Supporting patients with body dissatisfaction: a survey of the experiences and training needs of European multi-disciplinary healthcare professionals, NursingPlus Open 4 (2018) 8–13, https://doi.org/10.1016/j.npls.2018.05.001.
[61] C.A. Divecha, M.A. Simon, A.A. Asaad, H. Tayyab, Body image perceptions and body image dissatisfaction among medical students in Oman, Sultan Qaboos Univ. Med. J. 22 (2022) 218–224, https://doi.org/10.18295/squmj.8.2021.121.
[62] B.C. Santos, D.D. Almeida, N.V. Guilarducci, R.R. Machado, Body dissatisfaction among undergraduate medical students in the city of Juiz de Fora, Minas Gerais, Brazil, Einstein 20 (2022), eAO6648, https://doi.org/10.31744/einstein_journal/2022ao6648.
[63] L.S. Duarte, M.N. Chinen, E. Fujimori, Distorted self-perception and dissatisfaction with body image among nursing students, Rev. Esc. Enferm. USP 55 (2021), e03665, https://doi.org/10.1590/s1980-220x2019042903665.
[64] A. Jaworowska, G. Bazylak, An outbreak of body weight dissatisfaction associated with self-perceived BMI and dieting among female pharmacy students, Biomed, Pharmacother 63 (2009) 679–692, https://doi.org/10.1016/j.biopha.2008.08.005.