Int. J. Sport Psychol., 2025; 56: 089-104
doi: 10.7352/IJSP.2025.56.089

Chinese version of the craving for rest and volitional energy expenditure scale: Confirmatory analysis and gender invariance measurement

Linxian Zeng*, Garry Kuan*, Ayu Suzailiana Binti Muhamad*, Fang Zhao, Zhutang Liu*, Yee Cheng Kueh*****

(*) Exercise and Sports Science Programme, School of Health Sciences, Universiti Sains Malaysia, 16150 Kubang Kerian, Kelantan, Malaysia.
(**) Biostatistics and Research Methodology Unit, School of Medical Sciences, Universiti Sains Malaysia, 16150 Kubang Kerian, Kelantan, Malaysia.
(***) Nanning Normal University, 508 XinNing Rd. WuMing District, NanNing, Guangxi, China.

PURPOSE: *Physical activity and sedentary motivational states may vary over time. The CRAVE scale (Craving for Rest and Volitional Energy Expenditure) is designed to assess transient needs and desires for exercise. This study aimed to conduct confirmatory analysis of the Chinese version of the CRAVE-C scale and measure gender invariance using confirmatory factor analysis (CFA).*

METHOD: *This study conducted a cross-sectional survey among three middle schools in Kunming, Yunnan, China. The (CRAVE) scale was translated into Chinese (CRAVE-C) using standard forward-backward translation methods. CFA was used to analyze the data using Mplus8. The intraclass correlation coefficient (ICC) (95% CI) within a one-week interval was used to evaluate the test-retest reliability.*

RESULTS: *The results showed that the content validity index (I-CVI) of all items in the Chinese version of the CRAVE-C scale exceeded 0.80, and the content validity was good. A goodness-of-fit analysis was performed to compare the 13-item and 10-item scale models, and the 10-item validation version outperformed the 13-item version. The final model showed satisfactory goodness of fit (CFI=0.992; TLI= 0.990; SRMR= 0.022; RMSEA= 0.038 (0.020-0.054), indicating that the 10-item CRAVE-C scale has good validity among high school students in Yunnan, China. The reliability of the CRAVE-C scale was satisfactory, with Cronbach alpha (moving: α= 0.913, resting: α= 0.894), and the internal consistency was good. The ICC of the items in the CRAVE-C scale ranges from 0.729 to 0.928, indicating good test-retest reliability. The results of the invariance analysis for different genders showed that the ΔCFI and ΔTLI values were both less than 0.01, supporting the gender invariance of the CRAVE-C scale among high school students.*

Correspondence to: Garry Kuan, Exercise and Sports Science Programme, School of Health Sciences, Universiti Sains Malaysia, 16150 Kubang Kerian, Kelantan, Malaysia. (E-mail: garry@usm.my)

Conclusion: *The CRAVE-C scale demonstrated good reliability and validity. The 10-item CRAVE-C scale is suitable for assessing physical activity and sedentary motivation among Chinese high school students, and measurement invariance was observed across different genders. It is suggested that this scale can be used to measure the "now" need/desire for movement and rest among high school students.*

Key Words: *CRAVE-C, high school students, Confirmatory factory analysis, physical activity, sedentary behaviour.*

Introduction

Physical inactivity and sedentary behavior are risk factors for noncommunicable diseases (Tcymbal et al., 2020). Physical activity is defined as any body movement produced by skeletal muscles that results in energy expenditure (Caspersen, Powell, & Christenson, 1985). In related studies, Levine et al. mentioned that physical activity thermogenesis can be subdivided into thermogenesis from volitional movements (such as sports and fitness-related activities) and thermogenesis from non-volitional movements (such as activities of daily living, fidgeting, spontaneous muscle contractions, and not lying down). Maintain posture while sitting) (Levine, Eberhardt, & Jensen, 1999). The World Health Organization (WHO) recommends that children and adolescents should engage in an average of at least 60 minutes of moderate-to-vigorous intensity physical activity per day per week (Bull et al., 2020), although achieving the levels of physical activity recommended by the WHO can protect health (Geneva, 2010), but studies have shown that sedentary behaviour is also an independent risk factor for many non-communicable diseases (Biddle et al., 2016; Biswas et al., 2015; Rezende, Rey-López, Matsudo, & Luiz, 2014; Wilmot et al., 2012). Sedentary behaviour is defined as "any awake behaviour characterised by energy expenditure ≤1.5 METs while sitting or reclining" (Bames et al., 2012). Studies have shown that diseases such as obesity, cardiovascular disease (CVD), and diabetes are associated with lack of muscle exercise and minimal energy expenditure (Piercy et al., 2018).

Recent research shows that children and adolescents around the world, especially girls, do not meet World Health Organization recommendations for moderate-to-vigorous physical activity (MVPA) for this age group, a condition that is associated with obesity and other comorbidities (Romero-Parra, Solera-Alfonso, Bores-García, & Delfa-De-La-Morena, 2023). However, intrinsic motivation is a better determinant of participation in MVPA (Romero-Parra et al., 2023). According to self-determination theory, students' motivation has an impact on maintaining physical activity levels (Duncan, Hall,

Wilson, & Jenny, 2010; Levine et al., 1999; Sevil, García-González, Abós, Generelo, & Aibar, 2019).

