This is a preprint version of doi: 10.1080/21642850.2020.1797507. You may access the full article at: https://www.tandfonline.com/doi/full/10.1080/21642850.2020.1797507

Gender-Based Food Intake Stereotype Scale (GBFISS) for adolescents: Development and psychometric evaluation

Abstract

Objective. The study aimed to develop and test the validity and reliability of a gender-based food intake stereotype scale (GBFISS) to further the understanding of gender stereotype influences on food intake. **Design.** Two cross-sectional studies were conducted among adolescents. In the first one (*n*= 611), exploratory and confirmatory factor analyses were performed on subsamples to identify and cross-validate the scale's structure. Evidence of concurrent validity (correlation with sexism) was also examined. In the second study (*n*= 813), confirmatory factor analysis was conducted to confirm the scale's dimensionality on a different sample. Further evidence of construct validity (correlations with food intake and social desirability) was examined. Invariance was tested for different features as well. **Main outcome.** The Gender-Based Food Intake Stereotype Scale. **Results.** Factor analyses on the first and second studies helped identify and confirm the GBFISS as a three-dimensional scale. The studies also provided evidence of construct validity. Support for invariance by gender and age was found, and reliability was acceptable. **Conclusion.** The evidence suggests that the GBFISS is valid and reliable. Further research is recommended. The contribution of gender stereotypes, as measured by the GBFISS, to well-established health behavior models should be examined.

Keywords: Sexism, gender-based stereotypes, food intake, scale development

27

Introduction

Gender stereotypes refer to the set of social roles and behavioral norms and practices that are considered socially appropriate for men and women, so that, based on them, a person is deemed as masculine or feminine in the context of a specific culture and historical period (De Lemus et al., 2013). Across different cultures, masculinity is constructed in opposition to femininity, or to what it means to be feminine (Ellemers, 2018).

An implication of stereotyping two groups as opposites is that any movement away from the stereotype of one group is, by definition, a movement toward the other group (Lips, 2020). For example, a man who is perceived as acting less rationally than the male stereotype is seen not only as less masculine but also as more feminine. Conversely, a woman who is perceived as acting less emotionally than the female stereotype is viewed not only as less feminine but also as more masculine (Lips, 2020).

Health behaviors are part of broader social practices through which gender identities are continuously (re) constructed. Positive health beliefs or behaviors are also socially constructed as forms of idealized femininity (Cornwall, 2000; Lyons, 2009). As such, they are potentially feminizing influences that men must oppose using diverse strategies and mechanisms, depending on what other resources are accessible or are being utilized in the construction of masculinity. It has been demonstrated that the resources available for constructing masculinity are mostly unhealthy (e.g., consuming excessive amounts of alcohol (and drugs), not seeking professional help, being violent and aggressive, engaging in risky sexual and driving behaviors, and adopting an unhealthy diet) (Ellemers, 2018; Lyons, 2009). Men and boys often use these resources and reject healthy beliefs and

behaviors to demonstrate and achieve what is considered as manhood. A man's success in adopting (socially feminized) health-promoting behaviors, as well as his failure to engage in (socially masculinized) physically risky behaviors, can undermine his ranking among men and relegate him to a subordinated status (Ellemers, 2018). Based on cultural norms, men and boys tend to construct masculinity in opposition to the health beliefs and behaviors of women and less masculine (i.e., "feminized") men and boys. In the same way, women and girls tend to construct femininity in opposition to behaviors related to masculinity.

Several authors (Clément-Guillotin et al., 2011; Hannon et al., 2009; Hardin & Greer, 2009; Plaza et al., 2017) have shown that the practice of some physical activities is usually incompatible with the common constructions of feminine behavior. Sports are gender-based activities, with value and power associated with masculine traits (Birrell, 2013).

Gender differences in terms of food preferences have also been reported and might be partially explained by gender stereotypes (Al-Sobayel, Al-Hazzaa, Abahussain, Qahwaji, & Musaiger, 2015; Caine-Bish & Scheule, 2009). Consumption of meat and high-energy-dense foods (e.g., fast food, sugar-sweetened beverages) has been identified as a marker of masculinity. In contrast, consuming vegetables, fruits, and other healthy foods is identified as a marker of femininity. Women that conform to this conception of femininity reduce the amount of food they consume and eat slowly compared to men (Arganini et al., 2012; Carey et al., 2017; Cavazza et al., 2015a; Monge-Rojas et al., 2015; Vartanian et al., 2007; Young et al., 2009).

A body of evidence suggests that healthy dietary habits established during adolescence persist into adulthood (Cruz et al., 2018; Movassagh et al., 2017).

Consequently, adolescence has been suggested as the best time to introduce dietary modifications that seek to enhance health-conscious dietary habits (Cruz et al., 2018; Mikkilä et al., 2005; Schneider et al., 2016). However, since adolescents might be quite sensitive to social norms (Lombardi et al., 2019), it is particularly valuable to develop a better scientific understanding of gender-based stereotypes and their role in the establishment of unhealthy eating habits during this period of life. Several studies (Herman et al., 2019; Igenoza, 2017; Le, 2019; Timeo & Suitner, 2018) have shown that eating-related traditional femininity victimize girls into stereotypical body shapes and harmful weight-control behaviors (like dietary restraint). On the other hand, the high-energy-dense foods related to masculinity make adolescent boys more susceptible to developing a deleterious lipid profile and overweight/obesity in the short term. Furthermore, adolescents with unhealthy eating habits have a higher risk of developing cardiometabolic syndrome and its related complications in adulthood (Craigie et al., 2011; Cruz et al., 2018; Movassagh et al., 2017).

Methods used to study gender-based food intake stereotypes include qualitative interviews and focus groups (Carey, Saules, & Carr, 2017; Monge-Rojas et al., 2015), as well as self-reports (including correlational and experimental/quasi-experimental designs) (Cavazza et al., 2015b, 2015a; Kimura et al., 2009). However, to our knowledge, no scale has been developed and validated to measure such gender-based stereotypes.

Despite their likely contribution to the understanding of some health behaviors – especially those where gender differences are frequently reported– gender stereotypes are not explicitly included in major health behavior models (e. g., Ajzen, 1991; Prochaska & DiClemente, 1982; Schwarzer, 2008). Arguably, some health behavior models address

social norms (e. g., Ajzen, 1991), but their focus is not necessarily on gender. The development of a scale for gender-based food intake stereotypes may help examine their role in the mechanisms described by major health behavior models and determine their influence on the adoption of healthy eating habits during adolescence.

An unhealthy diet during adolescence has harmful short- and long-term health consequences. Consequently, identifying the factors that act as barriers to adopting a healthy diet during adolescence provides timely information to public policymakers for the definition of effective strategies aimed at establishing healthy eating habits during this life period.

Gender-based stereotypes, sexism, and food intake

From a theoretical standpoint, the construct of gender-based food intake stereotypes should relate to two kinds of variables: sexism and dietary food intake. *Sexism* has been defined as the endorsement of discriminatory or prejudicial beliefs and feelings based on sex, and it is usually linked to stereotypical conceptions of the sexes and the adoption of a traditional gender-role ideology (Moya & Expósito, 2001). Sexism has also been described as a system of inequality based on gender, which involves beliefs and discriminatory treatments based on the assumed superiority and privileges of men (Brown, 2010; Pistella et al., 2018).

Currently, psychologists identify two primary types of sexist ideologies: hostile and benevolent (Glick & Fiske, 1996). Hostile sexism is a derogatory view of women based on resentment, distrust, and the perception that women are seeking control over men. Benevolent sexism is a subjectively positive view of women as "pure creatures," who need to be protected and adored based on the perception of women as weak and best relegated to traditional gender roles. The endorsement of sexist views has been related to homophobic attitudes (Pistella et al., 2018). The belief that men are superior and that traditional gender

roles should hold may also be expressed as hostile beliefs towards individuals not fitting these roles, such as homosexuals.

Ambivalent sexism has been related to different types of masculinity and femininity (Glick et al., 2015). Masculinity is viewed as a social location, a set of practices and characteristics understood as "masculine" and having effects on bodily experience, individuals, relationships, and social structures (Schippers, 2007). Thus, instead of "possessing or having masculinity, individuals move through and produce masculinity by engaging in masculine practices" (Schippers, 2007). One salient type of masculinity found in gender studies literature is known as "hegemonic masculinity" (Connell, 1995; Connell & Messerschmidt, 2005; Messerschmidt, 2019). Connell (1995) defines it as a specific form of masculinity in a given historical and society-wide social setting that legitimizes unequal gender relations between men and women, between masculinity and femininity, and among masculinities. Hegemonic masculinity influences men's identities and behaviors (e.g., being strong, aggressive, tough, independent, courageous, invulnerable). Some masculine practices and characteristics are hegemonic, and others are not (e.g., supporting household activities, looking after body and personal appearance, having refined manners, being emotional) (Messerschmidt et al., 2018). Furthermore, different masculinities are continuously being renegotiated through different practices, arise out of different social contexts, and are not necessarily linked to different groups of men (Cornwall & White, 2000).

