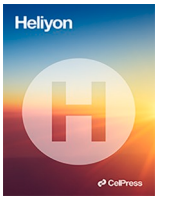
Heliyon 9 (2023) e17623

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon



Validity and Reliability of the Arabic Version of the Three-Factor Eating Questionnaire-R18

Salma Alhebshi [a], Serene Hilary [a], Samir K.H. Safi [b], Habiba I. Ali [a], Leila Cheikh Ismail [c,d], Ayesha Al Dhaheri [a], Lily Stojanovska [a,e,*]

[a] *Department of Nutrition and Health, College of Medicine and Health Sciences, United Arab Emirates University, PO Box, 15551, Al Ain, United Arab Emirates*
[b] *Department of Statistics and Business Analytics, College of Business and Economics, United Arab Emirates University, PO Box, 15551, Al Ain, United Arab Emirates*
[c] *Department of Clinical Nutrition and Dietetics, College of Health Sciences, University of Sharjah, Sharjah, 27272, United Arab Emirates*
[d] *Nuffield Department of Women's and Reproductive Health, University of Oxford, Oxford, United Kingdom*
[e] *Institute for Health and Sport, Victoria University, Melbourne, VIC 3011, Australia*

ARTICLE INFO

Keywords:
TFEQ-R18
Three-Factor Eating Questionnaire
Eating behaviour
Obesity
Validation
Arabic translation

ABSTRACT

The present study translated and validated an Arabic version of the Three-Factor Eating Questionnaire-R18 (TFEQ-R18), which assesses eating behaviour traits among Arabic speakers in the three domains: cognitive restraint (CR), uncontrolled eating (UE), and emotional eating (EE). Three bilingual experts first translated the English questionnaire into Arabic, and the questionnaire's comprehension was tested among 298 Arabic speakers. Confirmatory factor analysis (CFA) tested the model fit and flagged four items in the questionnaire with low consistency. These items were modified, and the revised Arabic version of TFEQ-R18 was validated in a population of 513 individuals. Validity was evaluated using CFA and Pearson's correlation coefficients, internal consistency using Cronbach's α, and the reproducibility of the questionnaire was confirmed with Bland-Altman analysis and T-test. The construct validity of the Arabic TFEQ-R18 was evaluated by comparing the response score in the population based on gender, BMI and age group. CFA confirmed that the model fit is good and strongly agrees with the collected data. Except for item 15 in the CR domain, all questions showed moderate to high correlation within their respective domains. The Cronbach's α for UE, EE, and CR domains recorded was 0.778, 0.784 and 0.588, respectively. Item 15 had the lowest consistency in the Arabic-TFEQ-R18, and Cronbach's α increased to 0.608 with its removal. There was no significant difference between the first and second attempts of the Arabic TFEQ-R18, indicating good test-retest reliability. Moreover, UE and EE were positively correlated to the BMI of the participants ($r = 0.159$, $p = 0{\cdot}000$; $r = 0.158$, $p = 0{\cdot}000$, respectively). The study concludes that the Arabic TFEQ is a valid and reliable tool for studying the three psychometric domains of UE, EE and CR among Arabic speakers.

1. Introduction

The high prevalence of obesity and overweight individuals is a growing concern worldwide. The global estimates of obesity have

* Corresponding author: Lily Stojanovska
E-mail address: lily.stojanovska@uaeu.ac.ae (L. Stojanovska).

https://doi.org/10.1016/j.heliyon.2023.e17623
Received 21 June 2022; Received in revised form 30 May 2023; Accepted 5 June 2023
Available online 2 July 2023

tripled since 1975, and as per estimates from the World Health Organization (WHO), in 2016, more than 1.9 billion adults were overweight or obese [1]. The increasing prevalence of obesity is also evident in the Gulf Cooperation Council (GCC), consisting of the seven Arab nations; Bahrain, Kuwait, Qatar, Oman, Saudi Arabia, and the United Arab Emirates (UAE). 48% of males and 35% of females were overweight, and 24% of males and 40% of females were reported to be obese in the GCC region [2]. The rise in metabolic disorders in the UAE is particularly alarming [3]. The intake of calorie-dense, processed, and prepackaged meals of poor nutritional value in the population is high owing to the shift from traditional dietary habits [4]. As per the UAE national action plan to combat non-communicable diseases published in 2017, the prevalence of overweight and obesity in the population aged 18 and above is 74% and 37.2%, respectively [5]. The same report also indicated that 14.7% of the population in the country is living with hyperglycaemia.

Several strategies, such as local, state and national level health efforts to spread awareness of the importance of physical activity and healthy diets, have been initiated over the years to combat the growing obesity epidemic [1,6]. However, the cause of obesity is more complex, involving psychological and behavioural factors [7]. Earlier studies have reported eating behaviour traits as a strong predictor for weight gain in a population [8–10], although the mechanistic role is poorly understood. Hence, for the success of community interventions to combat obesity, it is vital to understand the interaction between psychometric factors and eating behaviour traits that govern food choices and individuals' energy intake levels. Eating behaviour traits such as cognitive dietary restraint, a conscious restriction of dietary intake to manage body weight, have been associated with a higher intake of healthy foods such as green vegetables [8,9]. Disinhibition or uncontrolled eating is another eating behavioural trait associated with a higher intake of energy-dense foods [10]. In addition to cognitive dietary restraint for body image and uncontrolled eating, understanding other eating behaviour traits is vital to the success of weight-related interventions, such as overeating triggered by negative emotions [11]. Several psychometric tools are available to study eating behaviour, such as the Three-Factor Eating Questionnaire, The Dutch Eating Behavior Questionnaire, and the Restraint Scale. Among these tools, the Three-Factor Eating Questionnaire (TFEQ), a 51-item measure of restraint, disinhibition and hunger subscales published by Stunkard and Messick, was the most popular tool used to study eating behaviour [12].

