New developments in the assessment of weight-related psychological inflexibility (AAQW-Revised)

Lara Palmeira, Marina Cunha, José Pinto Gouveia,
Sérgio Carvalho, & Jason Lillis

2016
Journal of Contextual Behavioral Science
5(3): 193-200
doi: 10.1016/j.jcbs.2016.06.001

New developments in the assessment of weight-related psychological inflexibility (AAQW-Revised)

Lara Palmeira, Marina Cunha, José Pinto Gouveia, Sérgio Carvalho, & Jason Lillis

Abstract

Experiential avoidance, defined as attempts to control or change unwanted internal experiences when doing so causes harm, has been consistently associated with physical and mental health problems and has been traditionally measured using the Acceptance and Action Questionnaire. Several studies have highlighted the importance of developing content-specific measures to better capture relevant processes for specific populations. One such measure is the Acceptance and Action Questionnaire for Weight-Related Difficulties (AAQ-W), which measures experiential avoidance of unwanted weight related thoughts, feelings and actions. The AAQW factor structure still requires further examination.

The present study aims to contribute to the further development of the AAQW by conducting a confirmatory factor analysis (CFA) based on the existent factor structures and testing the measurement invariance across groups, through a multi-group analysis. Three distinct samples were used: the CFA used 215 women from the general population with BMI < 25 (sample 1); 210 overweight or obese treatment seeking (sample 2); AAQW's temporal stability and sensitivity to change was assessed using a sample of 58 overweight and obese women enrolled in Kg-Free (sample 3). Results supported a revised and shorter version of the AAQW (10 items) that we call AAQW-R (revised) with a three-factor structure (food as control, weight as barrier to living, weight-stigma) that showed a good fit to the data. Also, the measurement invariance across groups was confirmed. Finally, AAQW-R proved to be a reliable, stable measure and sensitive to clinical changes. Overall, this study offers new advances in the assessment of weight-related experiential avoidance, proposing a revised version of the AAQW. Moreover, it provides evidence for the usefulness of the AAQW-R both in general and clinical populations.

Key-words: weight-related experiential avoidance; confirmatory factor analysis; overweight and obesity; multi-group analysis;

1. Introduction

Obesity has been considered one of the most serious worldwide health problems, with increasing prevalence despite the availability of weight loss treatments (Fassino et al., 2002; WHO, 2011). There has been an increasing interest in literature regarding the psychological factors associated with poorer weight loss outcomes (e.g., Avenell et al., 2004; Byrne, Cooper, & Fairburn, 2003; Elfhag & Rössner, 2004). Individuals who are more likely to regain weight after a weight loss program have shown a tendency to report lower self-esteem, higher emotional eating, impulsivity and rigid control of eating and avoidance-based motivations for losing weight (e.g. to avoid being criticized by others or by the self; Avenell et al., 2004; Byrne et al., 2003; Fassino et al., 2002; Ogden, 2000). These characteristics can be conceptualized as a pattern of experiential avoidance regarding weight-related internal negative experiences (Byrne et al., 2003; Lillis, Hayes, Bunting, & Masuda, 2009; Kayman, Bruvold, & Stern, 1990).

Acceptance and Commitment Therapy (ACT; Hayes, Wilson, & Strosahl, 1999) defines experiential avoidance as the unwillingness to be in contact with unwanted difficult internal experiences (e.g., thoughts, emotions, physical sensations, urges) and attempts to control, suppress or avoid them. It has been proposed that eating behavior difficulties can be conceptualized in part as ineffective attempts to regulate internal experiences perceived as negative and unwanted (Baer, Fischer, & Huss, 2006; Merwin et al., 2011).

The literature has consistently shown that experiential avoidance is associated with overall psychopathological problems, diminished quality of life (e.g., Bond et al., 2011; Pinto-Gouveia, Gregório, Dinis, & Xavier, 2012) and functioning related to chronic medical conditions (Gifford et al., 2004; Gregg, Callaghan, Hayes, & Glenn-Lawson, 2007; Tapper, et al., 2009). The primary measure used to assess experiential avoidance is the Acceptance and Action Questionnaire (AAQ-II; Bond et al., 2011). However given that the AAQ-II was developed with a general mental health focus, recent studies have found that using a content specific measure of experiential avoidance can be more powerful, particularly in chronic health domains such as diabetes (Gregg et al., 2007), irritable bowel syndrome (Ferreira, Eugenicos Morris, Gillanders, 2013), and epilepsy (Lundgren, Dahl, & Hayes 2008).

The AAQW (*Acceptance and Action Questionnaire for Weight-Related Difficulties;* Lillis & Hayes, 2008) represents the first attempt at measuring experiential avoidance in relation to difficulties with eating, weight, and physical activity. Although the original version of the AAQW showed acceptable psychometric properties and temporal reliability, it was validated on a small sample size (n = 84) that did not allow for a full exploration of the factor structure using multiple groups. Indeed the authors stated that the factorial structure of AAQW still needed further analysis. The original study suggested a unifactorial structure (Lillis & Hayes, 2008), however a more recent study attempted to validate the AAQW for use with bariatric surgery patients and

suggested that a five factor structure might be appropriate in that context (Weineland, Lillis, & Dahl, 2012). However, three of the identified factors presented low internal consistencies (ranging between .44 to .67) and only 20 items of the original 22 were retained in the factor structure found.

More recently the psychometric properties of the Portuguese version of the AAQ-W was tested in a sample of 249 women with overweight and obesity seeking nutritional treatment. Results from the exploratory factor analysis did not entirely support the factor structure presented by Weineland et al. (2012). Instead, a three factor structure emerged (factor 1 - *food as control*; factor 2 – *emotional avoidance* and factor 3 - *weight-stigma)* explaining 50.94% of the AAQW total variance. From the original 22 items, only 15 items were retained in the final Portuguese version. The measure revealed good internal consistency (α=.81) and convergent and divergent validity (Cardoso, 2014).

Studies using the original 22-item version of the AAQW have found that weight-related experiential avoidance is associated with general psychopathology, body dissatisfaction, disordered eating attitudes and behaviors, binge eating symptoms and diminished quality of life (Cardoso, 2014; Lillis & Hayes, 2008; Lillis et al., 2009; Lillis, Hayes, & Levin, 2011; Weineland et al., 2012). Furthermore, Lillis et al. (2009) found that weight-related experiential avoidance mediated the impact of a 1-day ACT workshop on weight, weight self-stigma, psychopathological symptoms and health-related quality of life.

