Archives of Gerontology and Geriatrics xxx (2014) xxx–xxx

Contents lists available at ScienceDirect

Archives of Gerontology and Geriatrics

journal homepage: www.elsevier.com/locate/archger

The validity of the 12-item Bem Sex Role Inventory in older Spanish population: An examination of the androgyny model

Afshin Vafaei [a,*], Beatriz Alvarado [a], Concepcion Tomás [b], Carmen Muro [b], Beatriz Martinez [b], Maria Victoria Zunzunegui [c]

[a] Department of Public Health Sciences, Queen's University, Kingston, Canada
[b] School of Health Sciences, University of Zaragoza, Zaragoza, Spain
[c] Département de médecine sociale et préventive, Faculté de médecine, Université de Montréal, Montreal, Québec, Canada

ARTICLE INFO

Article history:
Received 17 April 2014
Accepted 31 May 2014
Available online xxx

Keywords:
Bem Sex Role Inventory
Gender roles
Factor analysis
Androgyny model
Older adults
Validation

ABSTRACT

The Bem Sex Role Inventory (BSRI) is the most commonly used and validated gender role measurement tool across countries and age groups. However, it has been rarely validated in older adults and sporadically used in aging and health studies. Perceived gender role is a crucial part of a person's identity and an established determinant of health. Androgyny model suggests that those with high levels of both masculinity and femininity (androgynous) are more adaptive and hence have better health. Our objectives were to explore the validity of BSRI in an older Spanish population, to compare different standard methods of measuring gender roles, and to examine their impact on health indicators. The BSRI and health indicator questions were completed by 120 community-dwelling adults aged 65+ living in Aragon, Spain. Exploratory factor analysis was performed to examine psychometric properties of the BSRI. Androgyny was measured by three approaches: geometric mean, *t*-ratio, and traditional four-gender groups classification. Relationships between health indicators and gender roles were explored. Factor analysis resulted in two-factor solution consistent with the original masculine and feminine items with high loadings and good reliability. There were no associations between biological sex and gender roles. Different gender role measurement approaches classified participants differently into gender role groups. Overall, androgyny was associated with better mobility and physical and mental health. The traditional four groups approach showed higher compatibility with the androgyny model and was better able to disentangle the differential impact of gender roles on health.

1. Introduction

1.1. Sex versus gender

Despite the fact that many researchers use the terms sex and gender interchangeably, they are two different entities. Sex refers to the biological and physiological characteristics that defines a person as a man or woman such as body size and shape, type and function of reproductive organs, and hormonal activities; whereas gender is the array of socially constructed roles and behaviors, personality traits, attitudes, and values that is considered to be socially appropriate for individuals of a specific sex in the context of a specific culture (Schmitz, 2010; Weiten, 1997). Gender roles are often a self-perceived construct and are based on how individuals identify themselves as masculine or feminine (Johnson, Greaves, & Repta, 2007). Studies have shown that gender roles independent of biological sex affect health (Doyal, 2000; Gale-Ross, Baird, & Towson, 2009). Generally, it is shown that having masculine gender role is associated with worse physical health (Månsdotter, Lundin, Falkstedt, & Hemmingsson, 2009) and high femininity relates to worse mental health (Bassoff & Glass, 1982). High expressiveness as an attribute of femininity was associated with lower mortality from coronary heart disease in middle-aged Scottish men (Hunt, Lewars, Emslie, & Batty, 2007) and was related to less frequent alcohol use in adolescents (Huselid & Cooper, 1992) and university students (Zimmermann, Sieverding, & Müller, 2011). Emotional suppression internalized during masculine-role socialization can result in hazardous health-related behaviors such as smoking and heavy drinking (Lee, 2002). Adherence to masculine role of excessive self-reliance and dominance, further

* Corresponding author at: Department of Public Health Sciences, Queen's University, Carruthers Hall 2nd Floor, 62 Fifth Field Company Lane, Kingston, Ontario K7L 3N6, Canada. Tel.: +1 416 561 2919.
E-mail address: 5av14@queensu.ca (A. Vafaei).

http://dx.doi.org/10.1016/j.archger.2014.05.012
0167-4943/

impede appropriate health service accessing (Macdonald, Gibbs, & Oliffe, 2004; Seymour-Smith, Wetherell, & Phoenix, 2002).