In 2000, the Three-Factor Eating Questionnaire revised 18-item (TFEQ-R18) version published by Karlsson et al. demonstrated that shortening the questionnaire did not alter its objective or validity [13]. The TFEQ-R18 is a more popular tool nowadays, and it consists of three domains: uncontrolled eating (UE), emotional eating (EE), and cognitive restraint (CR). EE is the tendency to eat food triggered to control negative emotions. UE is the tendency to eat more than usual and the inability to control binge eating. CR is cognitive dietary restrain, which is the conscious control of diet intake to control body weight and achieve weight loss. The UE, EE and CR domains of the

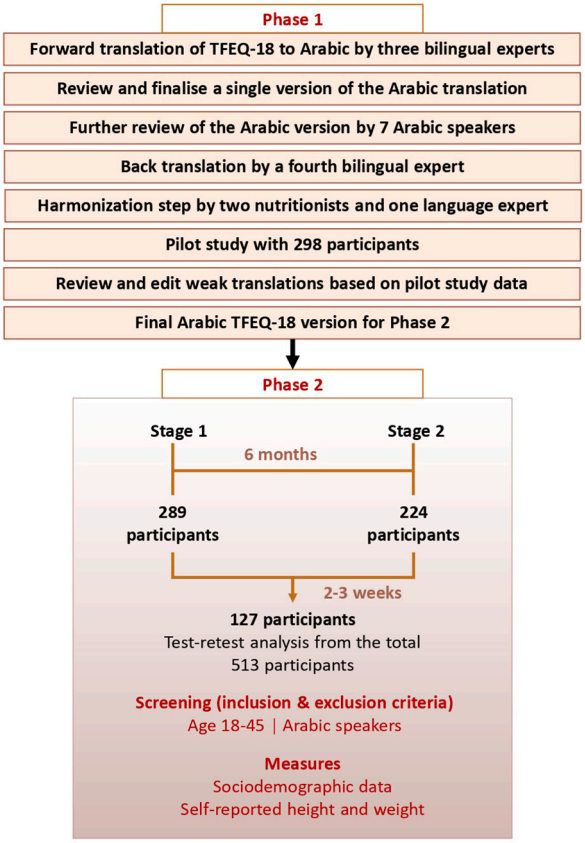


Fig. 1. Illustration of the study protocol.

TFEQ-R18 contain nine, three and six items, respectively. Over the years, TFEQ-R18 has been translated and adapted to various populations for assessing eating behaviour [14–17]. Given the high prevalence of obesity in GCC countries, it is essential to have a validated Arabic version of TFEQ-R18 for conducting similar studies among Arabic speakers. However, currently, such a version does not exist. Hence, the primary objective of this study was to translate and validate an Arabic version of the TFEQ-R18. This study aimed to analyse the Arabic TFEQ-R18's factor structure, internal consistency, construct validity (the correlation between the response and BMI of participants) and test-retest reliability.

2. Methods

2.1. Study participants and ethics approval

The study was conducted at the United Arab Emirates University among the university staff and student community. The ethics approval for the study was obtained from the United Arab Emirates University Human Research Ethics Committee (approval code: ERH-2020-6071). The study was executed in two phases. Phase 1 of the study involved translating the English questionnaire to Arabic according to the translation and cultural adaptation principles [18]. The translated questionnaire was initially tested among 298 participants to evaluate its comprehension. In phase 2 of the study, the finalised Arabic questionnaire was validated among 513 Arabic-speaking adults (Fig. 1). Participants for phases 1 and 2 were recruited from the university community using online communication platforms, and inclusion criteria were both genders in the age range of 18–60 years and Arabic speakers. The sample size was determined using Cochran's formula with a 95% confidence interval, 5% significance level (0.05) and marginal error of 0.04 [19]. Non-Arabic speakers were excluded.

2.2. Phase 1: Translation of the TFEQ-R18 and evaluation of the comprehension and acceptability of the questionnaire

The English TFEQ-R18 was translated and culturally adapted to Arabic by three bilingual nutritionists. The translation team reviewed the original English questionnaire's content clarity and adapted linguistic variation with its Arabic translation. Once the initial translation was finalised among the translation team members, sequential revisions of the content validity were performed by seven Arabic speakers on independent occasions. The Arabic TFEQ-R18 was then back-translated to English by one bilingual expert who was not involved in the forward translation step. The Arabic TFEQ-R18 was then finalised by the harmonisation step led by two nutritionists and a language expert to compare the translations with the original and modify the mistranslations in the Arabic version. It was then administered among 298 participants to identify potential weak questions in conveying the nuances of the English version. Based on the data from the pilot study, the Arabic TFEQ-R18 was modified and finalised for the Phase-2 validation study.

2.3. Phase 2: Validation of the Arabic TFEQ-R18

2.3.1. Study protocol

Phase 2 of the study was conducted in two stages (Fig. 1). In the first study stage, 289 Arabic-speaking adults were recruited, and the modified Arabic TFEQ-R18 from the Phase 1 study was administered. In the second stage of the study, 224 Arabic speakers (both males and females between the ages of 18–60) were recruited after a gap of 6 months to consider any potential variations in dietary intake that may occur due to a seasonal variation. The Arabic TFEQ-R18 was a self-administered online questionnaire. The questionnaire also collected sociodemographic data of the participants along with self-reported height and weight measures to compute BMI. We also collected contact information from participants willing to answer the questionnaire a second time for test-retest reliability analysis. Among those who consented to a second attempt, 127 participants representing one-third of the total 513 participants from stages 1 and 2, were contacted after 2–3 weeks of their initial attempt to answer the questionnaire a second time.

2.3.2. Administration of the Arabic TFEQ-R18

The Three-Factor Eating Questionnaire (TFEQ) is a scale with 18 questions that are categorised into three domains: cognitive restraint (CR), uncontrolled eating (UE), and emotional eating (EE). The Arabic TFEQ-R18 assesses these three domains; UE has 9 items, EE has 3 items, and CR has 6 items (supplementary document 1). Items 1–17 are measured on a four-point Likert scale, and item 18 is measured on an eight-point Likert scale. The eight-point scale of item 18 is converted to a four-point scale in the following format (1–2 = 1; 3–4 = 2; 5–6 = 3; 7–8 = 4). The mean score for each item was calculated using the coding procedure described by Anglé et al. for data analysis [20]. Items 14,15,16, and 17 in the questionnaire are coded in reverse, while others are coded normally.

2.3.3. Anthropometric measures

The participants self-reported body weight and height at the time of data collection. Hodge et al. had earlier tested the validity of using self-reported height and body weight measures [21]. BMI was calculated using the formula weight (kg) divided by height squared (m^2). Participants were then classified into four weight categories as follows: underweight (BMI <18.5 kg/m^2), normal weight (BMI 18.5–24.9 kg/m^2), overweight (BMI 25–29.9 kg/m^2), obese (BMI >30 kg/m^2).

