



## Cross-cultural measurement invariance of the Body Appreciation Scale-2 across five countries

Magdalena Razmus<sup>a,\*</sup>, Wiktor Razmus<sup>b</sup>, Tracy L. Tylka<sup>c</sup>, Marija Jović<sup>d</sup>, Marko Jović<sup>e</sup>, Hikari Namatame<sup>f</sup>

<sup>a</sup> Institute of Psychology, Maria Curie-Skłodowska University, Plac Litewski 5, 20-080, Lublin, Poland

<sup>b</sup> Institute of Psychology, The John Paul II Catholic University of Lublin, Al. Raclawickie 14, 20-950, Lublin, Poland

<sup>c</sup> Department of Psychology, The Ohio State University, 1465 Mt. Vernon Avenue, Marion, OH, 43302, United States

<sup>d</sup> Department of Marketing Management and Public Relations, University of Belgrade, Jove Ilića 154, 11000, Belgrade, Serbia

<sup>e</sup> Department of Plastic and Reconstructive Surgery, School of Medicine, University of Belgrade, Zvezanska 9, 11000, Belgrade, Serbia

<sup>f</sup> Faculty of Human Sciences, University of Tsukuba, 1-1-1 Tennodai, Tsukuba 305-8572, Japan

### ARTICLE INFO

#### Article history:

Received 2 February 2020

Received in revised form 9 July 2020

Accepted 9 July 2020

Available online 22 July 2020

#### Keywords:

Body appreciation

Measurement invariance

Adults

Cross-cultural examination

Gender

Age

### ABSTRACT

The Body Appreciation Scale-2 (BAS-2; Tylka & Wood-Barcalow, 2015a, 2015b) is a widely used measure of positive body image within many cultures and countries; yet, cross-cultural examinations are few. The present study aimed to investigate the measurement invariance of the BAS-2 across adults from five countries: Iran, Japan, Poland, Serbia, and the U.S. The sample included 2944 participants ranging in age from 18 to 82 years. The findings provided evidence that the BAS-2's one-dimensional structure is the same in these countries. Partial metric invariance (when some but not all items contribute to a latent construct equally for groups) indicated that nine out of 10 items contributed to the latent body appreciation construct to a similar degree across the countries. When a Multiple Indicators Multiple Causes (MIMIC) model was applied with participants' age and gender as covariates, evidence of differential item functioning was found. Results suggest that both age and gender influenced body appreciation, indicating that the body appreciation factor means are different at different levels of the covariates. In conclusion, cross-culturally body appreciation may be shaped by country, language, age, and gender correlates to different degrees. In future research, measurement invariance analyses should be conducted prior to cultural group comparisons on the BAS-2.

© 2020 Elsevier Ltd. All rights reserved.

### 1. Introduction

Positive body image has attracted a considerable amount of attention from researchers in the last decade (Tylka et al., 2019; Tylka & Wood-Barcalow, 2015b). Particularly, body appreciation has been the core positive body image construct within these investigations, and body appreciation is linked to psychological well-being in a variety of domains, such as self-compassion and life satisfaction (Cox, Ullrich-French, Tylka, & McMahon, 2019) Tylka et al., 2019). As originally defined by Avalos, Tylka, and Wood-Barcalow (2005), body appreciation is characterized by: (a) respecting and taking care of one's body by attending to its needs; (b) accepting one's body as it is, regardless of its imperfections;

and (c) protecting one's body by rejecting the internalization of unrealistic standards projected in mass media. Body appreciation is significantly associated with optimism, self-esteem, and proactive coping, even after controlling for appearance evaluation, body pre-occupation, and body dissatisfaction (Avalos et al., 2005). It is also significantly associated with eating behavior (positively related to intuitive eating and negatively related to disordered eating) after controlling for body dissatisfaction and appearance evaluation (Tylka & Wood-Barcalow, 2015a). Such findings highlight body appreciation's unique relevance to psychological well-being and demonstrate that body appreciation is broader in scope than body and appearance satisfaction.

To assess body appreciation, the 13-item Body Appreciation Scale (BAS) was developed and psychometrically evaluated by Avalos et al. (2005). The BAS had a one-dimensional factor structure as well as satisfactory construct validity and internal consistency in this parent study with college women from the U.S. The measure was adapted (i.e., translated when appropriate) and used both in Western (Jáuregui Lobera & Bolaños Ríos, 2011; Swami, Hadji-Michael, & Furnham, 2008; Swami, Stieger, Haubner, & Voracek,

\* Corresponding author.

E-mail addresses: [magdalena.razmus@gmail.com](mailto:magdalena.razmus@gmail.com) (M. Razmus), [wrazmus@gmail.com](mailto:wrazmus@gmail.com) (W. Razmus), [tylka.2@osu.edu](mailto:tylka.2@osu.edu) (T.L. Tylka), [marija.jovic@fon.bg.ac.rs](mailto:marija.jovic@fon.bg.ac.rs) (M. Jović), [markojovic01@yahoo.com](mailto:markojovic01@yahoo.com) (M. Jović), [hikari.namatame@gmail.com](mailto:hikari.namatame@gmail.com) (H. Namatame).

2008) and non-Western populations (Atari, Akbari-Zardkhaneh, Mohammadi, & Soufiabadi, 2015; Ng, Barron, & Swami, 2015; Swami & Chamorro-Premuzic, 2008; Swami & Jaafar, 2012; Swami et al., 2011; Swami, Hwang, & Jung, 2012; Swami, Mada, & Tovée, 2012). Interestingly, all above-mentioned non-Western studies on the BAS evidenced two factors which may be conceptualized as general body appreciation and body image investment, whereas American, Austrian, and Spanish versions (Avalos et al., 2005; Jáuregui Lobera & Bolaños Ríos, 2011; Swami, Stieger et al., 2008) supported a one-dimensional structure of the scale. These results suggested that body appreciation measured by the BAS may be differently considered across cultures.