Related research by Stults-Kolehmainen et al. (2020) suggested that the desire for exercise and rest may depend on two different but related factors. The distinction between the two factors reflects current views on physical activity and sedentary behavior (Gabriel, Morrow, & Woolsey, 2012; Owen, Healy, Matthews, & Dunstan, 2010). Physical activity and possibly motivation vary throughout the day (Stults-Kolehmainen et al., 2021). Stults-Kolehmainen et al. (2021) developed and validated a brief assessment (CRAVE: Craving for Rest and Volitional Energy Expenditure) to measure motivational states (wanting, desire, urge) for physical activity and sedentary behavior (Stults-Kolehmainen et al., 2021), This scale contains 13 items containing a problem statement about physical activity and sedentary life, appended to 11-point Likert items. Among them, 5 items are related to physical activity, 5 items are related to sitting for a long time, and 3 items are filler items not used for analysis (Stults-Kolehmainen et al., 2021). The scale is available in a "now" version and a "past week" version. An important limitation of some studies is not using the "past week" version (Stults-Kolehmainen et al., 2020) . According to Stults-Kolehmainen et al. (2020), physical activity and motivations may vary throughout the day. Motivations, such as "want," "crave," and "urge," exhibit dynamic characteristics and can frequently fluctuate over a short period of time. By focusing on participants' current feelings, researchers can accurately capture their present motivational states. In contrast, asking them to reflect on their feelings over the past week may lead to memory errors and recall biases. The objective of this study was to evaluate momentary motivational states; hence, we selected the "now" version of CRAVE.

Currently, there is no Chinese version of the CRAVE scale, and this scale has not been used in China to measure the motivational state of physical activity and sedentary behavior among adolescents. In addition, Studies have shown that there are differences between male and female in the desire to move/rest(Stults-Kolehmainen et al., 2021). Based on this point of view, this study also explored the equivalence of the CRAVE scale in different gender groups by conducting equivalent measurements on different gender groups, thereby improving the value of this equivalence scale. Therefore, the purpose of the study was to translate the English version of the CRAVE scale into the Chinese version of the CRAVE-C scale; determine the content validity of the CRAVE-C scale; CFA verifies the validity and reliability of the CRAVE-C scale; Tests the measurement invariance across gender.

Materials And Methods

PARTICIPANTS

This study recruited 530 high school students to voluntarily participate in the study from Yunnan Province, China. All participants obtained informed consent from their parents or legal guardians before participating in the study. This study received a total of 530 questionnaires, 25 invalid questionnaires were excluded. Invalid questionnaires were identified through double-entering data, removed extreme responders (e.g., those who answered all 10 or all 0 to all items), and screened for univariate outliers using SPSS version 28.0. and the remaining 505 were analysed. Hair et al. (2006) recommended a minimum sample size of 500 participants for CFA. Therefore, the sample size used in this study was 505, which met this criterion.

MEASURES

The original CRAVE scale was developed and validated by Stults-Kolehmainen et al. (2021), to measure the motivational state of physical activity and sedentary behavior in adolescents (Stults-Kolehmainen et al., 2021). The scale contains 13 items containing a problem statement about physical activity and sitting, attached to 11-point Likert items. Among them, 5 items are related to physical activity, 5 items are related to sitting for a long time, and 3 items are filler items not used for analysis (Stults-Kolehmainen et al., 2021). Participants rated the statements they agreed with at the time of the survey (and now) on an 11-point Likert scale, with "0" not at all and "10" even more. The internal consistency and reliability of the CRAVE English version was very good, McDonald's ω was very high (0.97 overall), indicating good reliability of the scale (Stults-Kolehmainen et al., 2021).

Questionnaire translation

After obtaining permission from the original author, the English version of the CRAVE scale was translated into Chinese. The typical forward-backward translation procedure was used to ensure the quality of the translation (Brislin, 1970). First, bilingual researchers familiar with the scale were invited to translate the English version of the CRAVE scale into Chinese while retaining the content and meaning of the scale. Secondly, the translated Chinese version was back-translated into English by a Chinese who was a senior English major and was proficient in his native language. Third, the two versions were re-examined and resolved by a panel of 6 experts in the fields of health psychology, sports medicine, physical education and health, physical education, athletic training and exercise psychology. The team members are all native Chinese speakers who also speak English and have more than 5 years of work experience in their respective fields of expertise. Experts evaluated these versions, comparing each item to the corresponding item in the original English version. All discrepancies have been appropriately corrected. The CRAVE-C scales were further evaluated and scored by an expert panel to determine whether they were culturally appropriate for the Chinese population.

The translated final version of the scale requires an assessment of content clarity (Kuan, Kueh, Abdullah, & Tai, 2019).This study invited 15 high school students to evaluate the clarity of the final Chinese version. They commented on the content of the questionnaire and the clarity of each item. Their opinions are unanimous, and no revision was needed.

Procedure

This study used a cross-sectional study design, the study received approval from the Universiti Sains Malaysia (USM) Human Research Ethics Committee (USM/JEPeM/22110706) and followed all the procedures outlined by the Declarations of Helsinki.