2.4. Statistical analyses

Statistical analyses for the experiments were performed using SPSS® version 26 (IBM Corporation, Armonk, NY, USA). Descriptive

statistics (Mean ± SD) were used to describe continuous variables, and categorical variables were expressed as counts and percentages (N, %). Kaiser-Meyer-Olkin and Bartlett's sphericity test analyzed the data from the Phase 1 study, after which confirmatory factor analysis (CFA) was carried out to verify the factorial structure of the model using AMOS software. For Phase 2 data, CFA was repeated, and item-to-total correlations were calculated to test the validity of the modified Arabic TFEQ-R18 in stages 1 and 2. Correlation coefficients of BMI and the three domains were also used to explain the construct validity of Arabic TFEQ-R18. The reliability of Arabic TFEQ-R18 was checked using Cronbach's α coefficients. Test-retest reliability of the Arabic TFEQ-R18 was analyzed by plotting Bland-Altman's and conducting a T-test of the two attempts. Normality of domain response scores was checked by the Kolmogorov-Smirnov test, which showed that the data were not normally distributed. Consequently, the non-parametric Kruskal-Wallis H test was used to assess the difference between the response score in the population divided based on gender, age group and BMI. A p-value <0.05 was considered statistically significant.

3. Results

3.1. Sociodemographic characteristics of participants

Table 1 details the main characteristics of the participants recruited for Phase 1 and Phase 2 of the validation study. Most participants were UAE nationals (85.9% in Phase 1 and 74.7%in Phase 2), while the rest, 14.1% in Phase 1 and 25.3% in Phase 2, comprised expatriate Arabic speakers. We had a significantly higher proportion of females recruited in the study than males, with females making up 86.9% in Phase 1 and 84.6% in Phase 2. The number of male participants in the study was limited to 39 (13.1%) in Phase 1 and 79 (15.4%) in Phase 2. The study population was also significantly younger, as 92.3% in Phase 1 and 88.9% in Phase 2 were between 18 and 26 years. The height and weight of the participants were not recorded in Phase 1 of the study. Among the 513 participants of Phase 2, 13.5% were underweight, 53.6% were normal weight, 22% were overweight, and 10.9% were obese.

3.2. Factorial structure of the Arabic TFEQ-R18 in Phase 1

Data from the Phase 1 study of the Arabic version of TFEQ-R18 are provided in Tables 2 and 3. The data suitability was assessed using the measure of sampling adequacy. The KMO and Bartlett's test of sphericity are shown in Table 2. The KMO measure of sampling adequacy was 0.809, 0.605 and 0.588 for UE, EE, and CR, respectively. The KMO values for all domains were higher than 0.5; hence the factor analysis was appropriate for this data. Bartlett's test was highly significant for all domains (p-value <0.001), indicating that the data was appropriate for CFA. Table 3 shows the factor loadings for the three-factor domain of Arabic TFEQ-R18 in Phase 1. The first factor, the UE domain, accounted for 33.29% of the total variance and comprised 9 items, of which item 9 ranked the highest and item 17 ranked the lowest. The second factor in the model was the EE domain which accounted for 67.23% of the total variance and comprised 3 items. Most items in this domain had relatively high factor loadings (≥ 0.802), with item 3 ranking the highest. The third factor in the model was the CR domain which accounted for 32.34% of the total variance and comprised 6 items. The majority of attributes had relatively high factor loadings (≥ 0.65). The highest ranked in this domain was item 11, and the lowest was item 2. The factor structure for the three-factor model with 18 items failed as the data from the pilot study could not uniquely estimate model parameters making it difficult to assess the model's fit to the data. Fig. 2 illustrates the factor structure of the Arabic questionnaire with 14 items after removing the four items from the CR domain (items 2, 15, 16 and 18) based on the analysis of pilot study data of 298 participants. These items were flagged for review and modification before phase 2 of the study. Table 4 provides the results of the model fit indices for the pilot study. The results indicate a ratio of the chi-square to degrees of freedom of 1.414, which is within the acceptable range for a good model fit. Additionally, the comparative fit index (CFI) of 0.968, Tucker & Lewis index (TLI) of 0.957, the

Table 1
Sociodemographic characteristics of participants in the study.

	Phase 1	Phase 2
	N = 298	N = 513
Nationality		
UAE	256 (85.9)	383 (74.7)
Others	42 (14.1)	130 (25.3)
Gender		
Female	259 (86.9)	434 (84.6)
Male	39 (13.1)	79 (15.4)
Age (years)		
18–26	275 (92.3)	456 (88.9)
27–35	15 (5.0)	29 (5.7)
36–45	8 (2.7)	28 (5.5)
BMI (kg/m^2)		
Underweight (BMI <18.5 kg/m^2)		69 (13.5)
Normal (BMI 18.5–24.9 kg/m^2)		275 (53.6)
Overweight (BMI 25–29.9 kg/m^2)		113 (22)
Obese (BMI >30 kg/m^2)		56 (10.9)

Data presented as N(%).

Table 2
Kaiser-Meyer-Olkin test and Bartlett's test of sphericity.

		UE Domain	EE Domain	CR Domain
KMO measure of Sampling Adequacy		0.809	0.685	0.588
Bartlett's Test of Sphericity	Approx. Chi-Square	452.882	216.278	225.397
	Df	36	3	15
	p-value	<0.001	<0.001	<0.001

Data from Phase 1 study of 298 participants. CR-cognitive restraint, UE-uncontrolled eating, and EE-emotional eating. Statistical significance set at $p \leq 0.05$.

Table 3
TFEQ-R18 sub-scales and factorial weights for items (N = 298).

UE Domain		EE Domain		CR Domain	
Variance = 33.29%		Variance = 67.23%		Variance = 32.34%	
Item	Factor loadings	Item	Factor loadings	Item	Factor loadings
Item 9	0.767	Item 3	0.849	Item 11	0.765
Item 13	0.707	Item 6	0.808	Item 12	0.714
Item 4	0.637	Item 10	0.802	Item 16	0.588
Item 1	0.593			Item 18	0.547
Item 8	0.553			Item 15	0.425
Item 7	0.537			Item 2	−0.140
Item 5	0.508				
Item 14	0.440				
Item 17	0.323				