Taking into account these inconsistencies and other psychometric limitations of the BAS, a refined version of the scale (BAS-2) was developed (Tylka & Wood-Barcalow, 2015a). Retaining five high-loading items from the BAS and devising five gender-neutral items (which more closely aligned with the positive body image literature that had emerged since the development of the original BAS), Tylka and Wood-Barcalow (2015a) reported that the BAS-2 had adequate psychometric properties across community and college samples from the U.S. Specifically, the original BAS-2 upheld a unidimensional structure and had high internal consistency reliability coefficients, and temporal stability over three weeks. Moreover, construct and incremental validity of the scale were evidenced. Measurement invariance (MI) analyses showed that the BAS-2 was similarly unidimensional across women and men. These findings were quite promising, and the authors encouraged scholars to further examine the psychometric characteristics of the BAS-2 across cultures, ethnicities, and geographic regions (Tylka & Wood-Barcalow, 2015a).

Thus far, scholars have investigated the psychometric properties of the BAS-2 across some national contexts. Specifically, the BAS-2 has been translated into Cantonese (Swami & Ng, 2015), Persian (Atari, 2016), Dutch (Alleva, Martijn, Veldhuis, & Tylka, 2016), Standard Chinese (Swami, Ng, & Barron, 2016), Serbian (Jovic, Sforza, Jovanovic, & Jovic, 2017), Icelandic (Pálmarsdóttir & Karlsdóttir, 2016), both Portuguese (Lemoine et al., 2018; Marta-Simões, Mendes, Oliveira, Trindade, & Ferreira, 2016) and Brazilian Portuguese (Alcaraz-Ibáñez, Cren Chiminazzo, Sicilia, & Teixeira Fernandes, 2017), Polish (Razmus & Razmus, 2017), French (Kertechian & Swami, 2017), Spanish (Swami, García, & Barron, 2017), Romanian (Swami, Tudorel, Goian, Barron, & Vintila, 2017), Japanese (Namatame, Uno, & Sawamiya, 2017), Danish (Lemoine et al., 2018), Swedish (Lemoine et al., 2018), Arabic (Vally, D'Souza, Habeeb, & Bensusmaide, 2019), Malay (Swami et al., 2019), Greek (Argyrides, 2019), Latin-American Spanish (Góngora, Cruz Licea, Mebarak Chams, & Thornborrow, 2020), and Lithuanian (Baceviciene & Jankauskiene, 2020). From these studies, it was demonstrated that all translated versions had a one-dimensional structure and were psychometrically sound measures of body appreciation with evidence of adequate validity and reliability. Moreover, studies in Portugal (Lemoine et al., 2018), Poland (Razmus & Razmus, 2017), Spain (Swami, García et al., 2017), China (Swami et al., 2016), Sweden (Lemoine et al., 2018), Mexico, Colombia, and Argentina (Góngora et al., 2020), and Lithuania (Baceviciene & Jankauskiene, 2020) provided evidence for full measurement invariance of the BAS-2 across gender, whereas research in Brazil (Alcaraz-Ibáñez et al., 2017), Denmark (Lemoine et al., 2018), France (Kertechian & Swami, 2017), Japan (Namatame et al., 2017), Malaysia (Swami et al., 2019), and Romania (Swami, Tudorel et al., 2017) demonstrated partial scalar invariance of the scale across men and women (i.e., whereby individuals' scores on some items stem not from real differences in latent means, but from diverse functioning of the scale across gender). The BAS-2 scores were also shown to be partially invariant across weight status (Alcaraz-Ibáñez et al., 2017).

Thus far, there have been only four studies exploring the BAS-2 measurement invariance across cultures. One investigated the scale equivalence across adolescents and young adults in three countries – Portugal, Denmark, and Sweden (Lemoine et al., 2018) and evidenced partial scalar invariance across Denmark and Portugal, Denmark and Sweden, and Portugal and Sweden. Two others showed that the BAS-2 scores were partially scalar invariant across adults from two major ethnic groups in Malaysia (Swami et al., 2019) and across adults from Malaysia and the United Kingdom (Todd & Swami, 2020). The last confirmed scalar invariance across Mexican, Argentinean, and Colombian adolescents (Góngora et al., 2020).

We aimed to investigate measurement invariance of the BAS-2 across cultures for three reasons. First, only the four above-mentioned studies have examined this issue. Second, since three of these studies demonstrated partial scalar invariance, which means that scores on some items may be related to being a part of a particular group, we intended to verify these findings in alternative cultural contexts. Third, these previously examined cultural comparisons included adolescents from three Latin-American countries (Mexico, Argentina, and Colombia; Góngora et al., 2020), adolescents and young adults aged 12–19 from two Nordic European countries (Denmark and Sweden) and a Latin European country (Portugal; Lemoine et al., 2018) and two ethnic groups inhabiting one country (Malaysian Malays and Malaysian Chinese; Swami et al., 2019). To the best of our knowledge, only one study has investigated measurement equivalence of the BAS-2 across adult samples between two countries (Todd & Swami, 2020). An investigation of the scale's measurement equivalence across adults from many diverse world regions has yet to be conducted. The present research was designed to address this gap.

Broadly, tests of MI were conducted to determine whether (a) a similar latent variable (i.e., body appreciation) is present across cultural groups as determined by the parameter patterns across groups (i.e., configural invariance), (b) the loadings for items on the latent factor are the same across groups (i.e., metric invariance), and (c) the intercept of the regression relating each item to the latent factor is the same across groups (i.e., scalar invariance). Methodologically, if the mentioned measurement invariance criteria are evidenced for the scale, it might be assumed that the scale is assessing the exact same latent variable across subgroups, thus comparisons of mean scores across these different national samples are allowed (Meredith, 1993). On the other hand, if properties of the questionnaire vary across groups, the observed scores cannot be directly compared as they are on different metrics. As a result, observed differences do not reflect differences on the latent variable, but are result of measurement bias. Therefore, some degree of measurement invariance is essential for valid interpretations of research results, specifically for valid group comparisons, protecting against erroneous conclusions and providing information about group differences and similarities. There are two degrees of invariance, namely full and partial invariance. The former holds when all parameters of all items are the same across subgroups, whereas the latter is established when parameters for only a subset of items are invariant. Specifically, in partial metric invariance the loadings for some items on the latent factor differ across groups, meaning that these items do not contribute to a latent construct equally for these groups. Yet, in partial scalar invariance intercepts for some items are not equal across samples, indicating that individuals with the same score on the latent variable do not obtain the same score on these items (Lai, Richardson, & Mak, 2019; Wang, Chen, Dai, & Richardson, 2018).