The primary aim of the current study is to perform confirmatory factor analysis of two possible factor structures (five-factor and three-factor) in a large sample of adult women. In addition, it explores the psychometric properties and construct validity of the AAQW. A multi-group factor analysis was also performed in order to test the measurement invariance of the AAQW in two different groups (women from general population and women with overweight and obesity seeking weight management treatment). Finally, a third sample composed of overweight and obese women enrolled in a 12 session compassionate mindfulness & acceptance group training (Kg-Free) was used to assess AAQW temporal stability and sensitivity to clinical change.

2. Methods

2.1. Participants

Sample 1 - Participants were 215 Portuguese women from the general population with BMI < 25. Mean age was 29.55 (*SD* = 9.52), with a mean years of education of 14.09 (*SD* = 2.57). Mean BMI was 21.49 (*SD* = 1.73). Concerning marital status, 67.9% of the sample was single and 21.9% married. The majority (46.7%) had a medium to high socio-economic status.

Sample 2 - This sample is comprised of 210 Portuguese women with overweight or obesity seeking nutritional treatment from both private and public health institutions in the district

of Coimbra, Portugal. Mean BMI was 31.14 (SD = 5.31), with a mean age of 40.14 (SD = 12.19) and a mean of 10.90 (SD = 3.81) years of education. The majority presented low to medium socio-economic status (65.7%) and, regarding marital status, 42% were single and 24.6% were married.

Sample 3 - Sample 3 is comprised of an additional 58 overweight or obese treatment seeking Portuguese women without binge eating disorder who were randomly assigned to one of two treatment conditions as part of a larger intervention trial: Kg-Free intervention (n = 28) or treatment as usual (TAU; n = 30). The Kg-Free intervention was developed to target weight stigma, shame and self-criticism and promote emotion regulation skills based on acceptance, mindfulness and self-compassion. At baseline and at the end of the intervention ($3^{1/2}$ months) participants were assessed. At baseline, the sample reported a mean BMI of 33.92 (*SD* = 5.22), a mean age of 42.67 (*SD* = 8.81) and a mean of 15.69 (*SD* = 3.80) years of education. Concerning marital status, 60.3% of the sample was married, 17.2% was single, and 13.8% was divorced. The majority (65.5%) came from a low to medium socio-economic status.

A summary of each sample as well as a description of the statistical procedures used with each sample are described in Table 1.

Table 1

Samples description

	Participants	Mean BMI (*SD*)	Statistical procedure
Sample 1 (n = 215)	Women from general population (not overweight)	21.49 (1.73)	Multi- Group analysis and independent t-tests
Sample 2 (n = 210)	Overweight or obese women treatment seeking	31.14 (5.31)	Multi- Group analysis and independent t-tests
Combination of Sample 1 and Sample 2 (n = 425)	Women from general population (not overweight) + Overweight or obese women treatment seeking	26.39 (6.31)	Confirmatory Factor Analysis, internal reliability, validity analyses
Sample 3 N = 58 (Kg-Free intervention n= 28; TAU n= 30)	Overweight or obese women enrolled in Kg-Free	33.92 (5.22)	Temporal stability (only TAU group) and sensitivity to change

Note: Sample 2 and Sample 3 are independent samples.

2.2. Measures

Demographic Data was obtained from participants self-report, including current height and weight. BMI (Wt/Ht^2) was calculated.

Acceptance and Action Questionnaire for Weight-Related Difficulties (AAQW; Lillis & Hayes, 2008) is a weight focused version of the original AAQ comprising 22 items that

specifically assesses experiential avoidance in relation to weight-specific thoughts, feelings, and bodily sensations. Items are rated on a 7-point scale (1 = "never true" or "not at all believable" and 7 = "always true" or "completely believable"). Higher scores reflect more weight-related experiential avoidance. The original version of the AAQ-W showed good internal consistency (α = .86) and test-retest reliability (Lillis & Hayes, 2008).

Acceptance and Action Questionnaire (AAQ-II; Bond, et al. 2011; Portuguese version by Pinto-Gouveia et al., 2012) is a widely used, 7-item questionnaire that assesses psychological inflexibility on a seven point scale. Higher scores indicate higher levels of psychological inflexibility (Bond et al. 2011). Both the original and Portuguese versions presented good psychometric properties (Bond et al. 2011; Pinto-Gouveia et al., 2012). In this study, the AAQ-II had a very good internal consistency of .92.

Other as Shamer Scale - Brief (OAS-2; Matos, Pinto-Gouveia, Gilbert, Duarte, & Figueiredo, 2015) is an 8-item questionnaire used to measure external shame, rated on a 5-point scale. Higher scores reflect higher levels of external shame (Matos et al., 2015). The scale showed an adequate internal consistency (α = .85), temporal stability and convergent and divergent validity (Matos et al., 2015). In the current study the internal consistency of OAS was very good .93.

Eating Disorder Examination Questionnaire (EDE-Q; Fairburn & Beglin, 1994; Portuguese version by Machado et al., 2014). The EDE-Q is a 36-item self-report measure that assesses disordered eating attitudes and behaviors. Although EDE-Q has four subscales, in the current study only the global score was calculated. EDE-Q has consistently shown to be a reliable measure of eating psychopathology (Fairburn, 2008). In this study the internal consistency of the EDE-Q was very good .92.

Binge Eating Scale (BES; Gormally, Black, Daston, & Rardin, 1982; Duarte, Pinto-Gouveia, & Ferreira, 2013) is a 16-item self-report measure assessing binge eating symptoms. For each item participants are asked to choose which sentence best describes their experience. Higher scores reflect higher severity of binge eating symptoms (from 0 to 46) and scores above 17 indicate the presence of binge eating symptoms (Duarte, Pinto-Gouveia, & Ferreira, 2015). Both the original and the Portuguese versions have shown good internal consistency, similar to the one found in this study (α = .89).