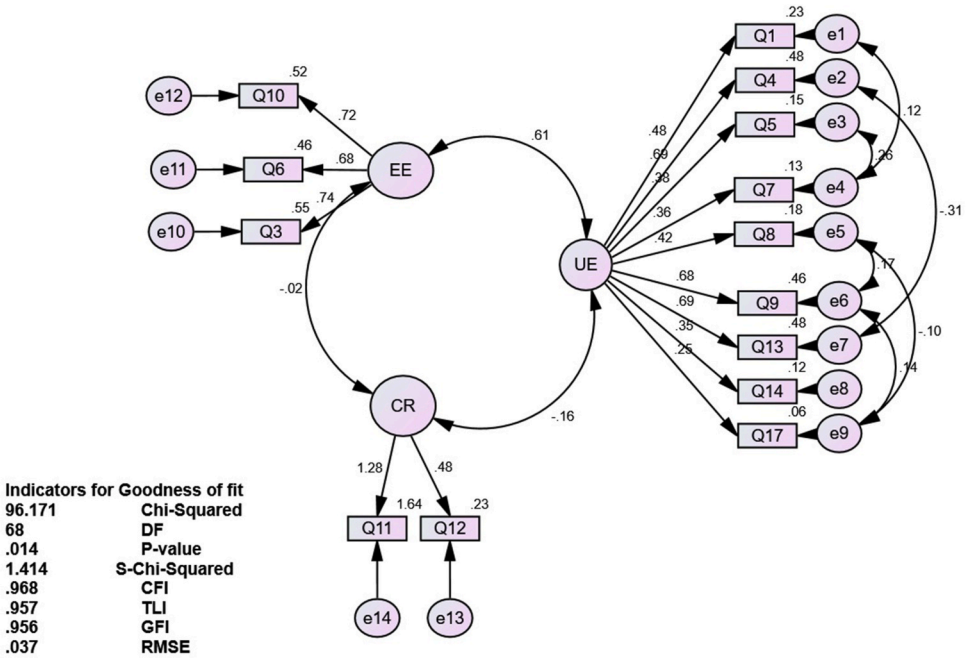


Fig. 2. Factorial Structure of the Arabic Questionnaire from the Pilot Study. CFI-comparative fit index, TLI-Tucker & Lewis index, RMSE- Root Mean Square Error. Data from Phase 1 study of 298 participants.

Table 4
Summary of fit indices.

Index	Model 14 Items (N = 298)	Model 18 Items (N = 513)
Scaled Chi-Squared	1.414	3.172
Comparative Fit Index (CFI)	0.968	0.882
Tucker-Lewis Index (TLI)	0.957	0.856
Goodness of Fit Index (GFI)	0.956	0.919
Root Mean Square Error (RMSE)	0.037	0.065

goodness of fit index (GFI) of 0.956, and Root Mean Square Error (RMSE) of 0.037 all demonstrated strong agreement between the model and data.

3.3. Validity of Arabic TFEQ-R18

The validity of the Arabic TFEQ-R18 questionnaire was estimated by repeating CFA and calculating the item-to-correlations in Phase 2 of the study. Fig. 3 displays the factor structure results of data from 513 participants in the second phase. The four items flagged for review (items 2, 15, 16 and 18) in the pilot stage were modified and added to the CR domain of the model, and CFA was repeated to evaluate the model fit of the 18-item configuration. From Table 4, we observe that the addition of four items to the cognitive restraint domain resulted in a chi-square to degrees of freedom ratio of 3.172, which is still within the range considered acceptable for a good model fit. The study results prove that the model fit is good and strongly agrees with the collected data, regardless of the number of items included. The indices of CFI at 0.882, TLI at 0.856, GFI at 0.919, and RMSE at 0.065 also suggest a good model fit. Therefore, these findings indicate that the inclusion of the four items did not significantly impact the overall goodness of fit of the model.

Pearson's correlation coefficients for all three domains are provided in Table 5. The table describes the correlation coefficients of the two stages of the study protocol separately and combined, where stage 1 had 224 participants and stage 2, conducted six months later, had 289 participants. The correlation coefficient of the two stages was similar. From the analysis, item 6 of the EE domain recorded the highest correlation in stages 1, 2 and total with coefficients 0.886, 0.856 and 0.870, respectively. The lowest correlation was observed for item 15 of the CR domain in stages 1, 2 and total (0.323, 0.331 and 0.328, respectively). All correlation coefficient values calculated were significant, with p-values of less than 0.01, indicating that this model's questions are consistent and objective,

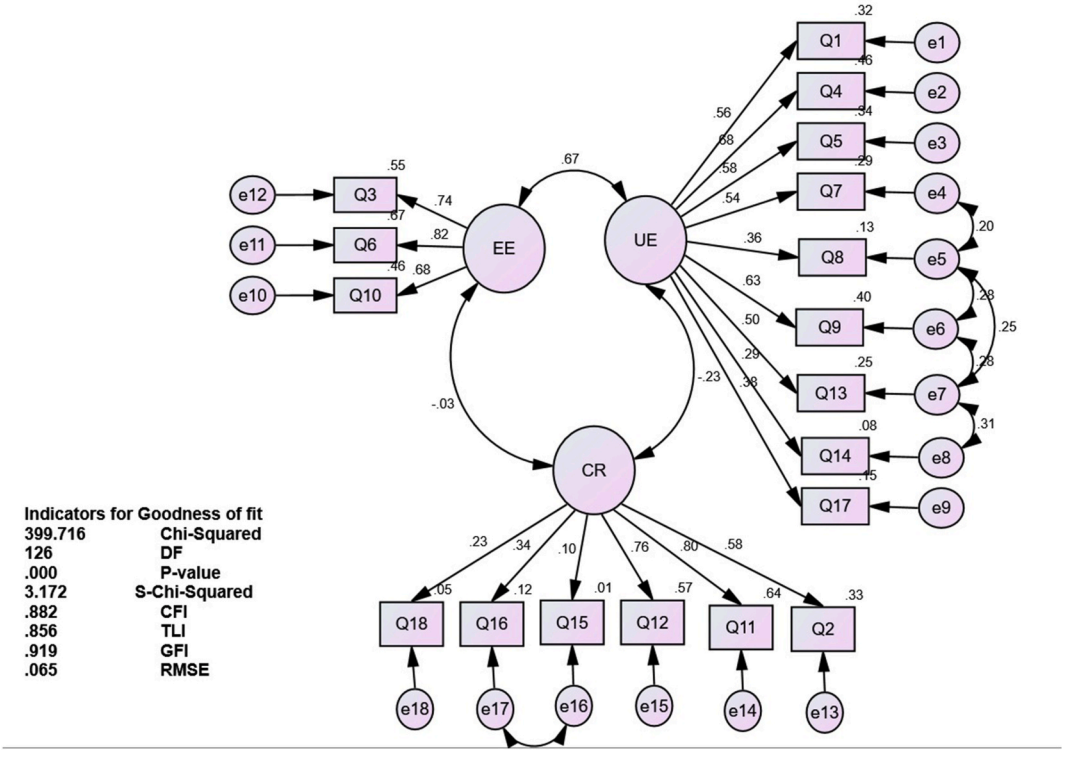


Fig. 3. Factorial Structure of the Arabic TFEQ-R18. CFI-comparative fit index, TLI-Tucker & Lewis index, RMSE- Root Mean Square Error. Data from Phase 2 study of 513 participants.