1.2. Measuring gender roles

Several approaches have been used to measure gender roles, from pure theoretical conceptualization to operationalization. These approaches are based on implicit assumptions that social forces produce gender roles and they are related to health (Spence, Helmreich, & Stapp, 1974). Scales such as the Personality Research Form ANDRO Scales (Berzins, Welling, & Wetter, 1978), the Adjective Check list (ACL) M and F scales (Lenney, 1991), the Personal Attributes Questionnaire (Spence et al., 1974), and the Bem Sex Role Inventory (BSRI) (Bem, 1974) have been used to measure gender roles and to examine their relationship to health.

BSRI was developed four decades ago by Bem (1974) and remains the most commonly used validated scale until today. In studying gender roles, she developed 60 personality traits and categorized them into two groups of instrumental and expressive traits. Instrumental traits such as having strong personality, being dominant, and taking the lead were considered as masculine gender roles; while expressive traits such as being tender, sympathetic, and gentle were thought to indicate feminine gender roles.

1.3. The androgyny model

Bem (1974) was the first person who argued against the exclusive dichotomy of gender roles and defined four gender roles: A person with high masculine and low feminine identification would be categorized as 'masculine', a person with high feminine identification and low masculine identification would be categorized as 'feminine', a person who had high identification with both characteristics would be categorized as 'androgynous', and finally a person who has low identification with both dimensions would be considered 'undifferentiated' (Bem, 1974; Hoffman & Fidell, 1979). She hypothesized that 'androgynous' individuals regardless of their biological sex, depending on the situational appropriateness can be instrumental and assertive or expressive and yielding (Bem, 1974). This flexibility in behavior results in being more adaptive and therefore more likely to have better mental health and higher competence (Bem, 1977). Research has provided support for this theory; in adults, androgyny has been associated with higher levels of optimal mental health (Lefkowitz & Zeldow, 2006) and better health practices (Shifren & Bauserman, 1996). A study of a sample of older adults in Windsor, Ontario, showed that having androgynous gender role is associated with better self-rated physical health, wellness, and life satisfaction (Gale-Ross et al., 2009).

Measurement of androgyny has also been subjected to debate. Generally, classification is done using participants' scores on the masculine and feminine scales. Using the medians from Bem's original normative samples (Lenney, 1991), the theoretical mean of the scale (Fernandez & Coello, 2010), or the sample medians (Gale-Ross et al., 2009; Lenney, 1991; Sedney, 1981), individuals are classified into four possible gender roles groups according to their masculinity and femininity scores (Fig. 1). A person is classified as androgynous if both scores are above the scale medians; and as undifferentiated if both scores are below the scale medians (Bem, 1977; Spence et al., 1974).

A person's androgyny may also be expressed as a continuous variable calculated according to the difference between femininity and masculinity scores divided by a term reflecting the variance of the two sets of ratings (Bem, 1974). Other methods include using the absolute value of differences between the masculine score and the feminine score (Heilbrun, 1981; Kalin, 1979), the arithmetic mean (Strahan, 1981), and the geometric mean (Bryan, Coleman, & Ganong, 1981).

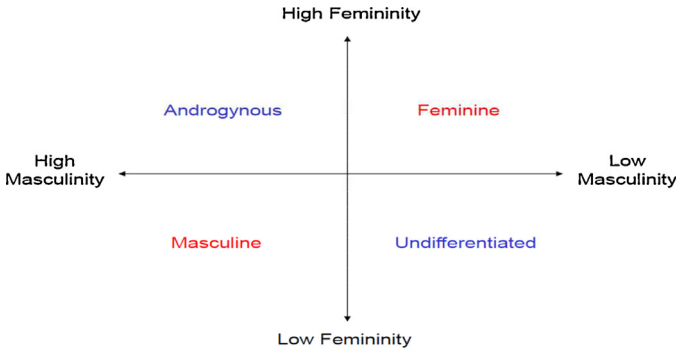

Fig. 1. The four traditional gender roles.