We decided to analyze the data collected with the original BAS-2 (U.S. sample) and the data sets from other countries selected on the basis of the following criteria: (a) the report on the BAS-2 linguistic version was published in a peer-reviewed journal; (b) the

**Table 1**  
Characteristics of the sample by country.

Variable	Iran (n = 550)	Japan (n = 672)	Poland (n = 721)	Serbia (n = 431)	U.S. (n = 570)
Women (%)	299 (54.4)	359 (53.4)	383 (53.1)	236 (54.8)	272 (47.7)
Men (%)	251 (45.6)	313 (46.6)	338 (46.9)	195 (45.2)	298 (52.3)
Age range	18–46	18–27	18–57	18–82	18–74
Age (M, SD)	25.17; 4.47	19.19; 1.31	35.18; 10.82	42.67; 14.24	33.72; 10.48
BMI range	no data	15.24–39.25	15.46–39.86	16.11 – 38.78	16.64–69.43
BMI (M, SD)	no data	21.02; 2.92	24.09; 3.75 <sup>a</sup>	24.58; 4.28 <sup>a</sup>	26.83; 7.24

Note. Mean values with the same superscript letter (a) were similar and no statistically significant differences were observed for these samples.

validity of a linguistic translation of the BAS-2 was demonstrated; (c) participants were at minimum 18 years old; (d) the sample consisted of both women and men; (e) and the authors of the BAS-2 translations agreed to participate in the study and provided their data. Thereby, our study and analysis included samples from five countries: Iran, Japan, Poland, Serbia, and the U.S. They represent three continents: Asia, Europe, and North America and four cultural clusters: Southern Asia (Iran), Confucian Asia (Japan), Eastern Europe (Poland and Serbia), and Anglo Cultures (the U.S.) and differ in terms of societal culture practices, societal values, and leadership style (Gupta, Hanges, & Dorfman, 2002; Nedeljković, Vukonjanski, Nikolic, Hadzic, & Sljukic, 2018). The included countries are also characterized by diverse socio-economic status indicators, such as average wages per month, education index, and Human Development Index (HDI). Specifically, the lowest average wages per month are reported in Iran (456\$), then in Serbia (533\$), Poland (1175\$), and Japan (3443\$), with the highest wages in the U.S. (5257\$) (World Data, 2019). The grading of the analyzed countries in terms of both education index and HDI corresponds to the rating on average wages (education index: Iran .71, Serbia .72, Poland .82, Japan .84, the U.S. .85; HDI: Iran .797, Serbia .799, Poland .872, Japan .915, the U.S. .920) (Human Development Report, 2019). Thus, our research does not focus on Western, Educated, Industrialized, Rich, and Democratic (WEIRD) societies (Henrich, Heine, & Norenzayan, 2010). What is more, it may be presumed that diversity of the selected countries refers also to body image-related variables. To our knowledge, there are no comparative studies on positive body image in countries which were included in our investigation. Yet, recent research demonstrates that these countries differ in terms of their levels of breast size dissatisfaction and its antecedents, such as frequency of local and Western media exposure (Swami et al., 2020). Thus, we hypothesize that individuals living in Iran, Japan, Poland, Serbia, and the U.S. may also demonstrate different levels of another body image construct: body appreciation.

Theory and research show that age and gender may be relevant covariates when examining cross-country comparisons of the BAS-2. Body appreciation may increase with age as the cultural pressure to meet sociocultural beauty standards decreases (Tiggemann & McCourt, 2013). Indeed, some researchers have found that body appreciation is positively related to age in various countries, such as the Netherlands, Iran, Australia, Serbia, and Hong Kong (Alleva et al., 2016; Atari, 2016; Jovic et al., 2017; Swami & Ng, 2015; Tiggemann & McCourt, 2013). Given that the BAS-2's measurement equivalence across age, to our knowledge, has not been investigated, we cannot be sure if this association does not stem from the scale's non-invariance. Therefore, we included age as a covariate in our analysis. In addition, women are more frequently exposed to unrealistic societal appearance ideals in everyday life compared to men (Buote, Wilson, Strahan, Gazzola, & Papps, 2011), and therefore are likely to have lower body appreciation than men. Indeed, a recent meta-analysis showed that men generally have a higher level of body appreciation than females (He, Sun, Zickgraf, Lin, & Fan, 2020). Investigations of the BAS-2's gender invariance in diverse cultural groups have provided equivocal results. For these reasons,

we decided to include gender as a covariate in our analyses. As still little is known about the BAS-2 measurement invariance in adults and differential functioning of the scale's items due to age and gender, we limited our investigation to two exploratory research questions: (1) Is there a significant direct effect of age and gender on the BAS-2 items? Specifically, do the items function differently due to participants' age and gender? and (2) Is the BAS-2 invariant across Iran, Japan, Poland, Serbia, and the U.S.?

## 2. Method

### 2.1. Participants and procedure

The total sample consisted of 2944 individuals from the five countries. Demographic characteristics of the sample by country are presented in Table 1.<sup>1</sup> As illustrated, differences in BMI between Poland and Serbia were nonsignificant, whereas other cultural samples differed in terms of BMI,  $F(3, 2356) = 153.57, p < .001$ . The mean values of BMI and their sequence in the investigated countries are consistent with World Health Organization data (World Health Organization, 2014) which indicate that BMI is the highest in U.S. and the lowest in Japan, with Poland and Serbia in the middle of the ratings. All five countries differed significantly in terms of mean age,  $F(4, 2939) = 224.48, p < .001$ .