This study has distributed questionnaires on site between December 2023 and January 2024, and participants signed informed consent forms before starting the study. The participant recruitment process included making and displaying posters, printing questionnaires, distributing questionnaires on-site, and collecting questionnaires on-site. By organizing school physical education teachers and head teachers to assist on-site supervision, students will fill in the questionnaires within the specified time and collect the questionnaires on the spot. In addition, after the first test, participants can choose to participate in the same study again, and the number of students and the list will be counted on the spot. After an interval of 7 days (a week, it will not be long enough to change the emotional state, nor short enough to clearly remember the content of the reply) (Wang, 2013), the researcher went to the site again to distribute the questionnaire and let the participants further fill in the questionnaire, and then received the retest questionnaire.

Statistical Analysis

Examination of the content validity of the CRAVE-C scale. Content validity is usually defined as a tool used to measure the appropriateness of questionnaire items. It requires at least 3 or no more than 10 experts (Lee, Waid, Tan, Andrew Chua, & Whitehead, 2012) who are responsible for evaluating and commenting on the relevance and understand ability of each item in the questionnaire and applying the calculation of the content validity index (CVI) in this step . In this study, 6 experts in related fields were invited to participate in the scoring, and the number was sufficient to meet the content validity test. I-CVI = (agreed item)/(number of rater), the acceptability range of three or more experts is 0.78 or higher, while the lower limit of acceptability of S-CVI is 0.80 (Polit & Beck, 2008). All items with consensus agreement above 0.80 were retained except for further analysis determined by experts.

This study used the statistical software Mplus 8.3 to conduct confirmatory analysis of the CRAVE-C scale. Pre-screening to identify incorrect data entry and missing values, the way to handle missing data is by using the listwise deletion method, which involves deleting all observations that contain missing values. Measurement data are expressed as mean ± standard deviation, and count data are expressed as frequency and percentage.

To ensure the reliability of the analysis, this study examined the data for multivariate normality. If the Mardia multivariate skewness ($p < .001$) and kurtosis ($p < .001$) are present, it indicates that the normality assumption is not met. Therefore, the robust maximum likelihood estimator (MLR) was chosen to handle such data in CFA and measurement invariance tests (Godin & Shephard, 1985).

The initial measurement model of the 13 items of the CRAVE-C scale was tested using CFA. Items with factor loadings below 0.40 are considered problem (DeVon et al., 2007). According to relevant studies, in order to determine the validity of a questionnaire, researchers usually propose multiple fit indices (Hair, 2006) . The recommended fit indices for the 13 items of the CRAVE-C scale in this study are: root mean square error of approximation (RMSEA) with expected value less than 0.07; comparative fit index (CFI) and Tucker and Lewis

index (TLI), with expected value greater than 0.90; Standardized root mean square residual (SRMR), the expected value is less than 0.08 (Hair, 2006). Refer to the CFA correction index for model renormalization to improve the model fitting index.

Reliability check. This study used composite reliability (CR) based on the Raykov method in Mplus. CR is similar to Cronbach's Alpha, indicating internal consistency. We set a minimum acceptable threshold of 0.60 and above for reliability (Tseng, Dörnyei, & Schmitt, 2006). Convergent validity was assessed using average variance extraction (AVE), with values above 0.5 indicating strong convergent validity (Fornell & Larcker, 1981).

Discriminant validity was assessed using correlation coefficients. Sufficient discriminant validity is established when factor correlations remain less than or equal to 0.85 (Browne & Cudeck, 1992). This study assessed test-retest reliability, including calculation of intraclass correlation coefficients (ICC), using a subsample of 89 participants. ICC values are divided into the following categories: less than 0.4 (poor), 0.4 to 0.59 (fair), 0.6 to 0.74 (good), and 0.75 or above (excellent) (Mason, Classen, Wersal, & Sisiopiku, 2021).

Measurement invariance. A multiple-group CFA was conducted to examine the invariance of CRAVE-C across gender in a high school student population. To verify sex invariance, a configured invariance model was built (Sabo, Kueh, Arifin, Kim, & Kuan, 2020). (1) Configural invariance. Config invariance model does not have gender equality restrictions on model parameters. (2) Weak invariance. Then construct and test the weak invariance. The weak invariance model imposes equal constraints on model factor loading across sex and ensures the similarity of measurement scales across sex to provide valid comparisons. (3) Strong invariance., developed and tested the strong invariance model. The strong invariance model imposes equal constraints on the factor loading and item intercept across sex, thus allowing cross-sex comparison of scale factors. (4) Strict invariance. developed and tested the strict invariance model (Zhao et al., 2024). To ensure that the error term variance remains constant over sex, the strict invariant model imposes equation constraints on factor loading, term intercept, and residual variance. This study used ΔCFI and ΔTLI to evaluate invariance, and if ΔCFI≤0.01 and ΔTLI≤0.01 were met, the invariance model was considered acceptable (Chen, 2007; Cheung & Rensvold, 2002).

Results

A total of 505 high school students participated in this study, aged 15 to 18 years old, 248 boys and 257 girls (as shown in Table I). There are three grades of students, including 178 in the first grade, 169 in the second grade, and 158 in the third grade.

Content validity

Based on the experts' ratings, the content validity index of I-CVI and S-CVI was calculated. The CVI assessment includes the relevance and representativeness of each item to a specific domain, where the I-CVI assesses each item in each domain and the S-CVI assesses the entire measurement scale (as shown in Table II).