Table 5
Pearson's correlation coefficient of Arabic TFEQ-R18.

		Stage 2 (N = 289)		Stage 1 (N = 224)		Total (N = 513)	
	Items	*r*	p-value[1]	*r*	p-value[1]	*r*	p-value[1]
UE Domain	Item 1	0.612**	0.000	0.588**	0.000	0.601**	0.000
	Item 4	0.650**	0.000	0.679**	0.000	0.661**	0.000
	Item 5	0.560**	0.000	0.622**	0.000	0.586**	0.000
	Item 7	0.611**	0.000	0.618**	0.000	0.614**	0.000
	Item 8	0.552**	0.000	0.654**	0.000	0.599**	0.000
	Item 9	0.718**	0.000	0.753**	0.000	0.732**	0.000
	Item 13	0.695**	0.000	0.665**	0.000	0.681**	0.000
	Item 14	0.492**	0.000	0.462**	0.000	0.478**	0.000
	Item 17	0.441**	0.000	0.450**	0.000	0.445**	0.000
EE Domain	Item 3	0.811**	0.000	0.868**	0.000	0.839**	0.000
	Item 6	0.856**	0.000	0.886**	0.000	0.870**	0.000
	Item 10	0.784**	0.000	0.814**	0.000	0.798**	0.000
CR Domain	Item 2	0.609**	0.000	0.606**	0.000	0.607**	0.000
	Item 11	0.711**	0.000	0.707**	0.000	0.709**	0.000
	Item 12	0.701**	0.000	0.698**	0.000	0.700**	0.000
	Item 15	0.331**	0.000	0.323**	0.000	0.328**	0.000
	Item 16	0.499**	0.000	0.581**	0.000	0.534**	0.000
	Item 18	0.597**	0.000	0.601**	0.000	0.599**	0.000

Pearson's correlation coefficient (*r*). [1]Statistical significance set at $p \leq 0.05$, **significant correlation. CR-cognitive restraint, UE-uncontrolled eating, and EE-emotional eating.

except item 15.

3.4. Reliability of Arabic TFEQ-R18

Cronbach's α for the Arabic TFEQ-R18 was 0.625, and the value for the UE, EE and CR domains were 0.778, 0.784 and 0.588, respectively. These results indicate good reliability for UE and EE domains. The reliability of the CR domain was poor and can be attributed to item 15, which also recorded a lower item-to-total correlation. We reanalysed the data after excluding item 15, which

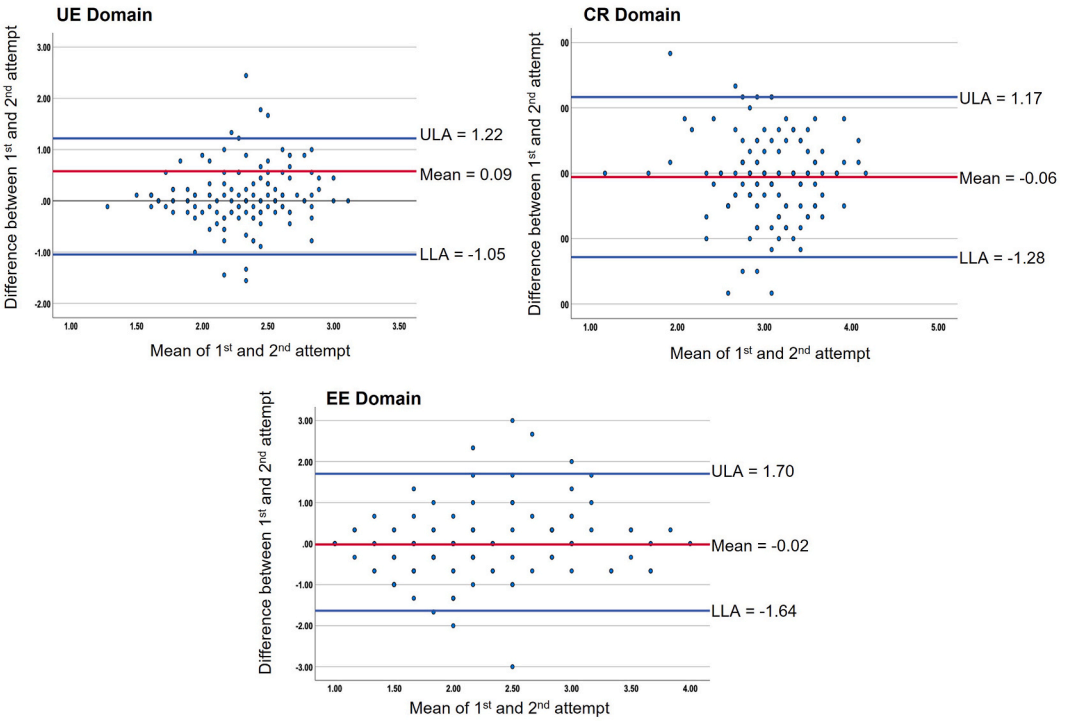


Fig. 4. Reproducibility of Arabic TFEQ-R18 with Bland-Altman Analysis. The mean difference between the scores of the 1st attempt (N = 513) and 2nd attempt (N = 127) was plotted against the mean of the two attempts for each domain of the TFEQ-R18 questionnaire in Arabic. ULA-the upper limit of agreement, LLA-the lower limit of agreement.

resulted in an increase in Cronbach's α value to 0.608, and the reliability of the whole model increased from 0.625 to 0.641. The test-retest reliability of Arabic TFEQ-R18 demonstrated the reproducibility of the scale when one-third of participants attempted the questionnaire a second time after 2–3 weeks from the first attempt. The Bland-Altman plots for the three domains clearly show that both response scores were within the limits of agreement between the first and second attempts (Fig. 4). Moreover, a T-test was performed to compare the differences between the response score of the two attempts. The results indicated no significant differences between the two; UE domain p-value 0.223, EE domain p-value 0.076, and CR domain p-value 0.563 (Table 6). The results confirm the agreement between the two attempts and no proportional bias.