Subjective Happiness Scale (SHS; Lyubomirsky & Lepper, 1999; Portuguese version by Pais- Ribeiro, 2012) contains 4 items rated on a 7-point scale. Two items ask participants to rate themselves through absolute and peer-related ratings and the other two ask participants to indicate the extent to which a given description describes them. The instrument has consistently shown good psychometric properties, with Cronbach alpha ranging from .79 to .94 (Lyubomirsky & Lepper, 1999). The Portuguese version also showed adequate internal consistency (α = .76; Pais-Ribeiro, 2012). In this study Cronbach alpha was .79.

2.3. Procedures

The current study was approved by the institutions where the samples were collected. The general population sample (sample 1) is a convenience sample from Coimbra, Portugal. This sample was collected in Coimbra´s Citizen´s Bureau from January to February 2014. Each participant was invited to participate by a member of the research team, who assured the voluntary and confidential nature of the data. Participants were given informed consent and the research goals were clarified.

Separately from sample 1, participants from the clinical sample (sample 2) were invited by a member of the research team to take part in the study on the day of their ongoing nutritional appointment at the hospital or private clinic. This sample was collected from October 2013 and June 2014.

Finally, sample 3 was recruited in Coimbra´s University Hospital (CHUC) at the endocrinology and internal medicine services. Participants were referred to the research team by the endocrinologist or resident, and then invited by a member of the research team to participate in the intervention study (Kg-Free). The sample, as well as all assessment moments occurred between May 2014 and September 2015.

Participants in all three samples were informed about the voluntary and confidential nature of their collaboration as well as the study's goals and gave their informed consent by a member of the research team. After they gave their consent, participants took approximately 20 min to complete the self-report measures.

2.4. Data analysis

Preliminary data analyses (Skewness and Kurtosis; Multicollinearity; Mahalanobis distance statistic for outlier analysis) were executed to examine the adequacy of the data. The AAQW psychometric properties were performed using IBM SPSS Statistics and the confirmatory factor analysis with AMOS Software.

Confirmatory Factorial Analysis (CFA) was conducted on the combined sample 1 and 2 to test and compare the two existent factor structures of the AAQW (the original version with 5 factors and the Portuguese version comprising 3 factors). The Maximum Likelihood (ML) estimation method was used as it is one of the most frequently used and suggested to be robust and appropriate for our goals (Brown, 2006; Iacobucci, 2010; Kline, 2005; Schermelleh-Engel, Moosbrugger, & Müller, 2003). Several goodness-of-fit indices and recommended cut-off points were used to evaluate the model fit (Brown, 2006; Kline, 2005): *Chi-Square* ($\chi 2$), *Normed Chi-Square* ($\chi 2$/d.f.), *Comparative Fit Index* (CFI $\geq$ .90, acceptable, and $\geq$ .95, desirable; Hu & Bentler, 1998), *Tucker-Lewis Index* (TLI $\geq$ .90, acceptable, and $\geq$ .95, desirable; Hu & Bentler, 1998), *Goodness of Fit Index* (GFI $\geq$ .90, good, and $\geq$ .95, desirable; Jöreskog & Sörbom, 1996), *Root Mean Square Error of Approximation* (*RMSEA* $\leq$ .05, good fit; $\leq$ .08, acceptable fit; $\geq$ .10,

poor fit; Brown, 2006; Kline, 2005) using a 90% confidence interval. In order to compare the two factor structures (original *versus* Portuguese structure) the chi-square difference test was used, with statistically significant difference (X^2 0.95) indicating better models. Moreover, Akaike Information Criterion (AIC) and Expected Cross-Validation Index (ECVI) were analyzed to compare alternative models (Schermelleh-Engel et al., 2003). Lower values on AIC and ECVI are considered indicators of superior models (Arbuckle, 2008).

Item standardized factor loadings (λ) and individual reliability (R^2) were examined as indicators of local adjustment. It has been stated that when $\lambda \geq .50$ the model has factorial validity and when $R^2 \geq .25$ items show individual reliability (Hair, Anderson, Tatham, & Black, 1998).

To examine AAQ-W reliability, *Cronbach's alphas* (cut-off of .70 is considered suitable; Field, 2013) and the item-total correlations were used to assess scale's internal consistency (Tabachnick & Fidell, 2007). Also, *Composite Reliability* and variance extracted measure (VEM should be ≥.50) were estimated. Composite reliability measures internal reliability of each construct and indicates the degree to which the individual indicators are all consistent with their common latent construct (values ≥.70 indicate acceptable reliability; Hair et al., 1998).

Convergent and divergent validities were assessed through Pearson correlation coefficients (Cohen, Cohen, West, & Aiken, 2003). We expected high correlations with general experiential avoidance (AAQ-II) and measures of eating pathology (BES, EDE-Q), moderate correlations with BMI, and lower (but still significant) correlations with subjective happiness (SHS) and external shame (OAS).

A *multi-group confirmatory factor analysis* was performed in order to assess structural invariance of the AAQW across different samples. To do so, two separate samples were used: sample 1 (composed by woman from general population with BMI < 25) and sample 2 (overweight or obese women seeking nutritional treatment). The invariance of the structural model for both groups was tested through the chi-square difference test (Byrne, 2010).

To examine differences in AAQW across two distinct groups (sample 1 and sample 2) we conducted *Independent sample t tests* and calculated *Cohen's d effect sizes* (e.g., Field, 2013). Cohen's guidelines were used to interpret effect size magnitude (1988 cited in Tabachnick & Fidell, 2007).

Temporal stability, as known as *test- retest reliability*, was performed in TAU group from Kg-free (Sample 3) by comparing results from the baseline assessment and the assessment after a three-month period, using Pearson product–moment correlations and *t*-tests for paired samples. Sample 3 was used to measure AAQW's s*ensitivity to change* through an analysis of covariance on 3-month follow-up scores with the baseline score as a covariate.

3. Results

3.1. Preliminary Data Analyses

Violations of normality were not found, as data Skewness and Kurtosis values were in acceptable ranges (SK < |3| and Ku < |8-10|). Multicollinearity was not problematic as all variables presented VIF values < 5 (Kline, 2005). Additionally, the Mahalanobis distance statistic (D^2) was calculated to examine the existence of multivariate outliers. Despite the fact that three cases did present values that indicate the presence of outliers, we decided to maintain them. This decision was based on the suggestion that data are more likely to be representative of the population when outliers are included (Kline, 2005; Tabachnick & Fidell, 2007).