1.4. International use of BSRI

The original 60 item and shorter versions of the BSRI have been used in various countries (e.g. France (Alain, 1987), Germany (Streiner & Norman, 2008), Spain (Mateo & Fernández, 1991), Japan (Katsurada & Sugihara, 1999; Sugihara & Katsurada, 2000), India and Malaysia (Ward & Sethi, 1986), China (Zhang, Norvilitis, & Jin, 2001), Turkey (Özkan & Lajunen, 2005), Canada (Gale-Ross et al., 2009), Brazil (Carver, Vafaei, Guerra, Freire, & Phillips, 2013; Hernandez, 2009)), and across different age groups of adolescents (Fontayne, Sarrazin, & Famose, 2000; Vega, 2007; Wilcox & Francis, 1997), adults (Behar, De la Barrera, & Michelotti, 2001; Bledsoe, 1983; Mateo & Fernández, 1991), and seniors (Carver et al., 2013; Gale-Ross et al., 2009; Windle & Sinnott, 1985) and appear to be reliable and valid across geography, age groups, and culture. However, to improve cultural fit a few items were removed when the BSRI was used in Zimbabwe (Wilson et al., 1990), Japan, and China (Katsurada & Sugihara, 1999). It has been suggested that age modifies the BRSI factor structure (Blanchard-Fields, Suhrer-Roussel, & Hertzog, 1994; Campbell, Gillaspy, & Thompson, 1997), thus the need for further validation in new older adult populations.

1.5. Gender and aging

Contrary to study of sex differences in health and well-being of older adults, there is limited research on gender and aging. Most available aging research has focused on health impacts of gender roles in separate populations of men (Downey, 1984) and women (Krames, England, & Flett, 1988) or have looked into changes in gender roles as people age not specifically gender-health relationships (Sinnott & Shifren, 2001; Twenge, 1997).

Given this gap in knowledge, the need for re-validation of the BSRI in different cultures and age groups, and the necessity for proper measurement of androgyny, the objectives of our study were: (1) to evaluate the construct validity of the Spanish version of 12-item BRSI scale in a Spanish older adults population, (2) to construct and compare different measures of androgyny, and (3) to compare effects of sex, self-perceived gender roles, and androgyny on physical and mental health of the aforementioned population.

2. Methods

2.1. Setting and participants

Community-dwelling older adults aged from 65 to 75, attending two rural and one urban primary care centers in Aragon, Spain were invited to participate in the study. All patients

Please cite this article in press as: Vafaei, A., et al., The validity of the 12-item Bem Sex Role Inventory in older Spanish population: An examination of the androgyny model. Arch. Gerontol. Geriatr. (2014), http://dx.doi.org/10.1016/j.archger.2014.05.012

regardless of the reason for visiting the primary care unit (chronic care, acute illnesses, prevention modalities) were eligible for this study. The study aims and procedure were explained to those in the study age range while they were in the waiting room of the neighborhood clinic. Those who agreed to participate were asked to sign an informed consent. Interviews were performed by three interviewers who were trained in questionnaire administration and data entry. Information on socio-demographic factors, gender roles, physical and mental health status, and mobility limitations were collected. The study was approved by the Research Ethics Board of the primary care centers.

2.2. Measures

2.2.1. Socio-demographic factors

Collected information included age, sex, marital status, level of education, and sufficiency of income to cover basic needs.

2.2.2. Gender roles

We used the shorter 12-item version of BSRI to measure participants' instrumental (masculine) and expressive (feminine) traits and thereby their gender roles. These items have been validated in the culturally similar population of Spanish students (Mateo & Fernández, 1991) and Brazilian older adults (Carver et al., 2013). Participants were asked to rank in five categories ranging from "strongly agree" to "strongly disagree" if they perceive themselves as gentle, warm, dominant, tender, sympathetic, affectionate, with leadership abilities, possessing strong personality and also if they act like a leader, defend own beliefs, make decision easily, and are sensitive to others' needs.

2.2.3. Health and physical functioning indicators

2.2.3.1. Physical health. Self-rated health (SRH) was utilized to estimate the level of general health in this population. SRH has been used extensively as an indicator of general health in various studies (Blakely, Lochner, & Kawachi, 2002; Gale-Ross et al., 2009; Subramanian, Kawachi, & Kennedy, 2001) and has proven to be a good predictor of mortality in studies of older adults worldwide (Idler & Benyamini, 1997) and in Spain (Sanchez-Santos, Zunzunegui, Otero-Puime, Cañas, & Casado-Collado, 2011). Also in a longitudinal study SRH was a determinant of functional ability in older adults (Idler & Kasl, 1995). Participants were asked to rate their health status on a five point Likert scale (very good, good, fair, poor, very poor). For analysis, 'very good' and 'good' answers were combined to create a *good* category and other options were considered the *poor* SRH category.