Authors of the five BAS-2 linguistic versions shared the data collected in their studies. In all countries, responses were anonymous or confidential, and participants gave informed consent before they completed questionnaires. In Iran, participants were selected from three large universities in Tehran. They completed the Persian translation of the BAS-2 (Atari, 2016) in a paper-and-pencil format in classroom settings. Iranian participants were not compensated. In Japan, participants were students of six universities. They completed paper-and-pencil version of BAS-2 in a classroom setting. Participation in the study was voluntary. In Poland, participants were selected using convenience sampling, recruited via personal contacts of data collectors using direct solicitation. They filled in paper-and-pencil version of BAS-2 in their homes or worksites and returned it in a sealed envelope, with participation being voluntary and without remuneration. In Serbia, data collectors directly recruited participants through their personal contacts. Respondents were given the paper and pencil questionnaire to complete. They participated voluntarily, were tested individually, and were not remunerated. In U.S., community women and men were recruited from Amazon Mechanical Turk, an online website that employs workers for completing tasks. Participants received \$1.00 USD for completing the online survey; 46 of the 50 states were represented in the data. Attention checks were embedded in the survey to detect and exclude participants who were not reading the questions with full attention.

<sup>1</sup> The subgroups profiles do not include information about participants' ethnicity, since in majority of the samples such data were not collected.

## 2.2. Measure

The BAS-2 consists of 10 items (e.g., *I feel good about my body*) which are rated on a 5-point scale ranging from 1 (*never*) to 5 (*always*). Items are averaged, with higher scores indicating higher body appreciation. Scores on the original English version have evidenced good psychometric properties, including internal consistency, 3-week test-retest reliability, and construct validity (Tylka & Wood-Barcalow, 2015a). Persian (Atari, 2016), Japanese (Namatame et al., 2017), Polish (Razmus & Razmus, 2017), and Serbian (Jovic et al., 2017) translations of the BAS-2 also demonstrated good reliability and validity, indicating that these versions of the scale are psychometrically sound measures of body appreciation within their respective country.

## 2.3. Statistical analyses

First, descriptive statistics and Cronbach's alpha coefficients for internal consistency of the scale in each sample were computed using SPSS version 24. Second, we examined confirmatory factor analysis (CFA) for each group separately. Considering the multivariate non-normality estimate, the one-dimensional model was tested using maximum likelihood estimation with robust standard errors. The following indices were considered when assessing model fit: Satorra-Bentler scaled chi square (S-B  $\chi^2$ ), Root Mean Square Error of Approximation (RMSEA), Comparative Fit Index (CFI), and Standardized Root Mean Square Residual (SRMR). Values below .08 for the RMSEA and .09 for the SRMR, and values higher than .90 for the CFI indicate an acceptable fit (Schweizer, 2010).

Third, after the model fit was established, to control the simultaneous influence of age and gender on the results and to investigate measurement invariance in the sample due to these covariates, a MIMIC model was estimated. Although participants from analyzed countries differed not only in terms of age but also BMI, including the latter variable into MIMIC analysis was not possible due to missing data in Iran. In the MIMIC modeling, data on the latent variable (here, body appreciation) and characteristics of subgroups (here, age and gender) are regressed on the scale's items to examine if subgroup characteristics produce an effect on the individual items. Such an effect demonstrates measurement noninvariance and provides evidence that the affected items work differently in subgroups, which is called differential item functioning (DIF). As a result, the items confound scores of the scale (Lúcio et al., 2017). We applied the MIMIC model in an exploratory fashion, namely we added the paths of the covariates on the latent variable and the paths of the covariates on the 10 BAS-2 items fixing all direct effects between age and gender and items at zero. Subsequently, we checked modification indices (MI) to examine if relevant direct effects would be present. The MI > 4 indicate that model fit could be improved if the path is freely estimated. If a significant direct effect of age and/or gender on the items occurred after a model estimation, it was treated as a proof of DIF meaning that for the same level of the underlying latent trait the groups of different age and/or gender will endorse different levels of the rating scale for that item. A direct effect of the covariates on body appreciation (latent variable) was interpreted as evidence of population heterogeneity (Brown, 2015).

Fourth, after estimating the MIMIC model, we investigated measurement invariance with Multigroup Confirmatory Factor Analysis (MGCFAs) by fitting a sequence of increasingly restrictive models. The MGCFAs results were controlled for age and gender effects. The procedure included testing configural invariance (unconstrained model), metric invariance (all factor loadings are constrained to be equal), and scalar invariance (all intercepts are constrained to be equal) (Meredith, 1993). Absolute difference in CFI ( $\Delta$ CFI) was calculated. Following recommendations (Chen, 2007), a value of

**Table 2**

Means, standard deviations, skewness, kurtosis, and internal consistency of the BAS-2 across countries.

Country	Score range	M	SD	Skew	Kurt	$\alpha$
Iran	1.20–5.0	3.80	0.71	−0.34	−0.09	.91
Japan	1.00–5.0	2.61	0.86	0.28	−0.43	.92
Poland	1.00–5.0	3.58	0.85	−0.38	−0.47	.95
Serbia	1.00–5.0	3.82	0.73	−0.69	0.69	.93
U.S.	1.00–5.0	3.42	0.91	−0.35	−0.28	.96

**Table 3**

Goodness-of-fit indexes for one-factor model of the BAS-2 in each country.

Country	$\chi^2$	df	p	RMSEA <sup>a</sup>	SRMR <sup>b</sup>	CFI <sup>c</sup>
Iran	154.740	35	<.001	.079	.035	.942
Japan	169.413	33	<.001	.078	.035	.952
Poland	157.231	34	<.001	.071	.022	.972
Serbia	114.171	35	<.001	.072	.034	.963
U.S.	125.834	35	<.001	.067	.021	.974

<sup>a</sup> Root Mean Square Error of Approximation.

<sup>b</sup> Standardized Root Mean Square Residual.

<sup>c</sup> Comparative Fit Index.

$\Delta$ CFI smaller than or equal to .010 would indicate measurement invariance (Cheung & Rensvold, 2002). In case of lack of measurement invariance, we released the misspecified parameters to look for partial invariance. The MIMIC and the MGCFAs were conducted using Mplus v.7.0 (Muthén & Muthén, 2012),

## 3. Results

### 3.1. Descriptive statistics

There were no missing data in the Polish and Serbian datasets. A maximum of 1.5 %, 0.4 %, and 0.2 % of data were missing on any one BAS-2 item from Iranian, Japanese, and U.S. datasets, respectively. Missing data were handled using full information maximum likelihood (FIML). Means, standard deviations, skewness, and kurtosis of total scores, as well as internal consistency of the scale in each country are presented in Table 2.