Hegemonic masculinity is not a trait-focused or fixed character concept: Connell (1995) emphasized its relational nature, which legitimates the superordination of some men over women and men with alternative forms of masculinity (Messerschmidt, 2019). These

masculinity subtypes are considered subordinate masculinities: those constructed as deviant to hegemonic masculinity.

The concept of hegemonic masculinity was formulated in tandem with emphasized femininity, a normative form of femininity that is practiced in a complementary, compliant, and accommodating subordinate relationship with hegemonic masculinity (Connell & Messerschmidt, 2005).

Literature from different theoretical frameworks suggests various mechanisms by which sexist ideologies might indirectly affect a wide range of behaviors (including those that are health-related), through gender stereotypes. For instance, the Expectancy-Value Model proffers that belief systems, cultural stereotypes, and social norms might determine behaviors through two core variables: success expectancies, that is, the perceived probability of success in a particular task, and subjective task value, which refers to the extent to which a task provides intrinsic interest and is perceived as useful and relevant by the individual (Eccles, 2011).

Expectancies and values are shaped over time by individual and contextual factors. These include personal and family features (e.g., gender, culture, SES), previous experiences of success and failure, individual self-concept, and the influence of different socializing agents (e.g., parents, teachers, peers, and schools).

Sexism may also indirectly affect various women's behavior through the internalization of hostile and benevolent sexist beliefs that may lead women to perceive substantial differences between genders (Hyde, 2005; Steele & Aronson, 1995), which in turn might affect their self-perception and motivations. In this regard, research has shown that women are more prone than men to support a generalized and diffuse system of

169 inequality after being exposed to benevolent sexism (Dardenne et al., 2007; Jost & Kay,
170 2005). Moreover, a substantial body of evidence states that stereotypes may influence
171 behavior when a member of a stereotyped group is placed in a situation in which his or her
172 behavior could be judged as evidence that the individual possesses stereotypical group
173 deficiencies. (Steele et al., 2004; Steele & Aronson, 1995).

174 Food intake is another variable that can be related to the construct of gender-based
175 stereotypes. Several qualitative studies have shown that the association of femininity and
176 masculinity with specific foods is often correlated with the food's profile (i.e., health value,
177 caloric and fat content), and with good/bad classifications that arise from these profiles.
178 Food intake in girls is usually higher in fruits, vegetables, and sweet foods, and lower in
179 fatty foods than in boys, suggesting that the girls' intake is healthier (Arganini et al., 2012;
180 Carey et al., 2017; Cavazza et al., 2015a; Kimura et al., 2009, 2011; Monge-Rojas et al.,
181 2015; Vartanian et al., 2007; Young et al., 2009).

182 Previous qualitative research on the influence of gender-based stereotypes on eating
183 behavior among Costa Rican adolescents (Monge-Rojas et al., 2015) suggests three salient
184 themes or categories of beliefs about food intake: consumption of moderate quantities of
185 nutritious food is related to femininity and boys' homosexuality; consumption of hearty
186 portions of unhealthy foods is associated with masculinity and boys' heterosexuality, and
187 body care among adolescent girls is an element of femininity and body image.

188 Food quantity and eating speed were also related: adolescent participants associated
189 faster eating with heterosexual masculinity, as opposed to femininity and men's
190 homosexuality (Monge-Rojas et al., 2015). This finding was consistent with previous
191 literature (Herman & Polivy, 2010). Although the qualitative findings of Monge-Rojas et

al. (2015) were used as the foundation for scale item generation (see Methods), the gender subtypes conceptualization by Connell (1995) and Messerschmidt (2019) remains in this proposal: we hold that there is a normative hegemonic masculinity from which the subordinate gender subtypes (feminine and masculine) are distinguished.

As suggested by the needs highlighted in this literature, we set out to develop a Gender-Based Food Intake Stereotype Scale (GBFISS) and to examine its psychometric properties (reliability and construct validity). We expect this new scale to be an instrument for further study of the influence of gender-based food intake stereotypes among adolescents.

Materials and methods

Participants and procedures

Two cross-sectional studies were conducted in sequence to assist in the development and assessment of the psychometric properties of a new scale about gender-based food intake stereotypes.

In the first study, we examined the theoretically expected convergence between a sexism scale and the GBFISS for construct validation and explored and cross-validated the scale's structure.

In the second study, we examined further evidence about the scale's dimensionality and, more importantly, we assessed a second theoretically grounded hypothesis as additional evidence of construct validity. The GBFISS was expected to be associated with food intake measures, and evidence of divergent validity was expected for the relationship between the GBFISS and social desirability scores. We also assessed the scale's fit to different subgroups (gender, age, and area of residence) and tested for invariance.

These studies included convergent and discriminant evidence of validity, in line with recommendations for testing new instruments (Campbell & Fiske, 1959). Further instrument characteristics were analyzed and reported in both studies (see Data Analysis).

The first study took place in 2016, with 611 adolescent participants aged 12 to 17 years (50.7% boys; mean age: 15.17 ± 1.6 years). The second study followed in 2018, with 813 adolescent subjects aged 12 to 17 years (36,5% boys; mean age: 15.03 ± 1.7 years).

Given that most Costa Rican adolescents (80%) are enrolled in school (Programa Estado de la Nación, 2019), these studies enlisted seventh to eleventh graders from rural and urban schools in the province of San José. San José is the Costa Rican province with the highest adolescent concentration (30%) in the country (UCR, 2013).

In determining the sample size of each study, we assumed a sampling error for a proportion of the population and applied a finite population correction. (Ryan, 2013). The sample was selected in three stages: 1) The schools were chosen using a proportional-size probability method (Skinner, 2014). The school sample from the first study (n=12) was different from the second study (n=16); 2) At each school, ten classes (2 from each grade level) were selected using simple random sampling, and 3) Participants were chosen randomly among those students who returned signed informed consent form (ICF) and informed assent form (IAF). Over 95% of adolescents returned the ICF signed by some of their parents, and 100% provided the IAF.

As part of the ethical procedures to protect human beings, the research team first contacted the adolescents at their schools to invite them to take part in the study. The IAF was explained to and read by interested students. Those in agreement with the IAF printed their names on it before an impartial witness who was not part of the research team. The

ICF was given to the students to take home and obtain parental permission to participate in the study. In compliance with the Costa Rican Biomedical Research Law (Asamblea Legislativa, 2014), parents who signed the ICF had to provide a copy of their ID to verify the stamped signatures. Parental signature was mandatory since the study participants were minors (under 18 years of age). Any adolescents that did not provide a signed ICF were excluded from the study. No other criteria were applied for selecting study participants.

At each school, participating students were gathered in a dedicated classroom during regular school hours. They were instructed on how to complete their sociodemographic information (age, gender, area of residence), fill the GBFISS, and answer a 22-item sexism scale. A researcher was available throughout to answer any questions. Afterward, a thorough explanation of how to collect food intake data was provided (see Measures). On average, the adolescents took 50 minutes to answer the scales. A bioethics committee, accredited by the Costa Rican Ministry of Health, approved the study, and all guidelines for human subject research were followed.

Measures

Sexism was measured using the *Ambivalent Sexism Inventory* (ASI; Glick & Fiske, 1996), adap ted to Latin American populations (Cárdenas et al., 2010). This is a paper and pencil 22-item instrument made up of two subscales: Hostile Sexism (HS), and Benevolent Sexism (BS). Examples of HS items are “Women seek to gain power by getting control over men” and “Women exaggerate problems they have at work”. Examples of BS items are “Many women have a quality of purity that few men possess,” and “Women should be cherished and protected by men.” Items are rated on a 5-point Likert scale.

Glick and Fiske (1996) reported Cronbach's alpha coefficients for the overall scale ranging from .80 to .90. For the HS subscale, alphas range from .80 to .90, while the BS subscale's alphas are lower, ranging from .70 to .85. Their validity studies yielded significant correlations between the ASI, especially the HS subscale, with other measures of sexism, racism, and gender bias. Further reports on psychometric properties as well as information on their application to different age and cultural groups have been provided (Cárdenas et al., 2010; Etchezahar & Ungaretti, 2014; Glick et al., 2002; North & Fiske, 2014). Regarding our data (first study), the overall scale reliability was α = .81, while the HS and BS subscale alphas were .84 and .70, respectively.

Social desirability was measured using the short form of the *Social Desirability Scale* developed by Crowne and Marlowe (1960) (MCSDS), with 13 true/false items. An example item is "I am always courteous, even to people who are disagreeable." The authors of the MCSDS considered it to have a single construct, namely, "the need for approval," defined as the extent to which an individual seeks the approval of others and tries to avoid their disapproval (Crowne & Marlowe, 1960; Leite & Beretvas, 2005). The rationale behind the items on the MCSDS is that an average individual would not always behave in a socially desirable manner. Consequently, a person with a higher need for approval would tend to present more socially desirable responses than the average (Leite & Beretvas, 2005). The use of the MCSDS has been extensive since its development (Beretvas et al., 2002), including its adaptation and use in different languages, contexts, and cultural backgrounds (e. g., Gutierrez, Sanz, Espinosa, Gesteira, & Paz Garcia-Vera, 2016; Kurz, Drescher, Chin, & Johnson, 2016; Perez, Labiano, & Brusasca, 2010;). This instrument has already been adapted and applied in Costa Rica (Smith-Castro, 2014). Further details and

discussions on the MCSDS structure, validity, and reliability have been provided elsewhere (e. g., Leite & Beretvas, 2005; Ventimiglia & MacDonald, 2012; Vésteinsdóttir, Reips, Joinson, & Thorsdottir, 2015). The reliability of our data (second study), as measured by the MCSDS, was α = .65.