TABLE I
The Demographic Characteristics (n = 505)

Characteristics	Frequency	Percentage	Mean ± SD
Gender			
Male	248	49.1%	
Female	257	50.9%	
Age (years)	15-18		16.7 ± 1.28
Class			
Grade one	178	35.2%	
Grade two	169	33.5%	
Grade three	158	31.3%	

TABLE II
Summary Cvi For All Of The Translated Constructs (n = 6 Experts)

Factor	Items	I-CVI	SCVI
CRAVE-C	13		
Moving the quantum scale	5	0.83-1.00	1
Rest subscale	5	0.83-1.00	1
Supplementary entry	3	1	1

As shown in Table II, the item-content validity index (I-CVI) of all items exceeded 0.80. The 6 experts had extremely consistent recognition of the scale items. high. For the moving subscale and the rest subscale, the I-CVI of some items ranged from 0.83 to 1.00. There were no noteworthy problems. Overall, the six experts had a high degree of consistency in the recognition of the items of the two scales.

MEASUREMENT MODEL FOR THE CRAVE-C

The initial model of CRAVE-C scale in this study consists of 13 items and 2 factors. Among them, items 1, 2, 6, 9 and 13 are set as the movement quantum scale, and items 3, 4 and 7 are set. Then, 8 and 10 are rest subscales, and 5, 11, and 12 are supplementary items (not involved in scoring). The

TABLE II BIS
Summary Of Rest And Volition Energy Expenditure Craving Scale-Crave-C Fitting Results (n = 505)

Path model	RMSEA (90% CI)	CFI	TLI	SRMR
Model-1	0.093 (0.084~0.103)	0.921	0.904	0.072
Model-2	0.038 (0.020~0.054)	0.992	0.990	0.022

results of the initially specified measurement model (Model-1) in this study showed a poor fit index (Table I). According to the analysis of the standardized regression coefficient results of the scale, the standardized factor loadings of all items of the moving subscale and the rest subscale are 0.40 and above, and the t test results of the standardized loadings meet the significance requirements ($p < 0.001$). However, the measurement item loadings of measurement items 5, 11, and 12 are relatively lower than 0.4 (Figure 1), and the standardized loading levels are 0.292, 0.207, and 0.139 respectively. Therefore, further fitting the model is considered after eliminating the above three items.

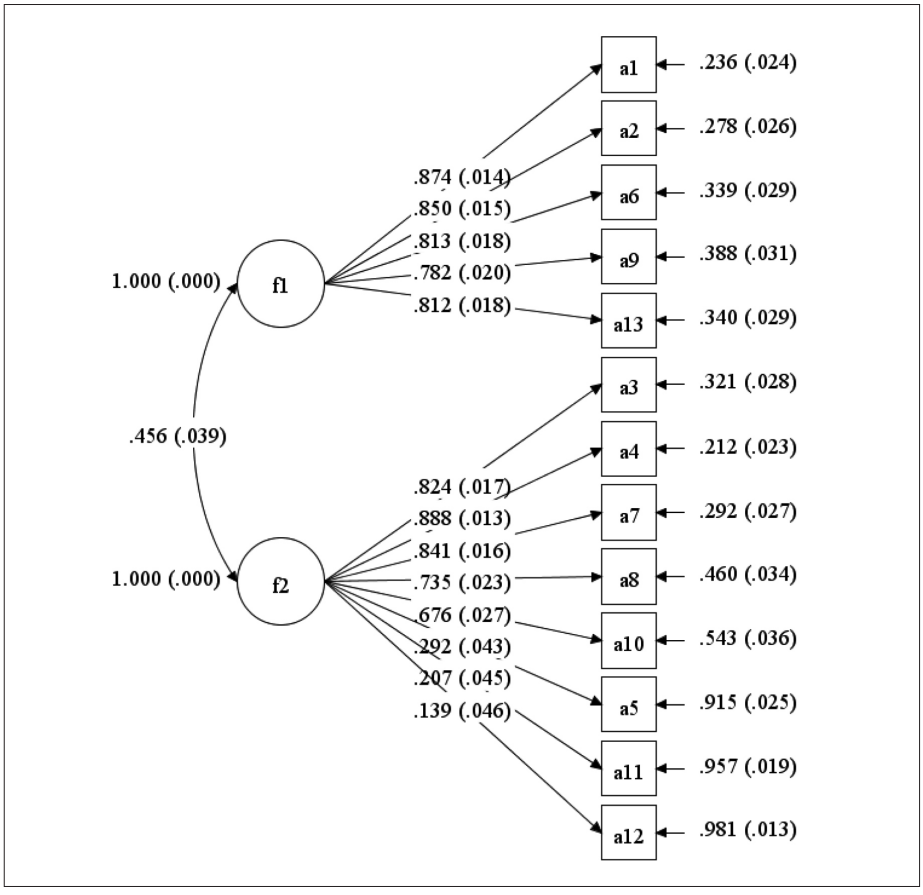

Figure 1. Measurement item loadings of the initial model.