3.5. Gender, BMI and age group differences in response scores

BMI of the participants was positively correlated to the domains UE ($r = 0.159$, p-value 0.000) and EE ($r = 0.158$, p = value 0.000) of the Arabic TFEQ-R18 affirming its construct validity. However, the participant's response in the CR domain was not correlated to BMI ($r = -0.039$, p-value 0.373). Additionally, the response score for all three domains was compared by grouping the population into gender, BMI and age groups. The Kruskal-Wallis H test results indicate that the response did not sufficiently change between males and females (Table 7), as the mean rank value stayed consistent for all response scores. Comparing the response scores of the four BMI weight classes also yielded consistent data points. For the UE domain, the obese participants had a significantly higher mean rank value (333.04) when compared to all other weight groups. However, for the EE domain, the response scores among underweight and normal-weight participants significantly differed from overweight and obese participants. But there was no difference in the domain score between obese and overweight participants. In the CR domain, the highest response score was recorded among overweight participants (283.47), significantly higher than normal, underweight and obese individuals. The difference in the response score of normal-weight participants compared to underweight and obese participants was also significant. Age-related differences were also noticed in UE and CR domain scores. For both UE and CR domains, the mean rank value for the age group of 18–26 years was significantly lower compared to older adults between 36 and 45 years.

4. Discussion

TFEQ-R18 remains one of the widely used tools translated and validated in different languages to investigate the different aspects of eating behavioural traits such as EE, UE and CR in population studies [16,20,22–25]. Based on earlier reports, significant associations have been reported between these eating behavioural traits and population health outcomes [7,26–29]. The main objective of our study was to translate and validate the TFEQ-R18 for Arabic speakers. To our knowledge, this is the first study validating an Arabic TFEQ-R18 in a large population sample. Our study has demonstrated that the Arabic version of the TFEQ-R18 is a valid and reliable tool to evaluate the three intended eating behaviours for educated adults. Our objective to generate a validated Arabic TFEQ-R18 for Arabic speakers was achieved by conducting the study in two phases; Phase 1, which tested the first translation among 298 respondents and Phase 2, which tested the validity of the modified translation among 513 participants. The Arabic TFEQ-R18 was tested in Phase 1 of the study by CFA, which confirmed the model's fitness and highlighted four items in the questionnaire that required modifications in the model. The modified TFEQ-R18 was then retested in a larger population in Phase 2.

To validate the modified questionnaire, CFA and item-to-total correlations were calculated for the data from both stages of Phase 2. Results of CFA confirmed that the modified Arabic TFEQ-18 with the 18 items used in Phase 2 of the study was a valid model. Pearson's correlation coefficient is a popular measure of internal consistency [30]. In the current study for all three domains, correlation coefficients were between 0.328 and 0.870 for the total population of 513. The Arabic questionnaire showed high internal consistency in the UE and EE domains. However, it highlighted a low consistency in the CR domain due to item 15 ('*How frequently do you avoid "stocking up" on tempting foods*?'). Previously, Anna Brytek-Matera et al. reported a low factor loading value of 0.31 for item 15 of the CR domain in the Polish version of TFEQ-R18 [24]. The factor loading for item 15 in Phase 1 was 0.425. In addition, the Spanish version of TFEQ-17, validated by Wrzecionkowska et al., also reported a weak correlation for item 15 [22]. In Phase 1 of our study, item 15 was flagged by the participants as confusing, and the modified translation of this question in Phase 2 only marginally improved its indices in the analysis.

Our results also highlighted the reliability of UE and EE domains of the Arabic TFEQ-R18, which recorded comparable Cronbach's α coefficient reported in both the Finnish and Spanish versions of the TFEQ-R18 [20,26]. Cronbach's α coefficient is a measure of reliability analysis, where values greater than 0.7 are generally considered reliable [31]. However, Schmitt (1996) has proposed that there is no general level such as 0.7 where α becomes acceptable, but tools with even a lower α value can still be helpful in certain circumstances [32]. The Turkish version of the original TFEQ-51 also adopted a similar strategy to adhere to the original scale when Cronbach's α value did not change by removing the items that contributed to α coefficient values lower than 0.7 [25]. The Cronbach's α coefficient of the CR domain was 0.588. Item 15 contributed to the poor result for the CR domain, and the domain reliability improved with its removal. Interestingly, a Spanish version of the TFEQ-R18 reported a Cronbach's α coefficient value of 0.59 for the CR domain even after removing three items from the analysis [22]. In our study, after removing item 15, Cronbach's α coefficient measured for the CR domain was 0.608 and for the whole model was 0.641, comparable to those reported in the literature [20,33]. An overall Cronbach's α coefficient of 0.699 was reported by Mohammadi et al. for the Persian version of TFEQ-R18 [33]. Another statistical approach to confirm the reliability of a questionnaire is test-retest reliability, and in this study, the Arabic TFEQ-R18 questionnaire showed excellent test-retest reliability. The second questionnaire attempt by 127 participants after a gap of 2–3 weeks showed no difference in domain score, indicating that the response to the Arabic questionnaire is consistent over time.

We also analyzed the construct validity by comparing the response score of Phase 2 participants classified by gender, BMI and age

Table 6
Test-retest reliability of the Arabic TFEQ-R18.

TFEQ-R18 Domains	Attempt	N	Score[1]	Coefficient	p-value[2]
UE Domain	I	513	2.35 ± 0.53	0.176	0.223
	II	127	2.26 ± 0.43		
EE Domain	I	513	2.309 ± 0.87	0.180	0.076
	II	127	2.25 ± 0.77		
CR Domain	I	513	2.99 ± 0.57	−0.067	0.563
	II	127	3.09 ± 0.59		

The response scores of both attempts were compared by T-test. [1] Domain scores represented as means ± s.d. [2] Statistical significance set at $p \leq 0.05$. CR-cognitive restraint, UE-uncontrolled eating, and EE-emotional eating.

Table 7
Comparison of Domain Responses of Gender, Age group and BMI in Phase 2.