Dietary food intake data were collected using 3-day food records (Ortega et al., 2015). Six trained nutritionists instructed the participants on how to complete accurate written food records for three consecutive days. Participants were asked to record detailed descriptions of all the foods and drinks consumed during the entire day, including food brand names when appropriate, methods of preparation, and recipes whenever possible. The participants also learned how to estimate portion sizes using a manual developed for Costa Rica (Chinnock, 2007). The manual includes photographs and diagrams of commonly consumed foods and preparations and includes 3 to 6 different portion sizes. The adolescents reported portion sizes using kitchen measurement tools (e.g., tablespoons, teaspoons, cups, glasses).

Current literature indicates that high-energy-dense foods are closely related to masculinity and dissociated to femininity (Arganini, Saba, Comitato, Virgili, & Turrini, 2012; Carey, Saules, & Carr, 2017; Cavazza, Guidetti, & Butera, 2015a; Monge-Rojas et al., 2015; Young, Mizzau, Mai, Sirisegaram, & Wilson, 2009). Hence, the consumption of fast food and sugary beverages was included as an external criterion. Skewness and kurtosis ranges for the consumption of beverages with added sugar and fast food were within the levels suggested by Kline (2011). Thus, transformation was not needed.

The information extracted from the food records was entered into a software application designed to assess the dietary composition of various foods in Costa Rica (Chinnock, 2010). Quantities were expressed in grams per day.

Data analysis

Item generation, exploratory factor analysis (EFA) and confirmatory factor analysis (CFA)

Based on the results of previous qualitative research by Monge-Rojas et al. (2015), themes about gender-based stereotypes among Costa Rican adolescents were identified. These themes were used to generate fifty items related to stereotypes in three gender subtypes: normative hegemonic masculinity, normative subordinate femininity, and non-normative subordinate masculinity. The items were applied to a sample of 611 students as part of a pilot study (Study 1). Dimensionality was first explored in a randomly selected subsample of 33% (N = 203). To improve interpretation, only items loading clearly in one dimension were selected (in exploratory factor analysis, the difference between loadings must be at least = .20). The final scale consisted of 21 items, with response options following a 5-point Likert format ranging from "strongly disagree" (1) to "strongly agree" (5). The original set of fifty items is provided as supplemental material (Appendix 1) as well as the final version of the scale (Appendix 2).

Exploratory factor analysis (EFA) was performed on the subsample data. Factors with eigenvalues > 1 were retained. For each of the dimensions identified, a McDonald's omega (ω) reliability analysis was conducted. Reports indicate that Cronbach's alpha is a statistically inappropriate estimation of the internal consistencies of scale items, and omega has been suggested as a better option (Crutzen & Peters, 2017; Gjalt Jorn Peters, 2014; Ventura-León & Caycho-Rodríguez, 2017). However, since many studies still include the alpha levels of scales, Cronbach's alpha (α) was also calculated and reported as additional information.

The factor solution found in the EFA was cross-validated on the complementary subsample (67%, N = 408) using a confirmatory factor analysis (CFA; estimation method: Maximum Likelihood). Reliabilities (McDonald's omega and Cronbach's alpha) and convergent validity (Pearson's correlation with sexism subscales) were examined in this subsample as well.

An additional CFA was performed on Study 2 using correlations (Pearson's *r*) with dietary food intake and social desirability as external criteria (for concurrent and discriminant validity). The aim was to replicate the results of the first study on a different sample of adolescents (N=813) and improve the robustness of the construct's validity (as suggested by Campbell and Fiske (1959), new scales require evidence of both concurrent and discriminant validity).

Criteria by Hu and Bentler (1999) and Cangur and Ercan (2015) were applied to examine fit in the CFA models. Both $\chi 2$ and $\chi 2/df$ were reported. For $\chi 2/df$, values close to 3.0 were considered acceptable, and lower values were taken as indicators of a better fit (Cangur & Ercan, 2015). The Comparative Fit Index (CFI), a measure of incremental fit, was also reported. In this index, values of .90 have been traditionally used as a cutoff, although more recently, values close to .95 are preferred (Cangur & Ercan, 2015; Hooper et al., 2008; Hu & Bentler, 1999). A CFI of .90 or higher was deemed acceptable, and a CFI of .95, satisfactory. Finally, a measure of absolute fit (Root Mean Square Error of Approximation (RMSEA)) was reported. Generally, an RMSEA value of .06 or lower is considered indicative of a good fit (Hooper et al., 2008; Hu & Bentler, 1999). Cangur and Ercan (2015) have been more specific with their interpretation of the RMSEA, suggesting

353 that a value of .05 or lower indicates convergence fit, a value between .05 and .08 indicates
354 a close-to-good fit, and a value between .08 and .10 is neither good nor bad.

355 In the second study, with the larger sample, model fit in different subgroups based
356 on gender, age, and area of residence was also examined. Where fit was acceptable,
357 invariance was also examined. There are several invariance levels (Furr, 2017), the weakest
358 of which is configural invariance. If this invariance level is met, it can be concluded that
359 items reflect the same latent constructs across (gender and age) groups. A more robust level
360 is known as strict invariance. If met, it indicates that the pattern of the factor loadings
361 across groups is the same, the exact values of the factor loadings are the same, the item
362 intercepts are the same, and—even further—the items' unique error variances are the same
363 (Furr, 2017). In hierarchical factor models such as the second-order factor model of the
364 proposed scale, additional invariance levels can also be tested (Chen et al., 2005). Table 1
365 shows the invariance models tested in this study in more detail. Each of these models was
366 specified as reported in Table 1. For the model examining invariance at a configural level,
367 no constraints between the men and women subgroups were specified in the hierarchical
368 CFA model. Constraints were added to each of the models so that higher invariance levels
369 assumed more invariance (and constraints) between gender subgroups. The same process
370 was repeated afterwards to test invariance by age groups. A statistical test was used to
371 compare more restrictive models, which assume stronger invariance, with the configural
372 and least restrictive model.

373 Insert Table 1 here

374 Traditionally, once an acceptable fit in the configural model has been found, chi-
375 square difference ($\Delta \chi^2$) is used to check if there is invariance in more restrictive models, as

compared to the configural model. However, the chi-square difference test has been criticized for being dependent on sample size. Other indices, such as the Comparative Fit Index difference test (Δ CFI), have been suggested as an alternative, with differences of < .01 between models required to establish invariance (Cheung & Rensvold, 2002). In this study, we use Δ CFI to examine for invariance.

Statistical analyses were performed using the Statistical Package for Social Sciences (SPSS Inc., version 23.0 for Windows, Chicago, Illinois), the Amos software package (Amos 23.0; SPSS Inc.), and the userfriendlyscience R package (Gjakt Jorn Peters et al., 2018).

Results

1. Study 1

1.1.1. Item generation and Exploratory Factor Analysis in Study 1

Items for each subscale originated from the results of the qualitative study of food-gender stereotypes among Costa Rican adolescents (Monge-Rojas et al., 2015). In the Exploratory Factor Analysis, three factors presented eigenvalues higher than 1. Overall, they explained 45.94% of the variance (first factor, 29.81%; second factor, 9.92%, and third factor, 6.23%). Table 2 shows the primary factor loadings of the rotated solution for each item. With regards to item content, the first factor represents a dimension of non-normative subordinate masculinity (stereotypical beliefs of what is considered typical in homosexual or effeminate boys), the second factor represents a dimension of normative subordinate femininity (stereotypical beliefs of what is considered ideal in heterosexual girls), and the

third factor represents a dimension of normative hegemonic masculinity (stereotypical beliefs of what is considered ideal in heterosexual boys).

Insert Table 2 here

The Pearson's correlations among dimensions were all between small and medium, and significant ($p < .001$). Non-normative subordinate masculinity had a correlation of r = .35 with normative hegemonic masculinity and r = .43 with normative subordinate femininity. The correlation between normative hegemonic masculinity and normative subordinate femininity was $r = .39$.

The overall mean of the GBFISS in this subsample was 2.32 (SD = .64). Individual dimension means were: non-normative subordinate masculinity, 1.61 (SD = .81); normative subordinate femininity, 2.45 (SD = .85), and normative hegemonic masculinity, 3.23 (SD = .95). Appendix 3a (Table 7) provides further information on item means, standard deviations, and inter-correlations.