Table II shows that the model fit index after removed items 5,11, and 12 is improved (Figure 2), and the fit index of the redesignated model (Model-2) is ideal (Table II), and all items are retained. The Model-2 results indicate a span level of 0.675 to 0.888, considered moderate to very good (Figure 2). Analysis the fitting results of the model fit index of the CRAVE-C scale, the RMSEA of Model-1 was 0.093 (0.084-0.103), CFI 0.921, TLI 0.904, SRMR 0.072, and RMSEA of Model-2 was 0.038 (0.020~0.054), CFI 0.992, TLI 0.990, and SRMR 0.022, indicating a good fit between the model and observed data.

Composite Reliability (cr) and Discriminant Validity

The CR of the scale was 0.915 for the moving subscale and the rest subscale was 0.896. The AVE for the mobile quantum scale and the rest sub-

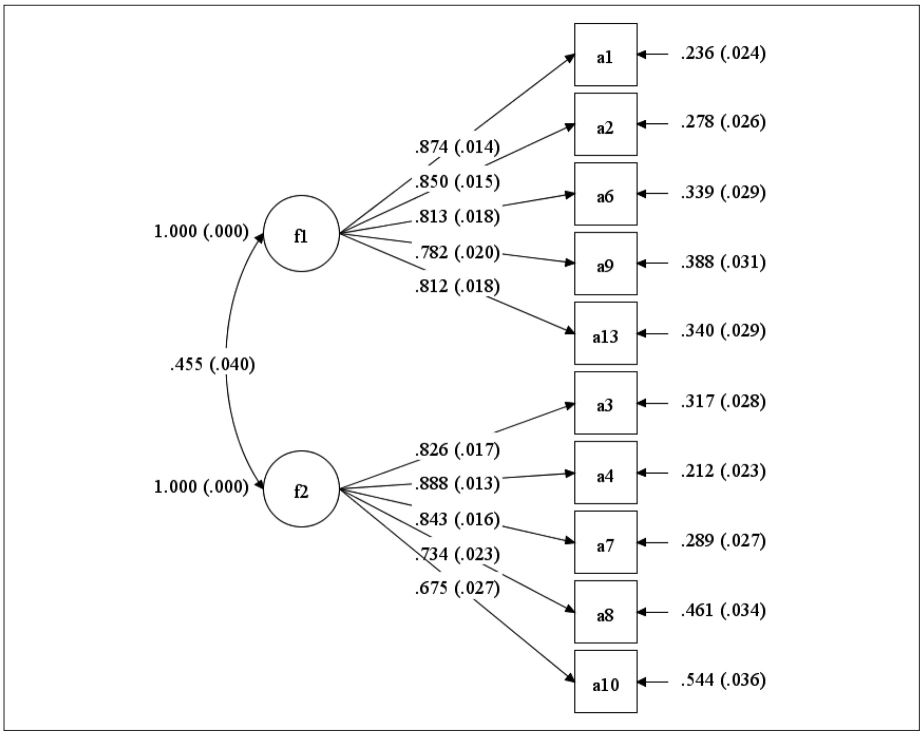

Figure 2. Measurement item loadings after removed

scales were 0.684 and 0.635, respectively. The CR and AVE values were higher than the recommended values of 0.60 and 0.50, respectively, thus concluding that the levels of convergent validity of the resting and volitional Energy-consuming Longing Scale-CRAVE-C were higher. The squared correlation coefficients for the moving quantum and rest subscales were 0.180, which is smaller than the AVE values of 0.684 and 0.635 for the moving quantum and rest subscales. Rest and volitional energy expenditure craving scale-CRAVE-C discriminant validity. CR and AVE values for rest and will energy expenditure craving scale-CRAVE-C and correlation coefficient and correlation coefficient squares are presented in Table III.

Test-retest reliability

In this study, the test-retest reliability of the CRAVE-C scale was tested by using 89 questionnaire data, and the scale consistency was measured by two tests (Test 1 and Test 2) and Intraclass Correlation Coefficient (ICC).

Table III
Composite Reliability (CR) And Discriminative Validity Results For CRAVE-C (n = 505)

Path model*	CR (95% CI)	Cronbach α	AVE	1	2	R2
F1	0.915 (0.782~0.874)	0.913	0.684	1	0.424**	0.180
F2	0.896 (0.675~0.888)	0.894	0.635		1	

Note: F1 is the moving subscale; F2 is the rest subscale.

Table IV
Test-Retest Reliability Results For CRAVE-C (n = 89)

	Test 1 Mean ± SD	Test 2 Mean ± SD	ICC (95% CI)
A1	5.506 ± 2.389	5.348 ± 2.384	0.874 (0.679,0.847)
A2	5.112 ± 2.156	5.225 ± 2.255	0.729 (0.417, 0.698)
A6	6.449 ± 2.398	6.640 ± 2.385	0.928 (0.802, 0.909)
A9	6.449 ± 2.722	6.461 ± 2.637	0.886 (0.704, 0.860)
A13	6.315 ± 2.054	6.461 ± 2.185	0.882 (0.696, 0.856)
A3	5.146 ± 2.284	5.022 ± 2.132	0.895 (0.724, 0.870)
A4	5.652 ± 2.079	5.640 ± 2.263	0.900 (0.736, 0.877)
A7	5.236 ± 2.190	5.270 ± 2.230	0.892 (0.719, 0.868)
A8	5.225 ± 2.325	5.079 ± 2.366	0.881 (0.695, 0.855)
A10	5.472 ± 2.106	5.303 ± 2.155	0.811 (0.553, 0.778)
Total score	56.5618 ± 14.380	56.450 ± 13.647	0.948 (0.901, 0.948)

Means and standard deviations of tests 1 and 2 varied across items (A1 to A10) but changed little overall. The ICC for items A1 to A10 was 0.729 to 0.928, respectively, corresponding to a 95% confidence interval (CI) of 0.417 to 0.909. The CRAVE-Cs scale has good test-retest reliability and is able to stably measure similar results even when tested at different time points. It indicates the reliability and stability of the scale in different contexts.