2.2.3.2. Mental health. Depression as an indicator of mental health was measured by the validated scale of *the Center for Epidemiological Studies Depression* (CESD) (Radloff, 1977). Items include depressive mood, feelings of worthlessness and helplessness, negative affect, and psychomotor retardation. Responses are based on the frequency of occurrence during the past week. CESD uses a 4-point ordinal scale: rarely or none of the time (less than 1 day); some or a little of the time (1–2 days); occasionally or a moderate amount of the time (3–4 days); most or all of the time (5–7 days). In this study, the 10-item version of the scale was employed with scores equal or greater than 10 suggestive of depression (Robison, Gruman, Gaztambide, & Blank, 2002).

2.2.3.3. Mobility limitations. Mobility was measured by two items of the Nagi standard tool (Nagi, 1976). Participants were asked to report whether they have difficulty in one or both of these two tasks: climbing 10 steps of stairs without resting and walking 1 km. Response options were 'none', 'some difficulty', 'a lot of difficulty', and 'unable to perform the activity'.

2.3. Statistical analysis

2.3.1. Descriptive

To obtain a general overview of the population under study, distributions of socio-demographic and health status variables across sexes were estimated and compared using *t*-test and Chi-square tests.

2.3.2. Factor analysis

To assess the construct validity of the BSRI-12 in this older population, responses to the 12 items were subjected to an exploratory factor analysis using ones as prior communality estimates. Since factor analysis procedures are very sensitive to missing data we imputed missing items by substituting medians of items' scores. Seven items had no or 1–2 missing data, four items had between 10 and 17 missing, and the item 'has leadership abilities' had the highest number of missing observations equal to 23.

The principal axis method was used to extract the components, and this was followed by an orthogonal rotation. The reason for choosing orthogonal solution over oblique solution was that the two extracted gender factors in prior research (Carver et al., 2013) and also in our data were not correlated. Kaiser Criterion, which retains factors with eigenvalues greater than 1 (Kaiser, 1960) and Cattell's (1966) screen test, which involves an examination of a plot of the eigenvalues for breaks or discontinuities were employed to determine the number of factors to retain. In interpreting the factor structure, an item was said to load onto a given factor if the factor loading was greater than or equal to 0.25. To establish a priori criteria, consistent with other studies (Carver et al., 2013; Hatcher, 1994) if the differences between loading on two or more factors were smaller than 0.10, the item was excluded from analysis. The internal consistency measure of Cronbach coefficient alpha (Cronbach, 1951) was computed for each of the retained factors to examine the degree to which items included in each factor are all measuring the same underlying concept. The robustness of the exploratory factor analysis was assessed by estimating diagnostic measure of sampling adequacy (the Kaiser–Meyer–Olkin measure) and by Bartlett's test of sphericity (to determine whether the correlations between the variables, examined simultaneously, do not differ significantly from zero, in other words to evaluate if two retained factors together are sufficient to explain correlations between items). After extracting items loading onto feminine and masculine perceived gender roles, two composite scores equal to sum of scores of the retained items were constructed representing *femininity* and *masculinity* scores for each individual.

2.3.3. Three approaches of constructing gender role groups based on Bem androgyny model

Using the Bem androgyny model as the conceptual base and following traditional approaches of separation of individual according to their gender groups (Shifren & Bauserman, 1996), we adopted the median split method (Lenney, 1991) to classify the study population into four gender role groups. First, the median of the masculine and the feminine scales was established for both. Then individual femininity and masculinity scores were compared to the median. If the individual's score was below the median on both the feminine and masculine scales, he/she was classified as 'undifferentiated'. If the individual's scores on both the masculine and feminine scales were equal to or above the median, that individual was classified as 'androgynous'. Those people who were equal to or higher than the median on the feminine scale and lower

Please cite this article in press as: Vafaei, A., et al., The validity of the 12-item Bem Sex Role Inventory in older Spanish population: An examination of the androgyny model. Arch. Gerontol. Geriatr. (2014), http://dx.doi.org/10.1016/j.archger.2014.05.012

on the masculine scale were classified as 'feminine'. Finally, those who were equal to or higher than the median on the masculine scale and lower on the feminine scale were classified as masculine. This method has been used in similar studies of older populations in Canada (Gale-Ross et al., 2009) and Brazil (Carver et al., 2013).