1.1.2. Reliability and validity on the exploratory subsample of Study 1

In the subsample used for the EFA, reliability results *were:* $\omega = .91$ and $\alpha = .91$ for non-normative subordinate masculinity; $\omega = .81$ and $\alpha = .81$ for normative subordinate femininity, and $\omega = .77$ and $\alpha = .77$ for normative hegemonic masculinity. The overall reliability of the scale was $\omega = .86$ and $\alpha = .88$.

Item-total correlations on all the subscales were between $r = .38$ and .76. Each of the gender stereotype dimensions was positively associated with both benevolent and hostile sexism. Correlations between hostile sexism and gender stereotype dimensions were: $r = .22$ ($p < .01$) for non-normative subordinate masculinity; $r = .35$ ($p < .001$) for

normative hegemonic masculinity, and r = .31 (p < .001) for normative subordinate femininity. Correlations between benevolent sexism and gender stereotype dimensions were: r = .30 (p < .001) for non-normative subordinate masculinity; r = .54 (p < .001) for normative hegemonic masculinity, and r = .38 (p < .001) for normative subordinate femininity.

1.2.1. Confirmatory Factor Analysis in Study 1

The scale structure was cross-validated with the remaining 66.7% of the sample (N = 408) using a CFA, where "gender stereotype" was specified as a second-order factor of the three first-order dimensions of non-normative subordinate masculinity, normative subordinate femininity, and normative hegemonic masculinity. Figure 1 presents the results of this analysis in terms of loadings and fit. The statistical significance of factor loadings provided evidence of convergent validity. In a previous CFA model using correlated first-order factors only, correlations were all between β = .39 and β = .42, indicating sufficient discriminant validity.

Insert Figure 1 here

The absolute fit of the model was considered satisfactory, or close to good, per Cangur and Ercan's terminology (2015). Incremental fit (Comparative Fit Index: CFI) was acceptable.

The GBFISS's mean was 2.33 (SD = .63), while the dimension means were: M = 2.51 (SD = .88), for normative subordinate femininity; M = 3.25 (SD = .88) for normative hegemonic masculinity, and M = 1.56 (SD = .77) for non-normative subordinate

masculinity. Appendix 3b (Table 8) provides details on item means, standard deviations and item correlations.

1. 2. 2. Reliability and concurrent validity of the confirmatory subsample in Study 1

Reliabilities for each dimension were: ω = .89 and α = .89 for non-normative subordinate masculinity; ω = .84 and α = .84 for normative subordinate femininity, and ω = .71 and α = .70 for normative hegemonic masculinity. The overall reliability of the scale was ω = .85 and α = .87. The associations between benevolent sexism and gender stereotype dimensions were r = .20 for non-normative subordinate masculinity (p < .01); r = .38 for normative subordinate femininity (p < .001), and r = .48 with normative hegemonic masculinity (p < .001). The associations between hostile sexism and gender stereotype dimensions were r = .24 for non-normative subordinate masculinity (p < .001); r = .37 for normative hegemonic masculinity (p < .001), and r = .36 for normative subordinate femininity (p < .001).

2.Study 2

2.1. Confirmatory Factor Analysis in Study 2

The CFA analysis was replicated in a larger sample using gender stereotypes as a second-order factor, and the dimensions of non-normative subordinate masculinity, normative subordinate femininity, and normative hegemonic masculinity as first-order factors. Figure 2 shows the results in terms of loadings and fit. The statistical significance of factor loadings provided evidence of convergent validity. In a previous CFA model using correlated first-order factors only, correlations were all between β = .28 and β = .44, indicating sufficient discriminant validity

Insert Figure 2 here

The absolute fit of this model was good (Cangur & Ercan, 2015). Even the upper level of the RMSEA's confidence intervals was below the cutoff value provided by Hu & Bentler (1999). Incremental fit was acceptable.

The GBFISS's mean was 2.14 (SD = .55), while the dimension means were: M = 1.28 (SD = .52), for non-normative subordinate masculinity; M = 2.26 (SD = .83), for normative subordinate femininity, and M = 3.32 (SD = .89) for normative hegemonic masculinity. Appendix 3 provides further information on item means, standard deviations, and item correlations.

Insert Table 3 here

Model fit for specific subgroups (gender, age, and residence area) was examined (see Table 3). The model was found to fit the data well for boys and girls, for younger (< 15 years) and older participants (> 15 years), and for participants living in rural areas. However, fit was not acceptable for participants from urban areas. Incremental fit in particular was below the recommended level (CFI < .90). Given these results, we further examined invariance by gender and age, but not by area of residence.

Table 4 presents a summary of invariance test results by gender and age. In both categories, the configural (not constrained) model presented good absolute fit, and incremental fit was acceptable, suggesting that the same set of items reflects the same constructs, independently of gender and age.

When further levels of invariance by gender were examined, the CFI difference test suggested there was invariance at the level of structural covariances (Δ CFI < .01 from the

metric level to the level of structural covariances). Also, there was marginal invariance at the level of structural residuals (Δ CFI = .011). These results indicate that, between boys and girls, the same set of items reflects the same set of constructs; the same first-order constructs represent the same second-order "gender stereotype" construct, which has the same meaning for boys and girls, and even that the structural residuals (disturbances) were almost equivalent.

Insert Table 4 here

Age invariance tests showed comparable results. Between younger and older participants, invariance was confirmed at the metric level (Δ CFI < .01) using the CFI difference test. Invariance was marginal from the scalar level to the level of the second-order (structural) residuals: the difference between the unconstrained model and the constrained models was slightly superior to the suggested maximum CFI difference (Δ CFI = .013). Overall, these results suggest that the same set of items represents the same dimensions in both age groups and that their latent meaning is similar across groups.

2.2.Reliability and validity in Study 2

Reliability was ω = .86 and α = .86 for non-normative subordinate masculinity; ω = .82 and α = .82 for normative subordinate femininity, *and* ω = .73 and α = .73 for normative hegemonic masculinity. Overall reliability was ω = .81 and α = .85.

Evidence of construct validity was provided by the negative association between the overall gender stereotypes scale and the consumption of unhealthy fast food, found only among girls (r = -.19, p < .01) but not among boys (r = .03, p = .70). This result makes sense from a theoretical standpoint because traditional femininity is related to body care and healthy

514 eating (Monge-Rojas et al., 2015). The negative association in girls was also found for the
515 dimensions of normative subordinate femininity (r = -.16, p < .01) and normative
516 hegemonic masculinity (r = -.11, p < .05), but not for non-normative subordinate
517 masculinity (r = -.08, p = .09).

518 Furthermore, the GBFISS general score was also positively associated with the
519 consumption of sugar-sweetened beverages among boys (r = .32, p < .001). This finding
520 agrees with the theoretical expectation and is, therefore, evidence of construct validity. The
521 positive association between gender stereotypes and beverage consumption was also found
522 for some dimensions of the GBFISS among boys: r = .32 (p < .001) for non-normative
523 subordinate masculinity, and r = .14 (p < .05) for normative hegemonic masculinity.
524 However, the correlation was non-significant (r = .03, p = .59) for normative subordinate
525 femininity. No association was found between the GBFISS and the consumption of sugar-
526 sweetened beverages among girls (r = .03, p = .54). Associations were not found (p > .05)
527 either for any of the GBFISS dimensions among girls.

528 The correlations between gender stereotype dimensions and social desirability
529 (MCSDS) were all small (Cohen, 1988), between r = .04 (p = .30), and r = .13 (p < .01),
530 suggesting the GBFISS was not strongly biased by a need for social approval.

531 **Discussion**
532

533 Despite all the research trying to disentangle the mechanisms by which gender-based
534 stereotypes might influence food choice and intake (e. g., Cavazza et al., 2015b; Kimura et
535 al., 2009; Rich et al., 2015), a valid self-report measure was still required to further the
536 understanding of gender-based stereotypes and their role in food intake behaviors. In this

manuscript, we have reported results from two studies on the development and assessment of the psychometric properties of a new scale that measures gender-based stereotypes on food intake, precisely. The scale is culturally sensitive, which is why its items reflect the practices, meanings, and values related to the gender-based cultural expectations of Costa Rican adolescents.

Our findings are encouraging since, overall, they suggest that the multidimensional GBFISS scale is supported by evidence of both concurrent and discriminant validity, as well as evidence of reliability. The dimensions identified across different samples were non-normative subordinate masculinity, normative hegemonic masculinity, and normative subordinate femininity.

In addition to providing support on construct validity, the relationship found between sexism and the GBFISS suggests that gender-based stereotypes about food intake are the expression of sexism applied to food choices. Moreover, the association of the GBFISS with different food intake behaviors provides further evidence of construct validity and suggests that sexism might account for eating behaviors. Nevertheless, we are aware that the association of gender stereotypes with the specific food preferences may vary because what is considered ‘masculine’ and ‘feminine’ might not be the same across cultures and even throughout the life span (Wardle et al., 2004).