### 3.2. Confirmatory factor analysis

We evaluated a one-factor structure of the BAS-2 by performing CFA in each sample. In 2 out of 5 samples, we took into consideration the error covariances between Items<sup>2</sup> 3 and 4 (Japan), Items 8 and 9 (Japan), and Items 1 and 5 (Poland), which is in line with what was demonstrated by analyses of Japanese (Namatame et al., 2017) and Polish (Razmus & Razmus, 2017) BAS-2 validations. The acceptable fit of the hypothesized model was established (Table 3). Factor loadings ranged from .58 to .79 in Iran, .59 to .84 in Japan, .67 to .87 in Poland, .65 to .85 in Serbia, and .71 to .90 in the U.S.

### 3.3. Differential item functioning

When direct paths from age and gender to each of the items were constrained to zero, the model presented adequate fit: S-B  $\chi^2(53) = 787.833$ ,  $p < .001$ , CFI = .959, RMSEA = .069 (90 % CI: .065, .073), SRMR = .027. Modification indices suggested to add paths from age to five items (Items 1, 7, 8, 9, and 10) and from gender to seven items (Items 1, 2, 3, 6, 8, 9, and 10), which means that age and gender were influencing responses to these items. After the modifications were implemented, the revised MIMIC model was run and the following model fit indices were demonstrated: S-B

<sup>2</sup> BAS-2 items are included in Table 4 for reference.

**Table 4**  
Effects of age and gender on BAS-2 items and body appreciation.

BAS-2 Item/Factor	Age			Gender		
	$\beta$	SE	<i>p</i>	$\beta$	SE	<i>p</i>
1. I respect my body.	.044	.012	.001	.041	.013	.002
2. I feel good about my body.	–	–	–	–.021	.011	.062
3. I feel that my body has at least some good qualities.	–	–	–	.060	.012	.001
4. I take a positive attitude towards my body.	–	–	–	–	–	–
5. I am attentive to my body's needs.	–	–	–	–	–	–
6. I feel love for my body.	–	–	–	–.015	.011	.174
7. I appreciate the different and unique characteristics of my body.	–.111	.013	.001	–	–	–
8. My behavior reveals my positive attitude toward my body; for example, I hold my head high and smile.	–.045	.014	.001	.067	.013	.001
9. I am comfortable in my body.	–.029	.011	.011	–.048	.012	.001
10. I feel like I am beautiful even if I am different from media images of attractive people (e.g., models, actresses/actors).	.050	.013	.001	.034	.013	.009
Body appreciation (total score)	.227	.018	.001	–.047	.019	.015

**Table 5**  
Measurement invariance across countries.

Country	$\chi^2$	<i>df</i>	<i>p</i>	CFI <sup>a</sup>	Model comparison	$\Delta$ CFI
M1. Configural invariance	910.797	205	<.001	.957	–	–
M2. Metric invariance	1151.049	245	<.001	.945	M2 vs M1	.012
M3. Partial metric invariance	1107.847	244	<.001	.948	M3 vs M1	.009

Note. <sup>a</sup> Comparative Fit Index.

$\chi^2(41) = 561.753$ ,  $p < .001$ , CFI = .971, RMSEA = .066 (90 % CI: .061, .071), SRMR = .019. This model provided a significantly better fit than model with paths constrained to zero,  $\Delta$ S-B  $\chi^2(12) = 226.080$ ,  $p < .001$ . Results demonstrated that age significantly affected the intercepts of five items. The negative values of  $\beta$  coefficients indicated that older participants reported lower scores on Items 7, 8 and 9, whereas the positive value of  $\beta$  coefficient suggested that they obtained higher scores on Items 1 and 10 (see Table 4). Moreover, older age predicted higher body appreciation ( $\beta = .227$ ,  $p < .001$ ). Findings also showed that gender significantly affected the intercepts of five items (Items 1, 3, 8, 9, and 10) and the latent variable ( $\beta = -.047$ ,  $p < .015$ ). Women scored higher than men on Items 1, 3, 8, and 10 and lower on Item 9 and on the latent variable. Significant direct effects of the covariates on the items and the latent variable indicate that the observed score differences between representatives of age and gender groups are not related to real differences in the trait level, but to noninvariance of the scale.

#### 3.4. Measurement invariance across countries

Next, using the MGCFA, we examined measurement invariance controlling for age and gender effects (see Table 5). To facilitate comparability across samples, no correlated error terms were included in the model. The configural model (M1) was supported and showed an adequate fit. Then, the factor loadings were constrained to be equal (M2). The drop in CFI was higher than assumed ( $\Delta$ CFI = .012); thus, full metric invariance was not supported. In the next step, the model was tested for partial metric measurement invariance releasing factor loadings with the largest misspecifications. The factor loading of Item 10 in Japan was relaxed. After releasing the factor loading partial metric invariance was achieved, with  $\Delta$ CFI = .009 (M3).

## 4. Discussion

The present study was designed to examine measurement invariance of the BAS-2 across five different countries: Iran, Japan, Poland, Serbia, and the U.S. To our knowledge, this is the first study which tested measurement equivalence of body appreciation across more than two countries using large adult samples. Findings

of the MGCFA provided evidence that the BAS-2 structure is the same in the selected countries. Moreover, partial metric invariance indicated that not all items contributed to the latent construct to a similar degree across the countries (Putnick & Bornstein, 2016); thus, meaningful comparisons of relationships between variables, but not of latent means, are permitted. Evidence of differential item functioning was found when MIMC modeling was applied, considering participants' age and gender. The findings suggest that both age and gender shaped responses to five out of 10 BAS-2 items. Our results demonstrate that age and gender significantly affect the body appreciation latent variable, with older age related to higher scores on the scale and women scoring higher than men on body appreciation. It implies that the observed differences between representatives of age and gender groups do not reflect, in fact, different levels of body appreciation, but are associated with noninvariance of the scale (i.e., there are differences in items responding which influence the score, regardless of the latent variable level).