Measurement invariance

This study used multiple group confirmatory factor analysis to test the sex-invariance of the CRAVE-C scale. Four models are developed: configuration invariance, weak measurement invariance, strong measurement invariance, strict measurement invariance. The results are shown in Table V. The CFI and TLI for all four models and females exceeded the threshold of 0.95, indicating a satisfactory model fit. In addition, the RMSEA is within 0.08, which indicates good fitting and meets the proposed criteria for model fitting.

The weak measurement invariance model was compared with the configural model (ΔCFI = 0.001, ΔTLI = 0.003), and these differences indicated satisfactory metric invariance scores across genders. The strong invariance model showed sufficient invariance (ΔCFI = -0.002, ΔTLI = 0.001). The strict invariance model has sufficient invariance (ΔCFI = 0.001, ΔTLI = 0.001). The results showed that male and female had comparable average item scores. The above results indicate that the relationship between CRAVE-C scale variables remains consistent across genders, with no significant differences between male and female samples.

Table V
CRAVE-C Baseline Model Fitting Results And Invariance Test Results (n = 505)

Path model*	$\chi 2$	Df	$\chi 2$/df	RMSEA (90% CI)	CFI	TLI	SRMR	ΔCFI	ΔTLI
Male	59.329	34	1.745	0.055 (0.030~0.078)	0.984	0.979	0.029	-	-
Female	56.010	34	1.647	0.050 (0.025~0.073)	0.987	0.983	0.029	-	-
M0 (configural)	115.339	68	1.696	0.053 (0.035~0.069)	0.986	0.981	0.029	-	-
Ml (Weak)	119.064	76	1.567	0.047 (0.030~0.063)	0.987	0.984	0.035	0.001	0.003
M2 (Strong)	113.631	84	1.353	0.048 (0.032~0.063)	0.985	0.984	0.037	-0.002	0.001
M3 (Strict)	137.890	89	1.549	0.047 (0.031~0.061)	0.985	0.985	0.038	0.000	0.001

Discussion

In this study, the researcher translated the English version of the CRAVE scale into Chinese, and then used confirmatory factor analysis to determine the validity and reliability of the Chinese version of the CRAVE-C scale in high school students in Yunnan, China, analyzing their gender invariance in male and female samples. The results show that all the study data provide evidence for the validity and reliability of the Chinese version of the CRAVE-C scale, achieving the purpose of the study.

In this study, the Chinese version of the CRAVE-C scale validated has good content validity. The Item-Content Validity Index (I-CVI) exceeded 0.80 for all items. I-CVI= (agreed item) / (number of rater), with an acceptability range of 0.78 or higher for three or more specialists, while the lower acceptability limit for the S-CVI was 0.80, indicating adequate internal validity of the scale.

The method used by CRAVE-C is CFA. CFA can evaluate the factors of the model indicators as a whole. CFA can be used for a variety of purposes, such as psychometric assessment, detection of method effects, construct validation, and variance measurement assessment (Hair, 2006; Kuan, Sabo, Sawang, & Kueh, 2020).The results of CFA showed that the 10-item CRAVE scale showed good structural validity. This is consistent with the results reported in the original 10-item studies of the English version of the CRAVE Scale (Stults-Kolehmainen et al., 2020). They found that 13 versions of the CRAVE scale (5 moves, 5 rest, 3 filled items) showed a slightly worse fitted model. Using the original 10-item version (remove "burn some calories", "lie down" and "rest the body") (Stults-Kolehmainen et al., 2021). In this study, the results of 13 CRAVE scale initially specified measurement models (Model-1) showed poor fit index, with measurement items 5,11 and 12 below 0.40 and items with a factor load of 0.40 being considered problematic and problematic items removed (DeVon et al., 2007). The CRAVE-C scale model was reconstructed by removing fill items 5,11 and 2. The results showed that 10 CRAVE-C showed a better fit in the student sample compared to 13 CRAVE-C. The factor loading of each item showed satisfactory measurement characteristics. These findings provide preliminary evidence for the validity of the 10-item Chinese version of the CRAVE-C scale. Therefore, the 10-item Chinese version of the CRAVE-C scale was used for the subsequent analysis.

This study used a composite reliability (CR) to test internal consistency, assess convergence validity using mean variance extraction (AVE) based on the Raykov method in Mplus, and use correlation coefficient to assess discrimination validity, test-retest reliability, including the calculated with-

in-group correlation coefficient (ICC). The results showed that CRAVE-C showed satisfactory internal consistency and reliability. In the analysis, the CR coefficient of different factors was greater than 0.60, with good internal consistency (MacCallum, Browne, & Sugawara, 1996). The mean variance extraction (AVE) is greater than 0.50, and the convergence effectiveness is very strong (Jak, Jorgensen, Verdam, Oort, & Elffers, 2021). The correlation coefficient (ICC) is greater than 0.70, and the retest reliability is strong Errore. L'origine riferimento non è stata trovata.. Combined with the above CFA results, favorable evidence was obtained for the internal consistency reliability and structural validity of CRAVE-C, which is consistent with the results of similar previous studies (Stults-Kolehmainen et al., 2021).