3.2. Confirmatory Factor Analyses (combined samples 1 & 2)

First, the model with the Weineland et al. (2012) five-factor structured was tested (model 1). This model presented a poor fit to the data (see Table 2). We then tested the three factor structure from the Portuguese exploratory factor analysis (model 2). Results showed an adequate fit for the model. Additionally, when the two models were compared, model 2 was statistically superior to model 1 (chi-square difference test: $\chi^2_{dif} = 378.118 > \chi^2_{0.95;\,(109)} = 134.370$) and presented lower values of comparisons indexes (AIC and EVCI; cf. Table 2) indicating a better fit to the data.

Nevertheless, some items from model 2 did not reach the recommended cut-off points for item factor loadings and Squared Multiple Correlations, namely items: 14 ($\lambda = .16$ and $R^2 = .03$), 18 ($\lambda = .30$ and $R^2 = .09$) and 5 ($\lambda = .42$ and $R^2 = .18$) and were excluded. Additionally, items 3 and 4 also presented local adjustment values just below the cut-off points ($\lambda = .24$ and $R^2 = .49$ for both items), had item-total correlation below .30 and did not contribute to the scale and subscale's internal consistency. In addition, both items failed to evidence theoretical consistency with their underlying factor. Thus, we decided to exclude them, based on both statistical and theoretical justification. The model was then respecified without those five items and the model showed a very good fit to the data (see Table 2, model 3), with the exception of the Chi-square value that remained statistically significant. Nevertheless, Chi-square is very sensitive to sample size and tends to be significant with large samples (Schermelleh-Engel et al., 2003).

Table 2

Goodness-of-fit statistics for comparative models of the AAQW (N =425).

Models	Chi-square	Df	X^2/df	CFI	TLI	GFI	RMSEA (90% C.I.; *p*)	AIC	ECVI
Model 1 (5 factors, 20 items)	545.552***	160	3.410	.848	.820	.885	.075*** (.069 to .082)	645.552	1.523
Model 2 (3 factors, 15 items)	167.434***	51	3.283	.940	.923	.932	.073** (.061 to .086)	245.434	.579
Model 3 (3 factors, 10 items)	90.242***	32	2.820	.966	.952	.957	.066 [.050 to .082; *p* =.052]	136.242	.321
Model 4 (2^{nd} order factor, 10 items)	90.242***	32	2.820	.966	.952	.957	.066 [.050 to .082; *p* =.052]	136.242	.321

Note. df = degrees of freedom; CFI = Comparative Fit Index; TLI = Tucker-Lewis Index; GFI = Goodness of Fit Index; RMSEA = Root Mean Error of Approximation; C.I. = Confidence Interval; AIC = Akaike Information Criterion; ECVI = Expected Cross-Validation Index.

**$p < .01$; ***$p < .001$;

Finally, we also tested a second-order CFA (see Fig. 1). This decision was based on the fact that the three factors were highly correlated with each other and with the scale's global score (e.g., Chen, Sousa, & West, 2005). In addition, the original publication of the AAQW (Lillis & Hayes, 2008) argued that the data supported a unified factor solution representing weight-related experiential avoidance. This change did not modify the model fit nor items factor loading and squared multiple correlations.

The final factor structure includes an underlying second-order factor. This factor is composed of the three sub factors that assess different aspects of weight-related experiential avoidance. The identified sub factors are: Factor 1 – *food as control,* which reflects the tendency to use food as a coping mechanism to deal with negative emotions. Factor 2 – *weight as barrier to living* includes items that assess the tendency to move away from a valued life due to one's

weight or body shape; Factor 3 – *weight-stigma* contains items that assess experiences of internalized stigma related to one's weight.

The final, 10-item version of the measure is referred to as the AAQW-R (revised) for the remainder of the manuscript.

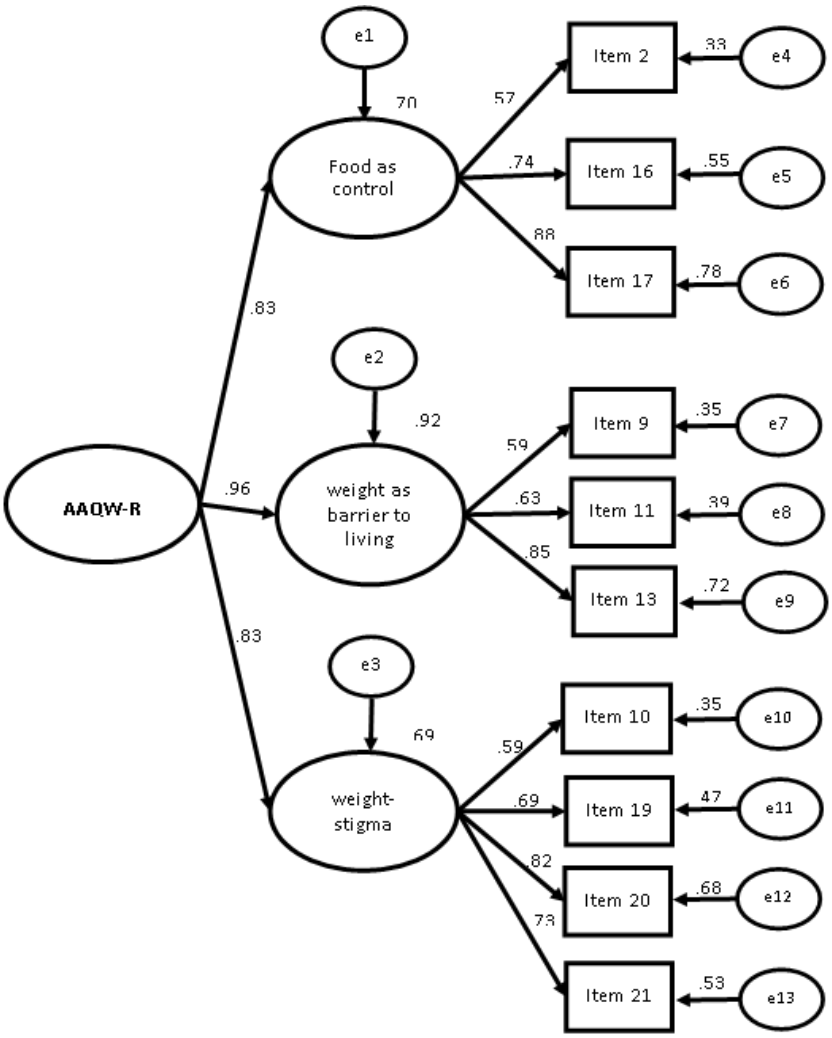

Fig. 1. Confirmatory Factor Analysis of the three-factor of the AAQW (N = 425). Standardized coefficients are shown; all paths are statistically significant (p < .001).