					P-value[1]	Pairwise Comparison
Gender	Male (N = 79)	Female (N = 434)				
UE	255.68	257.24			0.931	
EE	240.46	260.01			0.278	
CR	229.80	261.95			0.075	
Age group	I (18–26)	II (27–35)	III (36–45)			
	N = 456	N = 29	N = 28			
UE	263.56	238.14	169.75		0.004	I vs III*
EE	257.85	262.86	237.04		0.750	ns
CR	249.04	301.71	340.30		0.002	I vs III*
BMI	I (Underweight)	II (Normal)	III (Overweight)	IV (Obese)		
	N = 69	N = 275	N = 113	N = 56		
UE	243.67	245.45	255.57	333.04	<0.001	I vs IV*, II vs IV* III vs IV*
EE	219.57	235.23	298.89	325.49	<0.001	I vs III*, I vs IV* II vs III*, II vs IV*
CR	186.48	274.23	283.47	205.85	<0.001	I vs II*, I vs III* II vs IV*, III vs IV*

Results of the Kruskal-Wallis H test comparing the difference between the response score in the population divided based on gender, age group and BMI. * The mean rank difference is significant at the 0.05 level. CR-cognitive restraint, UE-uncontrolled eating, EE-emotional eating, BMI -body mass index. BMI categories are defined as ≤18.5 kg/cm^2 underweight, 18.5–24.9 kg/cm^2 normal weight status, 25.0–29.9 kg/cm^2 overweight and ≥30.0 kg/cm^2 obese. Age presented in years.

group. The response scores were comparable between both genders. However, the association between gender and the three domains of TFEQ are inconsistent with other reports on translated TFEQ-R18 scales [25,33]. However, in the current study, there was a positive relationship between BMI categories and the score of UE and EE, and this observation has been corroborated in several studies [20,22,34]. Previous research reported a significant association between the three domains of the TFEQ-R18 questionnaire and BMI [20,24,26]. It is noteworthy that in similar studies, the direction of the correlation is of greater interest than the strength of the association to establish the construct validity of the questionnaire. Previous studies using the TFEQ-R18 have also reported similar findings and have highlighted the importance of considering the direction of the associations. For instance, Mohammadi et al. reported significant correlation coefficients between body weight (r = 0.186), BMI (r = 0.176), and body fat% (r = 0.169) with the Persian version of the TFEQ-R18 [33]. Similarly, Timmerman et al. found a weak but significant relationship between the severity of binge eating behaviour and BMI [35]. Chearskul et al. also reported positive correlations between TFEQ factors and BMI (r = 0.17) and body fat percentage (r = 0.32) [36], while Lesdéma et al. reported a positive correlation between TFEQ and BMI [37]. All these studies are consistent with our results.

The findings of this study showed no significant correlation between BMI and the CR domain. The findings could be attributed to item 15, with a weak reliability score. However, Loffler et al. have suggested that the relationship between BMI and CR is complex [28]. The German version of the TFEQ-R18 reported a high EE and UE pattern with high BMI, while low domain scores were observed for CR in low BMI values [28]. Whereas Johnson et al. reported that individuals with high BMI scored high in CR due to their tendency to restrict food intake to control their weight [8]. Due to the complexity of CR's influence on weight gain, two different types of association have been suggested, flexible and rigid, where the former could help control body weight, but the latter is less effective in weight control [38]. The difference in domain scores of UE was significantly higher in obese individuals in our study with all other weight groups, which is corroborated by similar validation studies that have reported a positive association between UE and high BMI [20,22,28]. This association was expected and is supported by previous studies where EE domain scores were significantly higher in overweight and obese individuals. Our observations corroborate the study by Geliebter et al., which reported that high-emotional eaters, individuals who eat when anxious or depressed, were more often obese [39]. Another factor influencing eating behaviour is age, where a rising UE score is observed with increasing age [20]. Our study found that the domain scores for UE and CR were significantly higher among 36-45 year-olds than among 18-26 year-olds. Similar observations were made by Elfag et al., where the

score of the UE and CR was significantly lower among young adults than in the older age group [40].

5. Conclusions

The study demonstrated that the Arabic TFEQ-R18 is a reliable and valid tool to assess eating behaviour among Arabic speakers, indicating good construct validity. Our study, which included participants of various body weight ranges, revealed a significant association between BMI and the UE and EE domains, supporting previous research linking BMI with eating behaviour. Notably, the findings showed that the Arabic TFEQ-R18 is effective not only among obese individuals but also in the general population. As such, this tool is a crucial resource to gauge the level of three eating behaviours and to develop effective strategies to manage them among the Arab population. However, limitations exist, including the predominance of young adults and those with higher educational backgrounds as participants, reflecting the university setting of the study. Additionally, the self-reported weight and height measures may introduce response bias, requiring caution when generalizing the findings to the broader public.

Author contribution statement

Salma Alhebshi, Lily Stojanovska: Conceived and designed the experiments; performed the experiments; analyzed and interpreted the data; contributed analysis tools or data; wrote the paper.

Serene Hilary: Analyzed and interpreted the data; contributed analysis tools or data; wrote the paper.

Samir K. H. Safi: Analyzed and interpreted the data.

Habiba I. Ali: Performed the experiments.

Leila Cheikh Ismaill; Ayesha Al-Dhaheri: Performed the experiments; wrote the paper.

Data availability statement

Data will be made available on request.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

The authors would like to thank undergraduate students Fatima Rashid Al-Shamsi, Aaesha Salem Alhebsi, Moza Ali Al-Salami, Maitha Mohammed Al-Neyadi, Bushra Nassif, Sara Al-Qubaisi of the Department of Nutrition and Health, CMHS, UAE University for their involvement with the study as well as faculty and staff of UAEU for participating in the study.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.heliyon.2023.e17623.

References

[1] WHO. *Controlling The Global Obesity Epidemic*, 2021 [cited 2022 11 March]; Available from: https://www.who.int/activities/controlling-the-global-obesity-epidemic.
[2] L. Alhyas, et al., Prevalences of overweight, obesity, hyperglycaemia, hypertension and dyslipidaemia in the Gulf: systematic review, JRSM short reports 2 (7) (2011) 1–16.
[3] N. Sulaiman, et al., Prevalence of overweight and obesity in United Arab Emirates expatriates: the UAE national diabetes and lifestyle study, Diabetol. Metab. Syndrome 9 (1) (2017) 88.
[4] S.W. Ng, et al., Nutrition transition in the United Arab Emirates, Eur. J. Clin. Nutr. 65 (12) (2011) 1328–1337.
[5] Who, The Global Database on the Implementation of Nutrition Action (GINA) Policy - National Action Plan in Nutrition, World Health Organisation, United Arab Emirates, 2017.
[6] CDC, Strategies to Prevent & Manage Obesity, 2020 [cited 2022 11 March]; Available from: https://www.cdc.gov/obesity/strategies/index.html.
[7] U. Vainik, I. García-García, A. Dagher, Uncontrolled eating: a unifying heritable trait linked with obesity, overeating, personality and the brain, Eur. J. Neurosci. 50 (3) (2019) 2430–2445.
[8] F. Johnson, M. Pratt, J. Wardle, Dietary restraint and self-regulation in eating behavior, Int. J. Obes. 36 (5) (2012) 665–674.
[9] B. De Lauzon, et al., The Fleurbaix Laventie Ville Sante (FLVS) study group: the three-factor eating questionnaire-R18 is able to distinguish among different eating patterns in a general population, J. Nutr. 2380 (134) (2004), 2372-2004.
[10] E.J. Bryant, N.A. King, J.E. Blundell, Disinhibition: its effects on appetite and weight regulation, Obes. Rev. 9 (5) (2008) 409–419.
[11] A.J. Romain, J. Marleau, A. Baillot, Impact of obesity and mood disorders on physical comorbidities, psychological well-being, health behaviours and use of health services, J. Affect. Disord. 225 (2018) 381–388.
[12] A.J. Stunkard, S. Messick, The three-factor eating questionnaire to measure dietary restraint, disinhibition and hunger, J. Psychosom. Res. 29 (1) (1985) 71–83.