Our findings show that, among boys, normative hegemonic and non-normative subordinate masculinity were both related to the consumption of sugary beverages, but the endorsement of normative subordinate femininity beliefs was not related. Meanwhile, in the girls’ subsample, hegemonic masculinity and normative femininity were related to less fast food consumption, but subordinate masculinity presented no contribution. In boys, both

masculinity dimensions seem to work together as normative beliefs. In girls, hegemonic masculinity and normative femininity were negatively related to fast food intake, but the same was not found for subordinate masculinity stereotypes. It appears that, for boys, both normative hegemonic and non-normative subordinate masculinity stereotypes play some normative role on behavior, whereas in girls, non-normative subordinate masculinity beliefs have no effect.

Another compelling finding is that food intake was not equally related to gender stereotypes for both boys and girls. A possibility is that boys and girls, differently, might deem the consumption of fast food and sugary beverages as an expression of masculinity or femininity. So, sugary beverages could be considered masculine by boys, but neutral by girls, and fast food might be considered masculine or "non-feminine" by girls, but neutral by boys. Although previous investigations in Costa Rica and elsewhere (Arganini et al., 2012; Carey et al., 2017; Cavazza et al., 2015a; Kimura et al., 2009, 2011; Monge-Rojas et al., 2015; Vartanian et al., 2007; Young et al., 2009) concluded that adolescents consider unhealthy foods as "masculine" and healthy foods as "feminine," future research would benefit from a more detailed examination of this attributional process, segregated by sex and by specific food items. In other countries, studies have included the task of rating how "masculine" or "feminine" participants consider specific food items (Cavazza et al., 2015b; Timeo & Suitner, 2018).

There were some study limitations and challenges. Both studies were cross-sectional and, therefore, test-retest of the GBFISS was not assessed. Future research should provide information on this. We are also aware that the development of this instrument was based on qualitative data from adolescents in Costa Rica, and that evidence of its initial validity

583 and reliability also came from Costa Rican data. Psychometric studies from diverse cultural
584 backgrounds should be conducted. Additionally, we recognize that the relationship between
585 gender-related variables and food intake is complex and that the use of different food items
586 as expressions of masculinity and femininity might vary from item to item and culture to
587 culture. Future research should examine how masculinity and femininity are assigned to
588 food-related behaviors and avoid over-simplification of this phenomenon (and the use of
589 this scale).

590 In general terms, invariance of the multi-dimensionality identified by gender and
591 age was supported; i.e., the same items reflect the same constructs, and their meaning is
592 basically the same across the gender and age groups of adolescents. However, the fit among
593 those living in urban areas was slightly not acceptable, which raised some concerns related
594 to the residence area and suggests that further research is needed to elucidate the effect of
595 urbanization on gender-based stereotypes. In general, the challenge of research in this area
596 is to develop culturally sensitive measures that also allow for meaningful cross-cultural
597 comparisons that can help to understand the impact of cultural variables on eating
598 behaviors in different settings.

599 Finally, an intriguing research direction for the future is the one mentioned on the
600 introduction: a specific scale about gender-based food intake stereotypes in adolescents
601 may help to study the specific role of these variables in well-established health behavior
602 models (e.g., Ajzen, 1991; Prochaska & DiClemente, 2005; Schwarzer, 2008) as well as in
603 habit-formation processes (e.g., Lally & Gardner, 2013) among adolescent samples.
604 Depending on the results of these studies, gender-sensitive interventions, based on sound

theoretical models, should be designed and implemented among specific groups to address gender-related inequalities and unhealthy food intake patterns.

Compliance with Ethical Standards

Conflict of Interest. The authors declare that they have no conflict of interest.

Human Participants and/or Animals. All procedures performed in studies involving human participants followed the ethical standards and local regulations concerning research with human beings. No animals were involved in the research.

Informed Consent. Informed assent was provided by the high school students before their participation, and informed consent by their parents, per Costa Rican regulations concerning research with human beings.

Bibliographic references

Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, *50*(2), 179–211. https://doi.org/10.1016/0749-5978(91)90020-T

Al-Sobayel, H., Al-Hazzaa, H. M., Abahussain, N. A., Qahwaji, D. M., & Musaiger, A. O. (2015). Gender differences in leisure-time versus non-leisure-time physical activity among Saudi adolescents. *Annals of Agricultural and Environmental Medicine*, *22*(2), 344–348. https://doi.org/10.5604/12321966.1152091

Arganini, C., Saba, A., Comitato, R., Virgili, F., & Turrini, A. (2012). Gender differences in food choice and dietary intake in modern western societies. In J. E. Maddock (Ed.), *Public Health - Social and Behavioral Health*. Intech. https://doi.org/10.5772/37886

Asamblea Legislativa. (2014). *Ley Reguladora de Investigación Biomédica N° 9234. La Gaceta. San Jose, Costa Rica: Imprenta Nacional.* http://www.mep.go.cr/sites/default/files/page/adjuntos/ley-9234-regulacionn-investigacion-biomedica.pdf

Beretvas, S. N., Meyers, J. L., & Leite, W. L. (2002). A reliability generalization study of the Marlowe-Crowne social desirability scale. *Educational and Psychological Measurement*, *62*(4), 570–589. https://doi.org/10.1177/0013164402062004003

Birrell, S. (2013). Feminist Theories for Sport. In J. Coakley & E. Dunning (Eds.), *Handbook of Sports Studies*. Sage. https://doi.org/10.4135/9781848608382.n4

Brown, R. (2010). *Prejudice: Its Social Psychology* (2nd ed.). Wiley-Blackwell.

Caine-Bish, N. L., & Scheule, B. (2009). Gender differences in food preferences of school-aged children and adolescents. *Journal of School Health*, *79*(11), 532–540. https://doi.org/10.1111/j.1746-1561.2009.00445.x

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, *56*(2), 81–105. https://doi.org/10.1037/h0046016

Cangur, S., & Ercan, I. (2015). Comparison of Model Fit Indices Used in Structural Equation Modeling Under Multivariate Normality. *Journal of Modern Applied Statistical Methods*, *79*(11), 532–540. https://doi.org/10.22237/jmasm/1430453580

Cárdenas, M., Lay, S., González, C., Calderón, C., & Alegría, I. (2010). Inventario de Sexismo Ambivalente: Adaptación y validación y relación con variable psicosociales. *Salud & Sociedad*, *1*(2), 125–135. https://doi.org/10.22199/S07187475.2010.0002.00006

Carey, J. B., Saules, K. K., & Carr, M. M. (2017). A qualitative analysis of men's experiences of binge eating. *Appetite*, *116*, 184–195. https://doi.org/10.1016/j.appet.2017.04.030

Cavazza, N., Guidetti, M., & Butera, F. (2015a). Ingredients of gender-based stereotypes about food. Indirect influence of food type, portion size and presentation on gendered intentions to eat. *Appetite*, *91*, 266–272. https://doi.org/10.1016/j.appet.2015.04.068

Cavazza, N., Guidetti, M., & Butera, F. (2015b). The gender-based stereotype about food is on the table. Food choice also depends on co-eater's gender. *Psicologia Sociale*, *10*(2), 161–172. https://doi.org/10.1482/80763

Chen, F. F., Sousa, K. H., & West, S. G. (2005). Testing measurement invariance of second-order factor models. *Structural Equation Modeling*, *12*(3), 471–492. https://doi.org/10.1207/s15328007sem1203_7

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling*, *9*(2), 233–255. https://doi.org/10.1207/S15328007SEM0902_5

Chinnock, A. (2007). *Diario de Consumo de Alimentos. Instrumento para el registro de información*. UCR.

Chinnock, A. (2010). *Programa para el cálculo de valor nutritivo de los alimentos VALORNUT*. http://nutricion2.ucr.ac.cr/

Clément-Guillotin, C., Chalabaev, A., & Fontayne, P. (2011). Is sport still a masculine domain? *International Journal of Sport Psychology*, *43*, 67–78. https://doi.org/10.1037/t03782-000

Cohen, J. (1988). Statistical power analysis for the behavioral sciences. In *Statistical Power Analysis for the Behavioral Sciences* (Vol. 2nd). Erlbaum.

Connell, R. W. (1995). *Masculinities*. Polity Press.