3.3. Descriptive Statistics and Reliability Analysis

Table 3 presents the means, standard deviations, corrected item total correlation, Cronbach's alpha if item deleted and Cronbach's alpha for the total score and all subscales.

Table 3

Means (M), standard deviations (SD), corrected item-total correlations, Cronbach's alpha and Cronbach's alpha if item deleted for Acceptance and Action Questionnaire for Weight-Related Difficulties-Revised (AAQW-R) and its dimensions (N = 425).

Items	***M***	***SD***	**Corrected item-total *r***	**Cronbach's alpha**
AAQW-R_ food as control				**.77**
2. When I have negative feelings, I use food to make myself feel better	2.86	1.73	.51	.73
16. My eating urges control me	2.84	1.83	.62	.71
17. I need to get rid of my eating urges to eat better	3.49	2.15	.70	.65
AAQW-R_weight as barrier to living				**.73**
9. I need to feel better about how I look in order to live the life I want to	4.41	2.01	.50	.65
11. If I'm overweight, I can't live the life I want to	3.63	2.11	.56	.65
13. If I gain weight, that means I have failed	3.60	2.15	.60	.57
AAQW-R_ weight-stigma				**.79**
10. Other people make it hard for me to accept myself	2.28	1.67	.52	.78
19. If I eat something bad, the whole day is a waste	2.49	1.79	.54	.77
20. I should be ashamed of my body	2.19	1.73	.72	.68
21. I need to avoid social situations where people might judge me	2.25	1.75	.63	.73
AAQW-R Total				**.88**

The AAQW-R total score and the three subscales presented good internal reliability, with a Cronbach's alphas ranging from of .73 to .88. All items presented item-total correlations above .30 and contributed for the measure's internal consistency.

Moreover, the AAQW-R showed good composite reliability (.95 to AAQW-R's total score, .73 to AAQW-$R_{food\ as\ control,}$ .82 for AAQW-$R_{weight\ as\ barrier\ to\ living}$ and .90 to AAQW-$R_{weight\text{-}stigma}$ dimensions). Finally, all dimensions presented adequate variance extracted measure (VEM), specifically: .67 AAQW-$R_{food\ as\ control}$, .61 to AAQW-$R_{weight\ as\ barrier\ to\ living}$ and .63 to AAQW-$R_{weight\text{-}stigma}$. According to these results, it seems that the latent constructs are reflected in the items that comprise them.

3.4. *Construct validity*

Table 4 presents the means, standard deviations and Pearson's correlation coefficients for all variables in study. AAQW-R_{total} showed positive and moderate to high correlations with the BES and the EDE-Q, moderate positive associations with BMI, AAQ-II, and the OAS, and negative and low to moderate correlations with subjective happiness.

3.5. Multi-group analysis

A multi-group analysis was also conducted to test the measurement invariance of the AAQW-R across two samples, one from the general population (sample 1) and a clinical sample (sample 2) comprised of women with overweight and obesity seeking nutritional treatment (Meredith, 1993). Measurement invariance is suggested when measurement properties are structurally equivalent in different groups (Meredith, 1993; Schmitt & Kuljanin, 2008). The multiple-group CFA invariance was verified by comparing the unconstrained model (i.e., with free structural parameter coefficients) and the constrained model (i.e., where the parameters are constrained equally across groups; Byrne, 2010). The model presented a very good fit to the data for both groups: GFI= .95; CFI = .97; TLI = .96; RMSEA = .040, p[rmsea≤.05] = .903, I.C. 90%].026; .053[. Additionally, results confirm the invariance of measurement across groups for measurement weights (i.e., equal factor loadings) ($\chi^2_{dif\ (7)} = 9.603$, $p = .212 < \chi^2_{0.95;(7)} = 14.067$).

Table 4

Means (M), Standard Deviations (SD) and Intercorrelation scores on self-report measures (N =425).

Measures	*M*	*SD*	1	2	3	4	5	6	7	8	9
1. BMI	26.39	6.31	-								
2. AAQW-R _Total	30.04	13.16	.51***	-							
3. AAQW-R_food as control	9.20	4.74	.40***	.85***	-						
4. AAQW-R_weight as barrier to living	11.64	5.06	.42***	.87***	.62***	-					
5. AAQW-R_weight-stigma	9.20	5.45	.49***	.87***	.59***	.62***	-				
6. AAQ-II	19.99	9.33	.06	.45***	.39***	.33***	.45***	-			
7. OAS	5.36	5.54	.18***	.47***	.37***	.31***	.53***	.59***	-		
8. BES	25.56	7.27	.44***	.65***	.63***	.46***	.59***	.33***	.40***	-	
9. EDE-Q Total	1.37	1.14	.54***	.70***	.54***	.60***	.65***	.40***	.34***	.67***	-
10. SHS	4.94	1.05	.07	-.37***	-.28***	-.28***	-.39***	-.53***	-.44***	-.20**	-.27***

Note. BMI = Body Mass Index; AAQW-R = Acceptance and Action Questionnaire for Weight-Related Difficulties- Revised; AAQ-II = Acceptance and Action Questionnaire; OAS = Others as Shamer Scale; BES = Binge Eating Scale; EDE-Q = Eating Disorder Examination Questionnaire; SHS = Subjective happiness scale

*** p <.001; ** p <.01.

3.6. Group differences

Independent *t* tests were performed to explore differences in AAQW-R total score and its three factors regarding participant's with distinct BMI comparing participants from sample 1 (women from general population, BMI < 25) and sample 2 (women seeking nutritional treatment, BMI > 25). Table 5 shows means, standard deviations, *t*-test differences and Cohen's d for AAQW-R total score and all its dimensions. Results showed that the overweight and obese group (BMI > 25) presented significantly higher levels of weight-related experiential avoidance patterns than women from the general population. The differences reflect a large effect (Table 5).