The present study is the first which demonstrated partial metric invariance of the BAS-2 across various countries from diverse world regions, namely Iran, Japan, Poland, Serbia, and the U.S. Previously, either partial scalar (Lemoine et al., 2018; Swami et al., 2019; Todd & Swami, 2020) or full scalar invariance was evidenced (Góngora et al., 2020). Although the five BAS-2 versions included in the present study demonstrate very good psychometric properties for each country separately, results of measurement invariance and differential item functioning investigation do not allow direct comparability of body appreciation between the countries. These findings suggest that there is no need to use different forms of the BAS-2, but caution is recommended for Item 10 in Japan. Nevertheless, researchers should conduct measurement invariance analyses before conducting comparisons between cultures on the BAS-2. More specifically, body appreciation may be shaped not only by culture (country) and language, but also by age and gender to different degrees across populations. Thereby, when comparisons of the BAS-2 scores are conducted between countries, it is advised to control participants' age and gender. Although the present study's design does not allow to identify the exact nature of the differences in items functioning, there are several possible explanations for them. The noncomparability of the items between countries may be caused by differences in phenomenology of body appreciation

and its content, dissimilarities in cultural habits related to body, and/or by translations of the scale and linguistic issues. Considering age and gender, presumably representatives of different age and gender groups have diverse experiences and are exposed to diverse situations related to body image and have different ideas about what it means to appreciate characteristics of one's body. Further studies are needed to address this issue, as well as to identify other factors possibly contributing to the BAS-2 measurement non-invariance and differential item functioning.

Some limitations of the research should be acknowledged. First, only five of the 21 known BAS-2 translations were taken into consideration. Further studies would benefit from including more groups across diverse countries and applying different linguistic translations of the scale. Second, generalizability of the findings could be limited by using samples of convenience. Finally, although we demonstrated that the BAS-2 scalar measurement invariance does not hold across the five countries and age and gender influence item functioning, it is possible that there are other potential predictors of invariance which were not regarded in the current study. We were not able to include BMI in the research due to missing data in Iran. Future research is needed to address this gap, as well as to identify other sociodemographic variables (such as socioeconomic status), beyond country, age, and gender which may be possible sources of the noninvariant results.

In conclusion, the present study contributes to the examination of the BAS-2's psychometric properties across diverse populations by presenting results of measurement invariance and differential item functioning investigation of the scale across five countries. This is the first time that the BAS-2 equivalence across adult cross-country samples was evaluated using multiple language versions of the scale and a large sample size from various countries differing in terms of cultural background. It is notable that this study goes beyond WEIRD societies, which allows researchers to extend the generalizability of the findings to multicultural contexts. Moreover, a combined approach of MGCFAs and MIMIC analyses of the BAS-2 were employed. Prior to multigroup analysis, differential item functioning was investigated, taking into account participants' age and gender. It provided the opportunity to control age and gender effects on the results of measurement invariance analysis, which, thus far, has never been done in studies on the BAS-2.

### CRedit authorship contribution statement

**Magdalena Razmus:** Conceptualization, Methodology, Investigation, Writing - original draft, Writing - review & editing, Supervision. **Wiktor Razmus:** Conceptualization, Methodology, Investigation, Formal analysis, Writing - original draft, Writing - review & editing. **Tracy L. Tylka:** Methodology, Investigation, Writing - original draft, Writing - review & editing. **Marija Jović:** Methodology, Investigation. **Marko Jović:** Methodology, Investigation. **Hikari Namatame:** Methodology, Investigation.

### Declaration of Competing Interest

None.

### Acknowledgements

We would like to thank Mohammad Atari for sharing his Iranian BAS-2 data.

### References

Alcaraz-Ibáñez, M., Cren Chiminazzo, J. G., Sicilia, Á., & Teixeira Fernandes, P. (2017). Examining the psychometric properties of the Body Appreciation