Both males and females were involved in the baseline model. The configuration invariance model was then constructed by integrating the two baseline models with the same number of fixed and free factor loads (Sabo et al., 2020). (1) The configuration invariance model is very suitable for the data and has no equal constraints on any parameters across gender. (2) The weak measurement invariance model is well suited to the data and compared with the non-restrictive model (configuration model), and these differences indicate satisfactory measurement invariance scores across gender (ΔCFI = 0.001, ΔTLI = 0.003). The results show that these items have similar implications for both men and women. (3) The strong invariance model shows sufficient invariance (ΔCFI = -0.002, ΔTLI = 0.001). The results show that the factor load and intercept are constant by sex. (4) By applying equation constraints on factor load, item intercept and residual variance, the strict invariant model has sufficient variance (ΔCFI = 0.001, ΔTLI = 0.001). The findings showed that average item scores were comparable for men and women. All four models and the ΔCFI and ΔTLI were less than 0.01, and the findings consistently provide support for invariance at these levels (Classen, Wersal & Sisiopiku, 2021), indicating the stability of the CRAVE-C scale in different gender groups.

This study needs to consider some limitations and avenues for further inquiry. First, the participants in this study were only high school students in urban centers in Yunnan Province, China, and cannot represent all adolescent students. In order to enhance the applicability of the research results and confirm their robustness, future research may need to use more diverse samples, such as primary school students, junior high school students and other groups to be included in the research and may need to cover a wider range of regional representation, such as covering Rural and urban school students were included in the study. Second, given the satisfactory findings in a sample of high school students, the psychometric properties and measure-

ment invariance of the CRAVE-C scale in different groups can be examined in future studies.

Conclusion

The Chinese version of the CRAVE-C scale is considered to be valid and reliable for use among Chinese high school students. The 10-item CRAVE-C scale showed good model fit and measurement invariance across sex. Thus, this scale can be used as a tool to measure the physical activity and sedentary motivational state for that particular age group.

Acknowledgements

We would like to express our sincerest gratitude to all the participants, who has contributed to the study.

REFERENCES

Bames, J., Behrens, T. K., Benden, M. E., Biddle, S., Bond, D., Brassard, P., . . . Chaput, J. (2012). Standardized use of the terms" sedentary" and" sedentary behaviours". *Applied Physiology Nutrition and Metabolism-Physiologie Appliquee Nutrition Et Metabolisme, 37*, 540-542.

Biddle, S. J., Bennie, J. A., Bauman, A. E., Chau, J. Y., Dunstan, D., Owen, N., . . . Van Uffelen, J. G. (2016). Too much sitting and all-cause mortality: is there a causal link? *BMC Public Health, 16*(1), 1-10.

Biswas, A., Oh, P. I., Faulkner, G. E., Bajaj, R. R., Silver, M. A., Mitchell, M. S., & Alter, D. A. (2015). Sedentary time and its association with risk for disease incidence, mortality, and hospitalization in adults: a systematic review and meta-analysis. *Annals of Internal Medicine, 162*(2), 123-132.

Brislin, R. W. (1970). Back-translation for cross-cultural research. *Journal of Cross-Cultural Psychology, 1*(3), 185-216.

Browne, M. W., & Cudeck, R. (1992). Alternative ways of assessing model fit. *Sociological Methods & Research, 21*(2), 230-258.

Bull, F. C., Al-Ansari, S. S., Biddle, S., Borodulin, K., Buman, M. P., Cardon, G., . . . Chou, R. (2020). World Health Organization 2020 guidelines on physical activity and sedentary behaviour. *British Journal of Sports Medicine, 54*(24), 1451-1462.

Caspersen, C. J., Powell, K. E., & Christenson, G. M. (1985). Physical activity, exercise, and physical fitness: definitions and distinctions for health-related research. *Public Health Reports, 100*(2), 126.

Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling, 14*(3), 464-504.

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling, 9*(2), 233-255.

DeVon, H. A., Block, M. E., Moyle-Wright, P., Ernst, D. M., Hayden, S. J., Lazzara, D. J., . . . Kostas-Polston, E. (2007). A psychometric toolbox for testing validity and reliability. *Journal of Nursing Scholarship, 39*(2), 155-164.

Duncan, L. R., Hall, C. R., Wilson, P. M., & Jenny, O. (2010). Exercise motivation: a cross-sectional analysis examining its relationships with frequency, intensity, and duration of exercise. *International Journal of Behavioral Nutrition and Physical Activity, 7*, 1-9.

Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research, 18*(1), 39-50.

Gabriel, K. K. P., Morrow, J. R., & Woolsey, A.L.T. (2012). Framework for physical activity as a complex and multidimensional behavior. *Journal of Physical Activity and Health, 9*(s1), S11-S18.

Geneva, S. (2010). Global Recommendations on Physical Activity for Health.