[13] J. Karlsson, et al., Psychometric properties and factor structure of the Three-Factor Eating Questionnaire (TFEQ) in obese men and women. Results from the Swedish Obese Subjects (SOS) study, Int. J. Obes. 24 (12) (2000) 1715–1725.
[14] J.-l. Yabsley, et al., Validation of a child version of the Three-Factor Eating Questionnaire in a Canadian sample: a psychometric tool for the evaluation of eating behaviour, Publ. Health Nutr. 22 (3) (2019) 431–443.
[15] E.J. Bryant, et al., Development and validation of the child three-factor eating questionnaire (CTFEQr17), Publ. Health Nutr. 21 (14) (2018) 2558–2567.
[16] I. Jáuregui-Lobera, et al., Psychometric properties of Spanish version of the Three-Factor Eating Questionnaire-R18 (Tfeq-Sp) and its relationship with some eating-and body image-related variables, Nutrients 6 (12) (2014) 5619–5635.
[17] E. Kavazidou, et al., Structure Validity of the Three-Factor Eating Questionnaire-R18 in Greek Population, 2012.
[18] D. Wild, et al., Principles of good practice for the translation and cultural adaptation process for patient-reported outcomes (PRO) measures: report of the ISPOR task force for translation and cultural adaptation, Value Health 8 (2) (2005) 94–104.
[19] R.F. Woolson, J.A. Bean, P.B. Rojas, Sample Size for Case-Control Studies Using Cochran's Statistic, Biometrics, 1986, pp. 927–932.
[20] S. Anglé, et al., Three factor eating questionnaire-R18 as a measure of cognitive restraint, uncontrolled eating and emotional eating in a sample of young Finnish females, Int. J. Behav. Nutr. Phys. Activ. 6 (2009), 41-41.
[21] J.M. Hodge, et al., Validation of self-reported height and weight in a large, nationwide cohort of U.S. adults, PLoS One 15 (4) (2020), e0231229.
[22] D. Wrzecionkowska, S. Rivera Aragón, Three-factor eating questionnaire-R18 (TFEQ-R18) Spanish version: factor structure analysis among normal weight and overweight adults, Acta de investigación psicológica 11 (1) (2021) 84–94.
[23] L. Palmeira, et al., The revised Portuguese version of the Three-Factor Eating Questionnaire: a confirmatory factor analysis, BMC Health Serv. Res. 16 (2016) S3.
[24] A. Brytek-Matera, R. Rogoza, K. Czepczor-Bernat, The Three-Factor Eating Questionnaire-R18 Polish version: factor structure analysis among normal weight and obese adult women, Arch. Psychiatr. Psychother. 19 (3) (2017) 81–90.
[25] Ö. Küçükerdönmez, et al., Turkish version of the 'Three-Factor Eating Questionnaire-51' for obese individuals: a validity and reliability study, Publ. Health Nutr. 24 (11) (2021) 3269–3275.
[26] I. Jáuregui-Lobera, et al., Psychometric properties of Spanish version of the Three-Factor Eating Questionnaire-R18 (Tfeq-Sp) and its relationship with some eating- and body image-related variables, Nutrients 6 (12) (2014) 5619–5635.
[27] H. Konttinen, Emotional eating and obesity in adults: the role of depression, sleep and genes, Proc. Nutr. Soc. 79 (3) (2020) 283–289.
[28] A. Löffler, et al., Eating behaviour in the general population: an analysis of the factor structure of the German version of the three-factor-eating-questionnaire (TFEQ) and its association with the body mass index, PLoS One 10 (7) (2015), e0133977.
[29] A. Meule, The psychology of food cravings: the role of food deprivation, Current nutrition reports 9 (3) (2020) 251–257.
[30] W. Tang, Y. Cui, O. Babenko, Internal consistency: do we really know what it is and how to assess it, Journal of Psychology and Behavioral Science 2 (2) (2014) 205–220.
[31] M. Tavakol, R. Dennick, Making sense of Cronbach's alpha, Int. J. Med. Educ. 2 (2011) 53.
[32] N. Schmitt, Uses and abuses of coefficient alpha, Psychol. Assess. 8 (4) (1996) 350.
[33] M.R. Mohammadi, et al., Reliability and validity of the Persian version of Food Craving Questionnaire-Trait-Reduced (FCQ-Tr) in overweight and obese women, Journal of Nutrition, Fasting and Health 6 (3) (2018) 150–157.
[34] J.C. Cappelleri, et al., Psychometric analysis of the Three-Factor Eating Questionnaire-R21: results from a large diverse sample of obese and non-obese participants, Int. J. Obes. 33 (6) (2009) 611–620.
[35] G.M. Timmerman, J.S. Stevenson, The relationship between binge eating severity and body fat in nonpurge binge eating women, Res. Nurs. Health 19 (5) (1996) 389–398.
[36] S. Chearskul, et al., Thai version of three-factor eating questionnaire, Appetite 54 (2) (2010) 410–413.
[37] A. Lesdéma, et al., Characterization of the three-factor eating questionnaire scores of a young French cohort, Appetite 59 (2) (2012) 385–390.
[38] C.A. Timko, J. Perone, Rigid and flexible control of eating behavior in a college population, Eat. Behav. 6 (2) (2005) 119–125.
[39] A. Geliebter, A. Aversa, Emotional eating in overweight, normal weight, and underweight individuals, Eat. Behav. 3 (4) (2003) 341–347.
[40] K. Elfhag, Y. Linné, Gender differences in associations of eating pathology between mothers and their adolescent offspring, Obesity research 13 (6) (2005) 1070–1076.