- Scale-2 in Brazilian adolescents. *Psychology Society & Education*, 9(3), 505–515. <http://dx.doi.org/10.25115/psye.v9i3.1101>
- Alleva, J. M., Martijn, C., Veldhuis, J., & Tylka, T. L. (2016). A Dutch translation and validation of the Body Appreciation Scale-2: An investigation with female university students in the Netherlands. *Body Image*, 19, 44–48. <http://dx.doi.org/10.1016/j.bodyim.2016.08.008>
- Argyrides, M. (2019). The Body Appreciation Scale-2: Translation and validation in the Greek language. *The European Journal of Counselling Psychology*, 8(1), 108–117. <http://dx.doi.org/10.5964/ejcop.v8i1.179>
- Atari, M. (2016). Factor structure and psychometric properties of the Body Appreciation Scale-2 in Iran. *Body Image*, 18, 1–4. <http://dx.doi.org/10.1016/j.bodyim.2016.04.006>
- Atari, M., Akbari-Zardkhaneh, S., Mohammadi, L., & Soufiabadi, M. (2015). The factor structure and psychometric properties of the Persian version of Body Appreciation Scale. *American Journal of Applied Psychology*, 3(3), 62–66. <http://dx.doi.org/10.12691/ajap-3-3-3>
- Avalos, L., Tylka, T. L., & Wood-Barcalow, N. (2005). The Body Appreciation Scale: Development and psychometric evaluation. *Body Image*, 2(3), 285–297. <http://dx.doi.org/10.1016/j.bodyim.2005.06.002>
- Baceviciene, M., & Jankauskiene, R. (2020). Associations between body appreciation and disordered eating in a large sample of adolescents. *Nutrients*, 12(3) <http://dx.doi.org/10.3390/nu12030752>
- Brown, T. A. (2015). *Confirmatory factor analysis for applied research* (2nd ed.). NY: Guilford Press.
- Buote, V. M., Wilson, A. E., Strahan, E. J., Gazzola, S. B., & Papps, F. (2011). Setting the bar: Divergent sociocultural norms for women's and men's ideal appearance in real-world contexts. *Body Image*, 8, 322–334. <http://dx.doi.org/10.1016/j.bodyim.2011.06.002>
- Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling*, 14(3), 464–504. <http://dx.doi.org/10.1080/10705510701301834>
- Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling*, 9(2), 233–255. [http://dx.doi.org/10.1207/S15328007SEM0902\\_5](http://dx.doi.org/10.1207/S15328007SEM0902_5)
- Cox, A. E., Ullrich-French, S., Tylka, T. L., & McMahon, A. K. (2019). The roles of self-compassion, body surveillance, and body appreciation in predicting intrinsic motivation for physical activity: Cross-sectional associations, and prospective changes within a yoga context. *Body Image*, 29, 110–117. <http://dx.doi.org/10.1016/j.bodyim.2019.03.002>
- Góngora, V. C., Cruz Licea, V., Mebarak Chams, M. R., & Thornborrow, T. (2020). Assessing the measurement invariance of a Latin-American Spanish translation of the Body Appreciation Scale-2 in Mexican, Argentinean, and Colombian adolescents. *Body Image*, 32, 180–189. <http://dx.doi.org/10.1016/j.bodyim.2020.01.004>
- Gupta, V., Hanges, P. J., & Dorfman, P. (2002). Cultural clusters: Methodology and findings. *Journal of World Business*, 37(1), 11–15. [http://dx.doi.org/10.1016/S1090-9516\(01\)00070-0](http://dx.doi.org/10.1016/S1090-9516(01)00070-0)
- He, J., Sun, S., Zickgraf, H. F., Lin, Z., & Fan, X. (2020). Meta-analysis of gender differences in body appreciation. *Body Image*, 33, 90–100. <http://dx.doi.org/10.1016/j.bodyim.2020.02.011>
- Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *The Behavioral and Brain Sciences*, 33(2–3), 61–83. <http://dx.doi.org/10.1017/S0140525X0999152X>
- Human Development Report. (2019). *Beyond income, beyond averages, beyond today: Inequalities in human development in the 21st century*. <http://hdr.undp.org/sites/default/files/hdr2019.pdf>
- Jáuregui Lobera, I., & Bolaños Ríos, P. (2011). Spanish version of the Body Appreciation Scale (BAS) for adolescents. *The Spanish Journal of Psychology*, 14(1), 411–420.
- Jovic, M., Sforza, M., Jovanovic, M., & Jovic, M. (2017). The Acceptance of Cosmetic Surgery Scale: Confirmatory factor analyses and validation among Serbian adults. *Current Psychology*, 36(4), 707–718. <http://dx.doi.org/10.1007/s12144-016-9458-7>
- Kertechian, S., & Swami, V. (2017). An examination of the factor structure and sex invariance of a French translation of the Body Appreciation Scale-2 in university students. *Body Image*, 21, 26–29. <http://dx.doi.org/10.1016/j.bodyim.2017.02.005>
- Lai, M. H. C., Richardson, G. B., & Mak, H. W. (2019). Quantifying the impact of partial measurement invariance in diagnostic research: An application to addiction research. *Addictive Behaviors*, 94, 50–56. <http://dx.doi.org/10.1016/j.addbeh.2018.11.029>
- Lemoine, J. E., Konradsen, H., Lunde Jensen, A., Roland-Lévy, C., Ny, P., & Khalaf, A., et al. (2018). Factor structure and psychometric properties of the Body Appreciation Scale-2 among adolescents and young adults in Danish, Portuguese, and Swedish. *Body Image*, 26, 1–9. <http://dx.doi.org/10.1016/j.bodyim.2018.04.004>
- Lúcio, P. S., Salum, G., Swardfager, W., Mari, J., de, J., & Pan, P. M., et al. (2017). Testing measurement invariance across groups of children with and without Attention-Deficit/Hyperactivity Disorder: Applications for word recognition and spelling tasks. *Frontiers in Psychology*, 8, 1891. <http://dx.doi.org/10.3389/fpsyg.2017.01891>
- Marta-Simões, J., Mendes, A., Oliveira, S., Trindade, I. A., & Ferreira, C. (2016). *Validation of the Body Appreciation Scale-2 for the Portuguese population [Poster presented at the 3rd IPLEiria's International Health Congress]*.