Connell, R. W., & Messerschmidt, J. W. (2005). Hegemonic masculinity: Rethinking the concept. *Gender and Society*, *19*, 829–859. https://doi.org/10.1177/0891243205278639

Cornwall, A. (2000). Missing men? Reflections on men, masculinities and gender in GAD. *IDS Bulletin*, *31*(2), 1–6. https://doi.org/10.1111/j.1759-5436.2000.mp31002003.x

Cornwall, A., & White, S. C. (2000). Men, masculinities and development politics, policies and practice. *IDS Bulletin*, *31*(2), 1–6. https://doi.org/10.1111/j.1759-5436.2000.mp31002001.x

Craigie, A. M., Lake, A. A., Kelly, S. A., Adamson, A. J., & Mathers, J. C. (2011). Tracking of obesity-related behaviors from childhood to adulthood: A systematic review. *Maturitas*, *70*(3), 266–284. https://doi.org/10.1016/j.maturitas.2011.08.005

Crowne, D. P., & Marlowe, D. (1960). A new scale of social desirability independent of psychopathology. *Journal of Consulting Psychology*, *24*(4), 349–354. https://doi.org/10.1037/h0047358

Crutzen, R., & Peters, G. J. (2017). Scale quality: alpha is an inadequate estimate and factor-analytic evidence is needed first of all. *Health Psychology Review*, *11*(3), 242–247. https://doi.org/10.1080/17437199.2015.1124240

Cruz, F., Ramos, E., Lopes, C., & Araújo, J. (2018). Tracking of food and nutrient intake from adolescence into early adulthood. *Nutrition*, *55*, 84–90. https://doi.org/10.1016/j.nut.2018.02.015

Dardenne, B., Dumont, M., & Bollier, T. (2007). Insidious Dangers of Benevolent Sexism: Consequences for Women's Performance. *Journal of Personality and Social Psychology*, *93*(5), 764–779. https://doi.org/10.1037/0022-3514.93.5.764

De Lemus, S., Spears, R., Bukowski, M., Moya, M., & Lupiáñez, J. (2013). Reversing implicit gender stereotype activation as a function of exposure to traditional gender roles. *Social Psychology*, *44*(2), 109–116. https://doi.org/10.1027/1864-9335/a000140

Eccles, J. (2011). Gendered educational and occupational choices: Applying the Eccles et al. model of achievement-related choices. *International Journal of Behavioral Development*, *35*(3), 195–201. https://doi.org/10.1177/0165025411398185

Ellemers, N. (2018). Gender Stereotypes. *Annual Review of Psychology*, *69*, 275–298. https://doi.org/10.1146/annurev-psych-122216-011719

Etchezahar, E., & Ungaretti, J. (2014). Woman stereotypes and ambivalent sexism in a sample of adolescents from Buenos Aires. *Journal of Behavior, Health and Social Issues*, *6*(2), 87–94. https://doi.org/10.5460/jbhsi.v6.2.41328

Furr, R. M. (2017). *Psychometrics: an introduction*. SAGE Publications Sage CA: Los Angeles, CA.

Glick, P., & Fiske, S. T. (1996). The ambivalent sexism inventory. *Journal of Personality and Social Psychology*, *70*(3), 491–512. https://doi.org/10.1007/s00572-009-0231-8

Glick, P., Sakallı-Ugurlu, N., Ferreira, M. C., & Aguiar de Souza, M. (2002). Ambivalent

sexism and attitudes toward wife abuse in Turkey and Brazil. *Psychology of Women Quarterly*, *35*(3), 530–535. https://doi.org/10.1111/1471-6402.t01-1-00068

Glick, P., Wilkerson, M., & Cuffe, M. (2015). Masculine Identity, Ambivalent Sexism, and Attitudes Toward Gender Subtypes. *Social Psychology*, *46*(4), 210–217. https://doi.org/10.1027/1864-9335/a000228

Gutierrez, S., Sanz, J., Espinosa, R., Gesteira, C., & Paz Garcia-Vera, M. (2016). The Marlowe-Crowne Social Desirability Scale: Norms for the Spanish general population and development of a short version. *Anales de Psicología*, *32*(1), 206–2017. https://doi.org/10.4271/2013-01-2491

Hannon, J., Soohoo, S., Reel, J., & Ratliffe, T. (2009). Gender stereotyping and the influence of race in sport among adolescents. *Research Quarterly for Exercise and Sport*, *80*(3), 676–684. https://doi.org/10.1080/02701367.2009.10599608

Hardin, M., & Greer, J. (2009). The Influence of Gender-Role Socialization, Media Use and Sports Participation on Perceptions of Gender-Appropriate Sports. *Journal of Sport Behavior*, *32*(2), 207–226.

Herman, C. P., & Polivy, J. (2010). Sex and Gender Differences in Eating Behavior. In J. C. McCreary & D. R. Chrisler (Eds.), *Handbook of Gender Research in Psychology* (pp. 455–469). Springer. https://doi.org/10.1007/978-1-4419-1465-1_22

Herman, C. P., Polivy, J., Pliner, P., Vartanian, L. R., Herman, C. P., Polivy, J., Pliner, P., & Vartanian, L. R. (2019). Consumption Stereotypes and Impression Management: Food Choice. In *Social Influences on Eating* (pp. 79–94). Springer. https://doi.org/10.1007/978-3-030-28817-4_7

Hooper, D., Coughlan, J., & Mullen, M. R. (2008). Structural equation modeling: Guidelines for determining model fit. *Electronic Journal of Business Research Methods*, *6*(1), 53–60. https://doi.org/10.1016/j.cgh.2014.08.036

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Structural Equation Modeling*, *6*(1), 1–55. https://doi.org/10.1080/10705519909540118

Hyde, J. S. (2005). The gender similarities hypothesis. *American Psychologist*, *60*(6), 581–592. https://doi.org/10.1037/0003-066X.60.6.581

Igenoza, M. (2017). Race, Femininity and Food: Femininity and the Racialization of Health and Dieting. *International Review of Social Research*, *7*(2), 109–118. https://doi.org/10.1515/irsr-2017-0013

Jost, J. T., & Kay, A. C. (2005). Exposure to benevolent sexism and complementary gender stereotypes: Consequences for specific and diffuse forms of system justification. *Journal of Personality and Social Psychology*, *88*(3), 498–509. https://doi.org/10.1037/0022-3514.88.3.498

Kimura, A., Wada, Y., & Dan, I. (2011). Gender-based food stereotypes among young Japanese. In *Handbook of Behavior, Food and Nutrition* (pp. 2201–2213). Springer. https://doi.org/10.1007/978-0-387-92271-3_140

Kimura, A., Wada, Y., Goto, S., Tsuzuki, D., Cai, D., Oka, T., & Dan, I. (2009). Implicit gender-based food stereotypes. Semantic priming experiments on young Japanese. *Appetite*, *52*(2), 521–524. https://doi.org/10.1016/j.appet.2008.11.002

Kline, R. B. (2011). *Principles and Practice of Structural Equation Modeling*. The Guilford Press. https://doi.org/10.1017/CBO9781107415324.004

Kurz, A. S., Drescher, C. F., Chin, E. G., & Johnson, L. R. (2016). Measuring social desirability across language and sex: A comparison of Marlowe-Crowne Social Desirability Scale factor structures in English and Mandarin Chinese in Malaysia. *PsyCh Journal*, *5*(2), 92–100. https://doi.org/10.1002/pchj.124

Lally, P., & Gardner, B. (2013). Promoting habit formation. *Health Psychology Review*, *7*(1), 137–158. https://doi.org/10.1080/17437199.2011.603640

Le, T. P. (2019). The association of conformity to feminine norms with women's food consumption after a negative mood induction. *Appetite*, *133*(October 2018), 123–129. https://doi.org/10.1016/j.appet.2018.10.031

Leite, W. L., & Beretvas, S. N. (2005). Validation of scores on the Marlowe-Crowne social desirability scale and the balanced inventory of desirable responding. *Educational and Psychological Measurement*, *65*(1), 140–154. https://doi.org/10.1177/0013164404267285

Lips, H. M. (2020). Sex & Gender: An introduction. In *Sex & Gender: An introduction*. Waveland Press.

Lombardi, C. M. P., Coley, R. L., Sims, J., Lynch, A. D., & Mahalik, J. R. (2019). Social Norms, Social Connections, and Sex Differences in Adolescent Mental and Behavioral Health. *Journal of Child and Family Studies*, *28*(1), 91–104. https://doi.org/10.1007/s10826-018-1253-7

Lyons, A. C. (2009). Masculinities, Femininities, Behaviour and Health. *Social and Personality Psychology Compass*, *3*(4), 394–412. https://doi.org/10.1111/j.1751-9004.2009.00192.x

Messerschmidt, J. W. (2019). The Salience of "Hegemonic Masculinity." *Men and Masculinities*, *22*(1), 85–91. https://doi.org/10.1177/1097184X18805555

Messerschmidt, J. W., Messner, M. A., Connell, R., & Martin, P. Y. (2018). *Gender reckonings: New social theory and research* (N. Y. U. Press (Ed.)).

Mikkilä, V., Räsänen, L., Raitakari, O. T., Pietinen, P., & Viikari, J. (2005). Consistent dietary patterns identified from childhood to adulthood: The Cardiovascular Risk in Young Finns Study. *British Journal of Nutrition*, *93*(6), 923–931. https://doi.org/10.1079/bjn20051418

Monge-Rojas, R., Fuster-Baraona, T., Garita, C., Sánchez, M., Smith-Castro, V., Valverde-Cerros, O., & Colon-Ramos, U. (2015). The influence of gender stereotypes on eating habits among Costa Rican adolescents. *American Journal of Health Promotion*, *29*(5), 303–310. https://doi.org/10.4278/ajhp.130904-QUAL-462

Movassagh, E. Z., Baxter-Jones, A. D. G., Kontulainen, S., Whiting, S. J., & Vatanparast, H. (2017). Tracking dietary patterns over 20 years from childhood through adolescence into young adulthood: The Saskatchewan pediatric bone mineral accrual study. *Nutrients*, *9*(9), 990. https://doi.org/10.3390/nu9090990

Moya, M., & Expósito, F. (2001). Nuevas formas, viejos intereses: Neosexismo en varones españoles. *Psicothema*, *13*(4), 668–674.