Table 5

Means (M), standard deviations (SD), t-test differences and Cohen's d for effect size by group for AAQW-R dimensions (N = 425)

	Overweight and Obese (*n* = 210)		**Not Overweight (*n* = 215)**				
	M	***SD***	***M***	***SD***	***t(df)***	***p***	**Cohen's d**
AAQW-R-Total	35.75	13.65	24.47	9.75	9.790	<.001	.95
AAQW-R_food as control	10.90	4.94	7.52	3.87	7.847	<.001	.76
AAQW-R_weight as barrier to living	13.56	5.05	9.76	4.32	8.320	<.001	.81
AAQW-R_weight-stigma	11.29	6.06	7.19	3.82	8.333	<.001	.81

3.7. Temporal stability

Test-retest reliability, also referred to as temporal stability, is a measure of how reliable a scale is across two different time points. AAQW-R temporal stability was examined using the 30 participants from sample 3 who were allocated to TAU condition. Participants completed two assessments within a three month period. Results revealed a highly significant positive correlation between the first and second assessment of the AAQW-Rtotal score (r = .80), AAQW-Rfood as control (r = .81), AAQW-Rweight as barrier to living (r = .74) and AAQW-Rweight-stigma subscales (r = .78). Additionally, the *t*-tests (paired samples) analyses showed no significant

differences between the two assessments for AAQW-Rtotal score (t(29) = -.395, p = .696), AAQW-Rfood as control (t(29) = -.137, p = .137), AAQW-Rweight as barrier to living subscales (t(29) = .340, p = .736) and AAQW-Rweight-stigma (t(29) = .000, p = 1.000).

3.8. Sensitivity to change

In order to examine the AAQW-R's sensitivity to clinical change we compared pre and post treatment scores after 12 sessions of a mindfulness and acceptance-based intervention for overweight and obese woman program using sample 3 (N= 58). Analysis of covariances (using baseline score as covariate) showed that, at post treatment, the experimental group reported significantly lower levels of overall weight-related experiential avoidance (F (1,56)= 10.052, *p*= .002, partial $\eta^2 = .16$ – large effects), AAQW-$R_{\text{food as control}}$ dimension (F (1,56)= 12.791, *p*= .001, partial $\eta^2 = .19$ – large effects), AAQW-$R_{\text{weight as barrier to living}}$ dimension (F (1,56)= 7. 643, *p*= .008, partial $\eta^2 = .12$ – intermediate effects). However, results for AAQW-$R_{\text{weight-stigma}}$ did not reach a statistically significant result (F (1,56)= 3. 055, *p*= .086).

4. Discussion

Obesity is a significant public health problem and innovative treatment targets, such as experiential avoidance, are needed to drive technological evolution. Researchers have articulated the importance of developing content specific measures of experiential avoidance that are able to account for changes in important psychological processes that relate to treatment change (Lillis & Hayes, 2008; Sandoz, Wilson, Merwin, & Kellum, 2013). The AAQW is a widely used measure of experiential avoidance related to one's weight (Lillis & Hayes, 2008; Weineland et al., 2012) with a factor structure that required further development. The current study tested and compared the model fit of the two proposed factor structures in a mixed sample (women from the general population and women with overweigh and obesity seeking nutritional treatment).

The confirmatory factor analysis of a proposed five-factor structure showed a poor fit to the data. On the other hand, the three-factor structure derived from the Portuguese exploratory factor analysis (Cardoso, 2014) presented an adequate model fit and was superior when compared to the five-factor model. However, several items did not meet statistical and theoretical justification for inclusion and were eliminated. Additionally, the revised scale was tested as a second order, unified factor, which was shown to be statistically equivalent to the three-factor solution and may provide a more parsimonious interpretation of the data (Chen et al., 2005) while also being consistent with the original analysis of the AAQW (Lillis & Hayes, 2008).

The final revised version of the AAQW-R comprises 10 items (from the original 22 item AAQW) distributed in three-factors: AAQW-$R_{\text{food as control}}$ (items 2, 16 and 17); AAQW-$R_{\text{weight as}}$

$_{\text{barrier to living}}$ (items 9, 11 and 13) and AAQW-R$_{\text{weight-stigma}}$ (items 10, 19, 20 and 21). We suggest that this revised and shortened version be referred to as the AAQW-Revised (AAQW-R). The analyses presented in this manuscript support using the AAQW-R primarily as a unifactor measure of weight-related experiential avoidance. Additionally, when clinically or theoretically useful, it can also be used as a three-subfactor measure that can provide separate scores for food as control, weight as barrier to living, and weight-related stigma.

Results also support that the AAQW-R is a reliable measure, which is in line with the results found in previous studies (Cardoso, 2014; Lillis & Hayes, 2008; Weineland et al., 2012). Also, the three factors obtained adequate internal consistency, good composite reliability values, and adequate variance extracted measure, which seem to provide evidence for the AAQW-R reliability.

The current study is the first to confirm the measurement invariance of AAQW-R across two groups: women from the general population and women with overweight and obesity seeking nutritional treatment. These results suggest that the structure of the AAQW-R is consistent when assessing weight related experiential avoidance patterns across different BMI groups. In addition, participants who are overweight or obese (BMI < 25) presented significantly higher levels of *weight-related* experiential avoidance patterns when compared to women within a normative BMI range (BMI > 25), suggesting that the AAQW-R is sensitive to varying levels of weight-related experiential avoidance among groups of participants in varying BMI ranges.