Hair, J. (2006). Black W. Babin B., Anderson R. Tatham R. *Multivariate Data Analysis, 6*.

Jak, S., Jorgensen, T. D., Verdam, M. G., Oort, F. J., & Elffers, L. (2021). Analytical power calculations for structural equation modeling: A tutorial and Shiny app. *Behavior Research Methods, 53*, 1385-1406.

Kuan, G., Kueh, Y. C., Abdullah, N., & Tai, E. L. M. (2019). Psychometric properties of the health-promoting lifestyle profile II: cross-cultural validation of the Malay language version. *BMC Public Health, 19*, 1-10.

Kuan, G., Sabo, A., Sawang, S., & Kueh, Y. C. (2020). Factorial validity, measurement and structure invariance of the Malay language decisional balance scale in exercise across gender. *PloS One, 15*(3), e0230644.

Lee, Y. Y., Waid, A., Tan, H. J., Andrew Chua, S. B., & Whitehead, W. E. (2012). Validity and reliability of the Malay-language translation of the Rome III Diagnostic Questionnaire for irritable bowel syndrome. *Journal of Gastroenterology and Hepatology, 27*(4), 746-750.

Levine, J. A., Eberhardt, N. L., & Jensen, M. D. (1999). Role of nonexercise activity thermogenesis in resistance to fat gain in humans. *Science, 283*(5399), 212-214.

MacCallum, R. C., Browne, M. W., & Sugawara, H. M. (1996). Power analysis and determination of sample size for covariance structure modeling. *Psychological Methods, 1*(2), 130.

Mason, J., Classen, S., Wersal, J., & Sisiopiku, V. (2021). Construct validity and test-retest reliability of the automated vehicle user perception survey. *Frontiers in Psychology, 12*, 626791.

Owen, N., Healy, G. N., Matthews, C. E., & Dunstan, D. W. (2010). Too much sitting: the population health science of sedentary behavior. *Exercise and Sport Sciences Reviews, 38*(3), 105-113.

Piercy, K. L., Troiano, R. P., Ballard, R. M., Carlson, S. A., Fulton, J. E., Galuska, D. A., . . . Olson, R. D. (2018). The physical activity guidelines for Americans. *JAMA, 320*(19), 2020-2028.

Polit, D. F., & Beck, C. T. (2008). *Nursing research: Generating and assessing evidence for nursing practice*: Lippincott Williams & Wilkins.

Rezende, L. F. M. d., Rey-López, J. P., Matsudo, V. K. R., & Luiz, O. d. C. (2014). Sedentary behavior and health outcomes among older adults: a systematic review. *BMC Public Health, 14*, 1-9.

Romero-Parra, N., Solera-Alfonso, A., Bores-García, D., & Delfa-De-La-Morena, J. M. (2023). Sex and educational level differences in physical activity and motivations to exercise among Spanish children and adolescents. *European Journal of Pediatrics, 182*(2), 533-542.

Sabo, A., Kueh, Y. C., Arifin, W. N., Kim, Y., & Kuan, G. (2020). The validity and reliability of the Malay version of the social support for exercise and physical environment for physical activity scales. *PloS One, 15*(9), e0239725.

Sevil, J., García-González, L., Abós, Á., Generelo, E., & Aibar, A. (2019). Can high schools be an effective setting to promote healthy lifestyles? Effects of a multiple behavior change intervention in adolescents. *Journal of Adolescent Health, 64*(4), 478-486.

Stults-Kolehmainen, M. A., Blacutt, M., Fogelman, N., Gilson, T. A., Stanforth, P. R., Divin, A. L., . . . Ash, G. I. (2021). Measurement of motivation states for physical activity and sedentary behavior: development and validation of the CRAVE scale. *Frontiers in Psychology, 12*, 568286.

Stults-Kolehmainen, M. A., Blacutt, M., Gilson, T. A., Stanforth, P. R., Divin, A. L., Bartholomew, J. B., ... & Sinha, R. (2020). I want to move my body-right now! The CRAVE Scale to measure state motivation for physical activity and sedentary behavior. *MedrXiv.*

Tcymbal, A., Andreasyan, D., Whiting, S., Mikkelsen, B., Rakovac, I., & Breda, J. (2020). Prevalence of physical inactivity and sedentary behavior among adults in Armenia. *Frontiers in Public Health, 8*, 157.

Tseng, W.-T., Dörnyei, Z., & Schmitt, N. (2006). A new approach to assessing strategic learning: The case of self-regulation in vocabulary acquisition. *Applied Linguistics, 27*(1), 78-102.

Wang, Q. (2013). Gender and emotion in everyday event memory. *Memory, 21*(4), 503-511.

Wilmot, E. G., Edwardson, C. L., Achana, F. A., Davies, M. J., Gorely, T., Gray, L. J., . . . Biddle, S. J. (2012). Sedentary time in adults and the association with diabetes, cardiovascular disease and death: systematic review and meta-analysis. *Diabetologia, 55*(11), 2895-2905.

Zhao, M., Kuan, G., Zhou, K., Musa, R. M., Majeed, A. P. A., & Kueh, Y. C. (2024). Psychometric properties and gender invariance of the 8-item emotion regulation questionnaire (ERQ-8) among Chinese university students. *PloS One, 19*(1), e0296035.

Manuscript accepted for publication September 2024.