North, M. S., & Fiske, S. T. (2014). Social categories create and reflect inequality: Psychological and sociological insights. In *The Psychology of Social Status* (pp. 243–265). Springer. https://doi.org/10.1007/978-1-4939-0867-7_12

Ortega, R. M., Perez-Rodrigo, C., & Lopez-Sobaler, A. M. (2015). Dietary assessment methods: Dietary records [Métodos de evaluación de la ingesta actual: Registro o diario dietético]. *Nutricion Hospitalaria*, *31*(3), 38–45. https://doi.org/10.3305/nh.2015.31.sup3.8749

Perez, M., Labiano, M., & Brusasca, C. (2010). Escala de deseabildad social: analisis psicometrico en muestra argentina. *Evaluar*, *10*(1), 43–59.

Peters, Gjalt Jorn. (2014). The alpha and the omega of scale reliability and validity. *The European Health Psychologist*, *16*(2), 56–69. https://doi.org/0.31234/osf.io/h47fv

Peters, Gjakt Jorn, Verboon, P., & Green, J. (2018). *Package 'userfriendlyscience'. Vienna: R Foundation for Statistical Computing.* http://bioconductor.statistik.tu-dortmund.de/cran/web/packages/userfriendlyscience/userfriendlyscience.pdf

Pistella, J., Tanzilli, A., Ioverno, S., Lingiardi, V., & Baiocco, R. (2018). Sexism and Attitudes Toward Same-Sex Parenting in a Sample of Heterosexuals and Sexual Minorities: the Mediation Effect of Sexual Stigma. *Sexuality Research and Social Policy*, *15*(2), 139–150. https://doi.org/10.1007/s13178-017-0284-y

Plaza, M., Boiché, J., Brunel, L., & Ruchaud, F. (2017). Sport = Male… But Not All Sports: Investigating the Gender Stereotypes of Sports Activities at the Explicit and Implicit Levels. *Sex Roles*. https://doi.org/10.1007/s11199-016-0650-x

Prochaska, J. O., & DiClemente, C. (1982). Transtheoretical therapy: Toward a more integrative model of change. *Psychotherapy: Theory Research and Practice*, *19*(3), 276–288. https://doi.org/10.1037/h0088437

Prochaska, J. O., & DiClemente, C. (2005). The transtheoretical approach. In J. C. Norcross & M. R. Goldfried (Eds.), *Handbook of Psychotherapy Integration* (2nd ed., pp. 147–171). Oxford University Press.

Programa Estado de la Nación. (2019). *Sétimo informe estado de la educación. Masterlitho: San José, Costa Rica.* https://estadonacion.or.cr/informes/

Rich, E. P., Nkosi, S., & Morojele, N. K. (2015). Masculinities, alcohol consumption, and sexual risk behavior among male tavern attendees: A qualitative study in North West Province, South Africa. *Psychology of Men and Masculinity*, *16*(4), 382–392. https://doi.org/10.1037/a0038871

Ryan, T. P. (2013). Sample Size Determination and Power. In *Sample Size Determination and Power*. John Wiley & Sons. https://doi.org/10.1002/9781118439241

Schippers, M. (2007). Recovering the feminine other: Masculinity, femininity, and gender hegemony. *Theory and Society*, *36*(1), 85–102. https://doi.org/10.1007/s11186-007-9022-4

Schneider, B. C., De Carvalho Dumith, S., Lopes, C., Severo, M., & Assunção, M. C. F. (2016). How do tracking and changes in dietary pattern during adolescence relate to the amount of body fat in early adulthood? *PLoS ONE*, *11*(2), e0149299. https://doi.org/10.1371/journal.pone.0149299

Schwarzer, R. (2008). Modeling health behavior change: How to predict and modify the adoption and maintenance of health behaviors. *Applied Psychology*, *57*(1), 1–29. https://doi.org/10.1111/j.1464-0597.2007.00325.x

Skinner, C. J. (2014). Probability Proportional to Size (PPS) Sampling. In *Wiley StatsRef: Statistics Reference Online*. https://doi.org/10.1002/9781118445112.stat03346.pub2

Smith-Castro, V. (Ed.). (2014). *Compendio de Instrumentos de Medición IIP. Serie Cuadernos Metodológicos.* Instituto de Investigaciones Psicológicas.

Steele, C. M., & Aronson, J. (1995). Stereotype Threat and the Intellectual Test Performance of African Americans. *Journal of Personality and Social Psychology*. https://doi.org/10.1037/0022-3514.69.5.797

Steele, C. M., Spencer, S. J., & Aronson, J. (2004). *Contending with group image: The psychology of stereotype and social identity threat*. https://doi.org/10.1016/s0065-2601(02)80009-0

Timeo, S., & Suitner, C. (2018). Eating meat makes you sexy: Conformity to dietary gender norms and attractiveness. *Psychology of Men and Masculinity*, *19*(3), 418–429. https://doi.org/10.1037/men0000119

UCR. (2013). *Sistema de Información Estadística de Derechos de la Niñez y Adolescencia (SIEDNA). Personas menores de edad a la luz del censo 2011*. https://www.inec.go.cr/sites/default/files/documentos/inec_institucional/estadisticas/resultados/repoblaccenso2011-03.pdf.pdf

Vartanian, L. R., Herman, C. P., & Polivy, J. (2007). Consumption stereotypes and impression management: How you are what you eat. *Appetite*, *48*(3), 265–277. https://doi.org/10.1016/j.appet.2006.10.008

Ventimiglia, M., & MacDonald, D. A. (2012). An examination of the factorial dimensionality of the Marlowe Crowne Social Desirability Scale. *Personality and Individual Differences*, *52*(4), 487–491. https://doi.org/10.1016/j.paid.2011.11.016

Ventura-León, J. L., & Caycho-Rodríguez, T. (2017). El coeficiente Omega: un método alternativo para la estimación de la confiabilidad. *Revista Latinoamericana En Ciencias Sociales, Niñez y Juventud*, *15*(1), 26–27.

Vésteinsdóttir, V., Reips, U. D., Joinson, A., & Thorsdottir, F. (2015). Psychometric

properties of measurements obtained with the Marlowe-Crowne Social Desirability Scale in an Icelandic probability-based Internet sample. *Computers in Human Behavior*, *49*, 608–614. https://doi.org/10.1016/j.chb.2015.03.044

Wardle, J., Haase, A. M., Steptoe, A., Nillapun, M., Jonwutiwes, K., & Bellisle, F. (2004). Gender differences in food choice: the contribution of health beliefs and dieting. *Annals of Behavioral Medicine*, *27*(2), 107–116. https://doi.org/10.1207/s15324796abm2702_5

Young, M. E., Mizzau, M., Mai, N. T., Sirisegaram, A., & Wilson, M. (2009). Food for thought. What you eat depends on your sex and eating companions. *Appetite*, *53*(2), 268–271. https://doi.org/10.1016/j.appet.2009.07.021

881 **Table 1**. Description of invariance levels tested

Invariance level	Constraints involved	Interpretation
Configural level	No constraints between subgroups	The same set of items reflects the same latent constructs across subgroups.
Metric level (First-order measurement weights)	First-order factor loadings are constrained to be equal across groups.	The strength of the relationship between *each* item and its underlying construct is the same for both groups.
Scalar level (Intercepts of measured variables)	First-order factor loadings and intercepts are constrained to be equal across groups.	The same set of items reflects the same first-order latent constructs, and their meanings are the same across subgroups.
Structural weights level (Second-order factor loadings)	First-order factor loadings and intercepts, as well as second-order factor loadings, are constrained to be equal across groups	The strength of the relationship between *each* first-order construct and its underlying second-order construct is the same for both groups.
Structural covariances level (Second-order covariance)	First-order factor loadings and intercepts, as well as second-order factor loadings and covariance(s), are constrained to be equal across groups	The same set of items reflects the same first-order latent constructs, the same set of first-order constructs reflects the same second-order latent construct(s), and their meanings are the same across subgroups.
Structural residuals level (Disturbances of first-order factors)	First-order factor loadings and intercepts, as well as second-order factor loadings and covariance(s), are constrained to be equal across groups.	The same set of items reflects the same first-order latent constructs, the same set of first-order constructs reflect the same second-order latent construct(s), and their meanings are the same across subgroups. Additionally, there is no appreciable difference in the disturbances.