Concerning the relationships between AAQW-R and other measures, results corroborate previous research (Cardoso, 2014; Lillis & Hayes, 2008; Weineland et al., 2012) and generally fit our predicted pattern. As expected, the AAQW-R was found to be highly positively associated with eating pathology, moderately associated with BMI, and less strongly associated with subjective happiness, a construct that would be considered only weakly related to weight-related experiential avoidance. The correlation with external shame was slightly higher than expected, and the correlation with general experiential avoidance was slightly lower than expected, with both being squarely in the moderate range. In retrospect, the correlation with shame seems logical, given that the AAQW-R contains a stigma sub factor; a variable that is highly correlated with shame. The moderate correlation with general experiential avoidance is likely a positive indication that the AAQW-R is tapping into a related, but distinct domain, and is thus a welcome departure from our prediction. Overall, however, the general pattern was consistent with expectation and the AAQW-R seemed to display good convergent and divergent construct validity in the current study.

Test-retest reliability results supported the temporal stability of AAQW-R total as well as its three factors, which corresponds with the results found by Weineland et al. (2012) in a four-week period. The AAQW-R appears to show adequate temporal stability.

Finally, one of our main goals was to test whether AAQW-R was able to account for therapeutic changes after a 12-session mindfulness and acceptance-based group intervention for overweight and obese women. According to covariance analysis, the AAQW-R and its subscales (except for the *weight stigma* subscale) proved to be highly sensitive to clinical change, showing medium to large effects in a sample of 58 female participants.

This study has limitations that need to be taken into consideration when interpreting the results. First, the sample was comprised solely of adult women, which prevents from generalizing these results to adolescents females and males with overweight and obesity. Moreover, all data were collected via self-report, which can be biased.

Summary

Taken together, the current study offers new advances in the assessment of weight-related experiential avoidance. This study presents a revised, more rigorously tested version of AAQW, called the AAQW-Revised (AAQW-R), comprised of 10 of the original 22 AAQW items, containing three subfactors (food as control, weight as barrier to living and weight-stigma) and representing a global second-order factor, weight-related experiential avoidance. AAQW-R seems to be an improved, short, reliable, stable, and easy to use instrument to assess weight-related experiential avoidance, which has been consistently linked to negative health-related outcomes. In addition, the AAQW-R appears to have clinical utility, particularly for women with overweight and obesity.

Acknowledgments

This research was supported by the first author´s Ph.D. Grant (SFRH/BD/84452/2012), sponsored by FCT (Portuguese Foundation for Science and Technology) and European Social Fund (POPH). Also, the authors would like to acknowledge the nutritionists from Figueira da Foz District Hospital (HDFF), Eiras and Figueira da Foz primary care units for their valuable collaboration with the data collection.

References

Arbuckle, J.L. (2008). *Amos 17 users' guide*. Chicago, IL: SPSS.

Avenell, A., Brown, T. J., McGee, M. A., Campbell, M. K., Grant, A. M., Broom, J., Jung, R.T., & Smith, W. C. S. (2004). What are the long-term benefits of weight reducing diets in adults? A systematic review of randomized controlled trials. *Journal of Human Nutrition and Dietetics, 17*(4), 317-335. doi: 10.1111/j.1365-277X.2004.00531.x

Baer, R., Fischer, S., & Huss, D., (2006) Mindfulness And Acceptance In The Treatment Of Disordered Eating. *Journal of Rational-Emotive & Cognitive-Behavior Therapy*, *23*(4), 281-300. doi:10.1007/s10942-005-0015-9

Bond F.W., Hayes S.C., Baer R.A., Carpenter K.M., Guenole N., Orcutt H.K., Waltz T., & Zettle R.D. (2011). Preliminary psychometric properties of the Acceptance and Action Questionnaire-II: A revised measure of psychological flexibility and experiential avoidance. *Behavior Therapy*, *42*, 676-688.

Brown, T. A. (2006). *Confirmatory factor analysis for applied research*. New York: Guilford.

Byrne, S., Cooper, Z., & Fairburn, C. (2003). Weight maintenance and relapse in obesity: A qualitative study. *International Journal of Obesity and Related Metabolic Disorders*, *27*, 955–962. doi: 10.1038/sj.ijo.802305.

Byrne, B. M. (2010). *Structural equation modeling with AMOS* (2nd ed). New York: Routledge.

Cardoso, S. (2014). A Inflexibilidade Psicológica na Obesidade: estudo das propriedades psicométricas do AAQ-W [Psychological inflexibility in Obesity: Study of the psychometric properties of the AAQ-W]. (Unpublished master's thesis). Faculty of Psychology and Educational Sciences. Coimbra, Portugal.

Chen, F.F., Sousa K. H., & West S. G. (2005). Teacher's Corner: Testing Measurement Invariance of Second-Order Factor Models, Structural Equation Modeling. *A Multidisciplinary Journal, 12*(3), 471-492, doi: 10.1207/s15328007sem1203_7.

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences*. Hillsdale: Erlbaum.

Duarte, C., Pinto-Gouveia, J. & Ferreira, C. (2015). Expanding binge eating assessment: Validity and screening value of the Binge Eating Scale in women from the general population. *Eating Behaviors, 18,* 41–47 doi:10.1016/j.eatbeh.2015.03.007

Elfhag, K. & Rössner, S. (2004). Who succeeds in maintaining weight loss? A conceptual review of factors associated with weight loss maintenance and weight regain. *Obesity Reviews*, *6*, 67–85.

Fairburn, C. & Beglin, S. (1994). Assessment of eating disorders: Interview or self-report questionnaire? *International Journal of Eating Disorders*, *16*(4), 363–370. doi:10.1002/1098-108X(199412)16:4<363::AID-EAT2260160405>3.0.CO;2-#

Fassino, S., Leombruni, P., Piero, A., Daga, G. A., Amianto, F., Rovera, G., & Rovera, G. G. (2002). Temperament and character in obese women with and without binge eating disorder. *Comprehensive Psychiatry*, *43*(6), 431-437. doi: 10.1053/comp.2002.35906

Ferreira, N.B., Eugenicos, M.P., Morris, P.G., & Gillanders, D.T. (2013) Measuring acceptance in irritable bowel syndrome: preliminary validation of an adapted scale and construct utility. *Quality of Life Research*, *22*(7), 1761-6. doi: 10.1007/s11136-012-0299-z.

Field, A. (2013). *Discovering Statistics using IBM SPSS Statistics* (4th ed.). Los Angeles, London: SAGE.