- Meredith, W. (1993). Measurement invariance, factor analysis and factorial invariance. *Psychometrika*, 58(4), 525–543. <http://dx.doi.org/10.1007/BF02294825>
- Muthén, L. K., & Muthén, B. O. (2012). *Mplus user's guide* (7th ed.). Muthén & Muthén.
- Namatame, H., Uno, K., & Sawamiya, Y. (2017). Development of Japanese version of the Body Appreciation Scale-2. *The Japanese Journal of Psychology*, 88, 358–365. <http://dx.doi.org/10.4992/jjpsy.88.16216>
- Nedeljković, M., Vukonjanski, J., Nikolic, M., Hadzic, O., & Sljukic, M. (2018). A comparative analysis of Serbian national culture and national cultures of some European countries by GLOBE project approach. *Journal of the Geographical Institute "Jovan Cvijic" SASA*, 68, 363–382. <http://dx.doi.org/10.2298/IJGI180315002N>
- Ng, S.-K., Barron, D., & Swami, V. (2015). Factor structure and psychometric properties of the Body Appreciation Scale among adults in Hong Kong. *Body Image*, 13, 1–8. <http://dx.doi.org/10.1016/j.bodyim.2014.10.009>
- Pálmarsdóttir, P. G., & Karlsdóttir, A. E. (2016). *Psychometric properties of the Icelandic translation of the Body Appreciation Scale-2 Dissertation* [University of Iceland]. <http://skemman.is/en/item/view/1946/24746>
- Putnick, D. L., & Bornstein, M. H. (2016). Measurement invariance conventions and reporting: The state of the art and future directions for psychological research. *Developmental Review*, 41, 71–90.
- Razmus, M., & Razmus, W. (2017). Evaluating the psychometric properties of the Polish version of the Body Appreciation Scale-2. *Body Image*, 23, 45–49. <http://dx.doi.org/10.1016/j.bodyim.2017.07.004>
- Schweizer, K. (2010). Some guidelines concerning the modeling of traits and abilities in test construction. *European Journal of Psychological Assessment*, 26(1), 1–2. <http://dx.doi.org/10.1027/1015-5759/a000001>
- Swami, V., & Chamorro-Premuzic, T. (2008). Factor structure of the Body Appreciation Scale among Malaysian women. *Body Image*, 5(4), 409–413. <http://dx.doi.org/10.1016/j.bodyim.2008.04.005>
- Swami, V., & Jaafar, J. L. (2012). Factor structure of the Body Appreciation Scale among Indonesian women and men: Further evidence of a two-factor solution in a non-Western population. *Body Image*, 9(4), 539–542. <http://dx.doi.org/10.1016/j.bodyim.2012.06.002>
- Swami, V., & Ng, S.-K. (2015). Factor structure and psychometric properties of the Body Appreciation Scale-2 in university students in Hong Kong. *Body Image*, 15, 68–71. <http://dx.doi.org/10.1016/j.bodyim.2015.06.004>
- Swami, V., Campana, A. N. N. B., Ferreira, L., Barrett, S., Harris, A. S., & da C. G. C. F. Tavares, M. (2011). The Acceptance of Cosmetic Surgery Scale: Initial examination of its factor structure and correlates among Brazilian adults. *Body Image*, 8(2), 179–185. <http://dx.doi.org/10.1016/j.bodyim.2011.01.001>
- Swami, V., Mohd. Khatib, N. A., Toh, E., Zahari, H. S., Todd, J., & Barron, D. (2019). Factor structure and psychometric properties of a Bahasa Malaysia (Malay) translation of the Body Appreciation Scale-2 (BAS-2). *Body Image*, 28, 66–75. <http://dx.doi.org/10.1016/j.bodyim.2018.12.006>
- Swami, V., Ng, S.-K., & Barron, D. (2016). Translation and psychometric evaluation of a Standard Chinese version of the Body Appreciation Scale-2. *Body Image*, 18, 23–26. <http://dx.doi.org/10.1016/j.bodyim.2016.04.005>
- Swami, V., Tran, U. S., Barron, D., Afhami, R., Aimé, A., & Almenara, C. A., et al. (2020). The Breast Size Satisfaction Survey (BSSS): Breast size dissatisfaction and its antecedents and outcomes in women from 40 nations. *Body Image*, 32, 199–217. <http://dx.doi.org/10.1016/j.bodyim.2020.01.006>
- Swami, V., García, A. A., & Barron, D. (2017). Factor structure and psychometric properties of a Spanish translation of the Body Appreciation Scale-2 (BAS-2). *Body Image*, 22, 13–17. <http://dx.doi.org/10.1016/j.bodyim.2017.05.002>
- Swami, V., Hadji-Michael, M., & Furnham, A. (2008). Personality and individual difference correlates of positive body image. *Body Image*, 5(3), 322–325. <http://dx.doi.org/10.1016/j.bodyim.2008.03.007>
- Swami, V., Hwang, C.-S., & Jung, J. (2012). Factor structure and correlates of the acceptance of cosmetic surgery scale among South Korean university students. *Aesthetic Surgery Journal*, 32(2), 220–229. <http://dx.doi.org/10.1177/1090820X11431577>
- Swami, V., Mada, R., & Tovée, M. J. (2012). Weight discrepancy and body appreciation of Zimbabwean women in Zimbabwe and Britain. *Body Image*, 9(4), 559–562. <http://dx.doi.org/10.1016/j.bodyim.2012.05.006>
- Swami, V., Stieger, S., Haubner, T., & Voracek, M. (2008). German translation and psychometric evaluation of the Body Appreciation Scale. *Body Image*, 5(1), 122–127. <http://dx.doi.org/10.1016/j.bodyim.2007.10.002>
- Swami, V., Tudorel, O., Goian, C., Barron, D., & Vintila, M. (2017). Factor structure and psychometric properties of a Romanian translation of the Body Appreciation Scale-2. *Body Image*, 23, 61–68. <http://dx.doi.org/10.1016/j.bodyim.2017.08.001>
- Tiggemann, M., & McCourt, A. (2013). Body appreciation in adult women: Relationships with age and body satisfaction. *Body Image*, 10(4), 624–627. <http://dx.doi.org/10.1016/j.bodyim.2013.07.003>
- Todd, J., & Swami, V. (2020). Assessing the measurement invariance of two positive body image instruments in adults from Malaysia and the United Kingdom. *Body Image*, 34, 112–116. <http://dx.doi.org/10.1016/j.bodyim.2020.05.009>
- Tylka, T. L., & Wood-Barcalow, N. L. (2015a). The Body Appreciation Scale-2: Item refinement and psychometric evaluation. *Body Image*, 12, 53–67. <http://dx.doi.org/10.1016/j.bodyim.2014.09.006>
- Tylka, T. L., & Wood-Barcalow, N. L. (2015b). What is and what is not positive body image? Conceptual foundations and construct definition. *Body Image*, 14, 118–129. <http://dx.doi.org/10.1016/j.bodyim.2015.04.001>
- Tylka, T. L. (2019). Body appreciation. In T. L. Tylka & N. Piran (Eds.), *Handbook of positive body image and embodiment: Constructs, protective factors, and interventions* (pp. 22–32). Oxford University Press.
- Vally, Z., D'Souza, C. G., Habeeb, H., & Bensumaideia, B. M. (2019). The factor structure and psychometric properties of an Arabic-translated version of the Body Appreciation Scale-2. *Perspectives in Psychiatric Care*, 55, 373–377. <http://dx.doi.org/10.1111/ppc.12312>
- Wang, S., Chen, C.-C., Dai, C.-L., & Richardson, G. B. (2018). A call for, and beginner's guide to, measurement invariance testing in evolutionary psychology. *Evolutionary Psychological Science*, 4(2), 166–178. <http://dx.doi.org/10.1007/s40806-017-0125-5>
- World Data. (2019). *Average income around the world*. <https://www.worlddata.info/average-income.php>
- World Health Organization. (2014). *Global status report on noncommunicable diseases 2014*. <https://www.who.int/nmh/publications/nccd-status-report-2014/en/>