882

Table 2. Exploratory Factor Analysis: item-to-factor loading

	Items	Factor 1 Non-normative subordinate masculinity	Factor 2 Normative subordinate femininity	Factor 3 Normative hegemonic masculinity
3.	A man who only eats salads is definitely gay	.68		
4.	Men who bring fruits to school are usually effeminate	.67		
5.	Men who watch what they eat to avoid gaining weight are gay	.76		
6.	A man who eats little is gay	.82		
7.	Men who eat healthy food to stay in shape are effeminate	.73		
8.	Men who eat slowly are effeminate	.73		
9.	Queer men mind their manners when eating	.55		
10.	Men who eat little are gay	.78		
12.	Men prefer women who watch what they eat		.41	
13.	Women who eat quickly appear less feminine		.44	
14.	Beautiful women generally eat little		.56	
15.	Women who don't watch what they eat are not appealing to men		.67	
16.	The more feminine a woman is, the more fruits she eats		.64	
17.	If a woman wants to be successful with men, she must watch what she eats		.63	
19.	A woman who eats a lot looks manly		.59	
21.	Thin women are more feminine		.55	
1.	An average man eats a lot			.52
2.	Real men eat very quickly			.40
11.	Men don't care if the food they eat is greasy			.74
18.	Men eat whatever they want without remorse			.74
20.	Men do not care about what they eat			.58

Note: In this table, items are freely translated from Spanish into English. The original items in Spanish are provided in Appendix 1. KMO = .868, Bartlett test = 1853.05 ($p < .001$). Item numbers are reported based on the order they had in the study questionnaire.

Table 3. Fit of gender, age, and place of residence subgroups in Study 2

Fit by group categories	χ^2	$\chi^{2/}$df	CFI	RMSEA [90% CI]	χ^2	$\chi^{2/}$df	CFI	RMSEA [90% CI]
Gender			Boys				Girls	
	407.19	2.19	.90	.063 [.055, .072]	476.94	2.56	.91	.055 [.049, .61]
Age			Younger				Older	
	450.67	2.42	.92	.055 [.048, .061]	440.75	2.37	.90	.064 [.056, .71]
Residence area			Urban				Rural	
	588.41	3.16	.88	.073 [.066, .079]	400.17	2.15	.91	.053 [.046, .61]

Note: Degrees of freedom were 186 for all the analyses in these groups. There were 297 boys and 516 girls, 475 younger (< 15 years) and 338 older (> 15 years) participants, and 409 urban and 404 rural inhabitants.

Table 4. Invariance results by gender and age subgroups in Study 2

	Gender groups					Age groups				
Invariance level	χ^2	df	χ^2/df	CFI	RMSEA [90% CI]	χ^2	df	χ^2/df	CFI	RMSEA [90% CI]
Configural	884.26	372	2.37	.91	.041 [.038, .045]	891.48	372	2.39	.91	.041 [.038, .45]
Metric (Measurement weights)	943.35	390	2.42	.90	.042 [.038, .045]	929.19	390	2.38	.91	.041 [.038, .45]
Scalar (Measurement intercepts)	970.87	411	2.36	.90	.041 [.038, .044]	1004.55	411	2.44	.90	.042 [.039, .46]
Second-order loadings (Structural weights)	974.21	413	2.36	.90	.041 [.038, .044]	1008.64	413	2.44	.90	.042 [.039, .45]
Second-order covariance (structural covariance)	975.31	414	2.35	.90	.041 [.038, .044]	1011.70	414	2.44	.90	.042 [.039, .45]
Second-order residuals (structural residuals)	994.45	417	2.38	.90	.041 [.038, .045]	1013.98	417	2.43	.90	.042 [.039, .45]

Note: ***$p < .001$, **$p < .01$

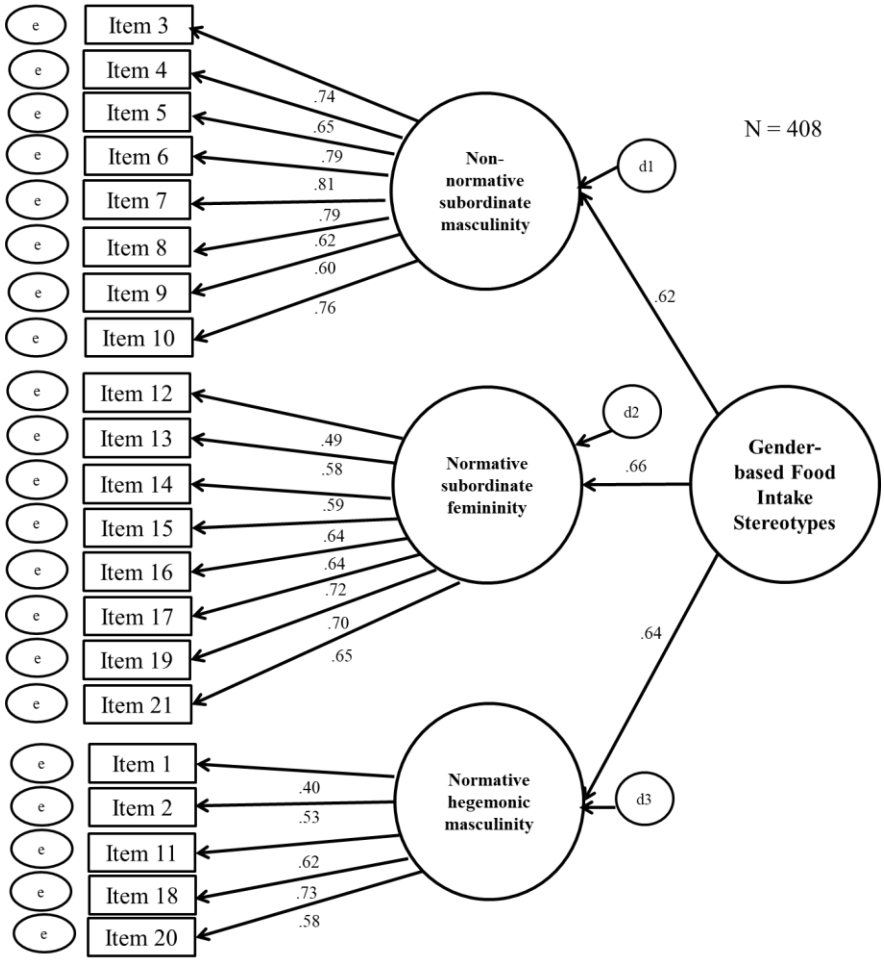

Figure 1. Note. Fit model: $\chi 2$ (186) = 457.27, $p < .001$, $\chi 2$ /df = 2.46, CFI = .91, RMSEA = .060, 90% CI [.053; .067]. Coefficients are standardized. No item-factor loading was below the recommended level of β = .30 (Kline, 2011). Loadings were all significant ($p < .001$).

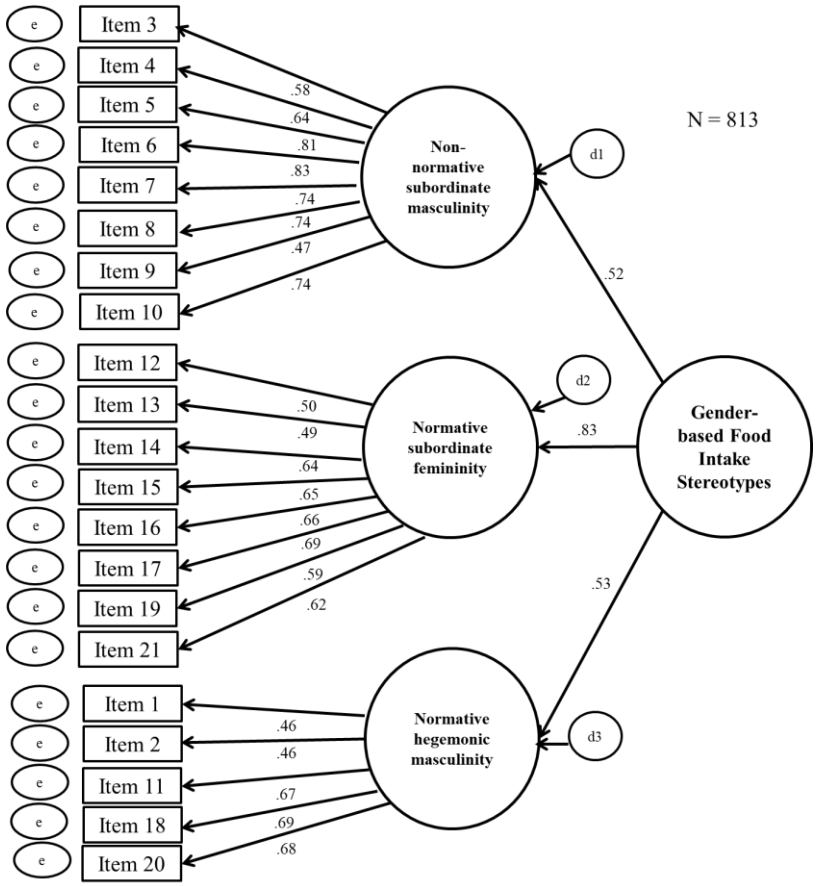

Figure 2. Note: $\chi 2$ (186) = 618.65, $p < .001$, $\chi 2$ /df = 3.32, CFI = .92, RMSEA = .053, 90% CI [.049; .058]. Coefficients are standardized. No item-factor loading was below the recommended level of $\beta = .30$ (Kline, 2011). Loadings were all significant ($p < .001$).