Gifford, E. V., Kohlenberg, B. S., Hayes, S. C., Antonuccio, D. O., Piasecki, M. M., Rasmussen-Hall, M. L., & Palm, K. (2004). Acceptance based treatment for smoking cessation. *Behavior Therapy*, *35*, 689–705. doi: 10.1016/S0005-7894(04)80015-7.

Gormally, J., Black, S., Daston, S., & Rardin, D. (1982). The assessment of binge eating severity among obese persons. *Addicive Behaviors*, *7*, 47-55.

Gregg, J. A., Callaghan, G. M., Hayes, S. C., & Glenn-Lawson, J. L. (2007). Improving diabetes self-management through acceptance, mindfulness, and values: a randomized controlled trial. *Journal of Consulting and Clinical Ps*ychology, *75*(2), 336-343.

Hair, J. Anderson, R., Tatham, R., & Black, W. (1998). *Multivariate data analysis*. *Upper Saddle River*, NJ: Prentice Hall.

Hayes, S. C., Strosahl, K., & Wilson, K. G. (1999). *Acceptance and commitment therapy: An experiential approach to behavior change.* New York: Guilford Press.

Hu, L., &Bentler, P. M. (1998). Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. *Psychological Methods, 3*, 424–453.

Iacobucci, D. (2010). Structural equations modeling: Fit indices, sample size, and advanced topics. *Journal of Consumer Psychology, 20*(1), 90-98. doi: 10.1016/j.jcps.2009.09.003

Jöreskog, K.G., & Sörbom, D. (1996). *LISREL 8 user's reference guide*. Chicago: Scientific Software International.

Kayman, S., Bruvold, W., & Stern, J. S. (1990). Maintenance and relapse after weight-loss in women - behavioral- aspects. *American Journal of Clinical Nutrition*, *52*(5), 800- 807.

Kline, R. B. (2005). *Principles and Practice of Structural Equation Modeling* (2nd ed.). New York: Guilford.

Lillis, J., & Hayes, S.C. (2008). Measuring avoidance and inflexibility in weight related problems. *International Journal of Behavioral Consultation and Therapy*, *4*(4), 348-354.

Lillis, J., Hayes, S. C., Bunting, K., & Masuda, A. (2009). Teaching acceptance and mindfulness to improve the lives of the obese: A preliminary test of a theoretical model. *Annual Behavior Medicine*, *37*, 58-69. doi: 10/1007/s12160-009-9083-x.

Lillis, J., Hayes, S., & Levin, M. (2011). Binge Eating and Weight Control: The role of experiential avoidance. *Behavior Modification*, *35*, 3252-3264. doi: 10.1177/0145445510397178.

Lundgren, T., Dahl, J., & Hayes, S. C. (2008). Evaluation of mediators of change in the Treatment of epilepsy with Acceptance and Commitment Therapy. *Journal of Behavior Medicine, 31*(3),225–235. http://dx.doi.org/10.1007/s10865-008-9151-x.

Lyubomirsky, S., & Lepper, H. S. (1999). A measure of subjective happiness: Preliminary reliability and construct validation. *Social Indicators Research, 46*(2), 137-155. doi: 10.1023/A:1006824100041

Matos, M., Pinto-Gouveia, J., Gilbert, P., Duarte, C., & Figueiredo, C. (2015). The Other As Shamer Scale – 2: Development and validation of a short version of a measure of external shame. *Personality and Individual Differences*, 74,6-11. doi: 10.1016/j.paid.2014.09.037

Machado, P., Martins, C., Vaz, A., Conceição, E., Pinto-Basto, A., & Gonçalves, S. (2014). Eating Disorder Examination Questionnaire: Psychometric properties and norms for the Portuguese Population. *European Eating Disorders Review*, *22*(6), 448–453. doi:10.1002/erv.2318

Meredith, W. (1993). Measurement invariance, factor analysis and factorial invariance. *Psychometrika*, *58*(4), 525-543. doi:10.1007/BF02294825

Merwin, R. M., Timko, C. A., Moskovich, A. A., Ingle, K. K., Bulik, C. M., & Zucker, N. L. (2011). Psychological Inflexibility and Symptom Expression in Anorexia Nervosa. *Eating Disorders*, *19*, 62-82. doi: 10.1080/10640266.2011.533606

Ogden, J. (2000). The correlates of long-term weight loss: A group comparison study of obesity. *International Journal of Obesity*, *24*(8), 1018-1025. doi:10.1038/sj.ijo.0801354

Pais-Ribeiro, J. L. (2012). Validação transcultural da escala de felicidade subjectiva de Lyubomirsky e Lepper. *Psicologia saúde & doenças,13*(2), 157-168.

Pinto-Gouveia, J., Gregório, S., Dinis, A., & Xavier, A. (2012). Validation of the Portuguese Version of Acceptance and Action Questionnaire (AAQ-II). *International Journal of Psychology & Psychological Therapy*, *12* (2), 139-156.

Sandoz, E. K., Wilson, K. G., Merwin, R. M., & Kellum, K. (2013). Assessment of body image flexibility: The Body Image – Acceptance and Action Questionnaire. *Journal of Contextual Behavioral Science*, *2*(1–2), 39–48. doi:10.1016/j.jcbs.2013.03.002i

Schermelleh-Engel, K., Moosbrugger, H., & Müller, H. (2003). Evaluating the fit of structural equation models: Test of significance and descriptive goodness-of-fit measures. *Methods of Psychological Research - Online*, *8*(2), 23-74.

Tabachnick, B & Fidell, L. (2007). *Using multivariate statistics*. New York: Pearson Education Inc.

Tapper, K., Shaw, C., Ilsley, J., Hill, A,J., Bond, F.W., & Moore, L. (2009). Exploratory randomised controlled trial of a mindfulness-based weight loss intervention for women. *Appetite*, *52*(2), 396–404. doi: 10.1016/j.appet.2008.11.012

Weineland, S., Lillis, J. & Dahl, J. (2012) Measuring experiential avoidance in a bariatric surgery population–Psychometric properties of AAQ-W. *Obesity Research & Clinical Practice*, *7*(6), e464–e475. doi:10.1016/j.orcp.2012.06.002

World Health Organization (2011). *World Health statistics*. WHO: France. ISBN 978 92 4 156419 9.