We measured androgyny scores by two additional standard methods. Following classical work of Bryan et al. (1981), we defined the androgyny score as the geometric mean of asculinity and femininity scores. Our last measurement method was based on the original definition of androgyny by Bem (1974) which expresses a subject's androgyny score as the difference between femininity and masculinity scores divided by a term reflecting the variance of the two sets of femininity and masculinity ratings. In fact it is a student's t score, also called the androgyny t-ratio, in which scores near zero reflect androgyny orientation. For each individual, t-ratio androgyny was calculated using these formulas:

$$t \text{ score for androgyny} = \frac{\bar{F} - \bar{M}}{\sqrt{\left(\frac{Sp2}{nF}\right) + \left(\frac{Sp2}{nM}\right)}}$$

where

$$Sp^2 = \frac{bF \cdot SF^2 + nM \cdot SM^2}{nF + nM - 2}$$

This approach will label two groups with high androgyny: those with high scores in both femininity and masculinity and those with low scores in both femininity and masculinity and therefore is unable to separate androgynous from undifferentiated individuals.

Obtained scores by the two last methods were categorized into two gender role groups of 'high androgyny' and 'low androgyny' after splitting on the median of the distribution.

2.3.4. Examining the association between gender roles and health status

The frequencies of health status indicators across two high and low androgyny groups (constructed based on median split method) were compared by chi-square tests for contingency tables. To evaluate the consistency between different conceptualization of gender roles and measures of androgyny, analyses were repeated with the four traditionally defined gender roles groups by Bem (1974).

3. Results

3.1. The sample characteristics

The study population consisted of 60 men and 60 women with an average age of 69.9 (SD: 3) with no significant age differences across sexes. Seventy three percent of the participants were married and 40% of men and 35% of women had at least some university education. The population was relatively well-off, 62% reported having adequate income, however, there were disparities across sexes; compared to 70% of men only 53% of women perceived their income as adequate ($p = 0.06$). The participants represented a healthy population, 69% reported 'good' or 'very good' health and only 19% suffered from depression, with no significant difference between men and women (Table 1).

3.2. Psychometric properties and the validity of the BSRI items

Given that the Kaiser–Meyer–Olkin (KMO) measure of sampling adequacy was greater than 0.60 (0.77), the size of the sample deemed adequate for a factor analysis (Cerny & Henry, 1977). Our predefined criteria guided retaining two factors which represents meaningful underlying measurable constructs for two feminine and masculine gender roles. The statistical significance of Bartlett's test suggested that there *were* a two-common factor solution and it was sufficient to explain the correlations. The factor analysis resulted in factor loadings presented in Table 2.

Table 1
Description of the socio-demographic and health status of the study population.

Variable (number, percentage)	Men (n, %)	Women (n, %)	Total (n, %)	P-value[a]
Age(mean, SD)	70 (2.9)	69.8 (3.0)	69.9 (2.97)	0.71
Marital status				
Single	8 (13.5)	4 (7)	12 (10)	0.51
Married	44 (73)	44 (73)	88 (73)	
Divorced/widowed	8 (13.5)	12 (20)	20 (17)	
Education				
University	24 (40)	21 (35)	45 (38)	0.69
Less than university	36 (60)	39 (65)	75 (62)	
Adequacy of income				
Adequate	42 (70)	32 (53)	74 (62)	0.06
Health status SRH				
Good-very good	44 (73)	39 (65)	83 (69)	0.32
Depression	8 (13)	15 (25)	23 (19)	0.10
No difficulty climbing stairs	45 (75)	40 (67)	85 (71)	0.32
No difficulty walking 1 km	46 (77)	39 (65)	85 (71)	0.16

[a] From t-test and Chi-square tests comparing distribution between men and women

The six items loaded onto factor 1 were: warm; affectionate; tender; gentle; sensitive to others' needs; and sympathetic. All were among feminine items in the original BSRI and showed high internal consistency (Cronbach alpha = 0.89). The remaining six items: acts like a leader; leadership abilities; strong personality; dominant; makes decisions easily; and defends own beliefs were all masculine with relatively high internal consistency (Cronbach alpha = 0.74).

3.3. Relationships between gender roles and biological sex

Means of androgyny scores measured by geometric mean was identical across sexes (Table 3). When androgyny was measured by t-ratio, women showed higher mean scores (16.2 vs. 13.8), however differences were not statistically significant ($p = 0.22$). After categorizing the population into two groups of high and low androgyny, using either geometric mean or t-ratio method almost

Table 2
Factor analysis pattern matrix.

Loading			
Factor 1 (feminine role)		Factor 2 (masculine role)	
Warm	0.92	Acts like a leader	0.74
Affectionate	0.86	Has leadership abilities	0.71
Tender	0.84	Strong personality	0.53
Gentle	0.70	Dominant	0.51
Sensitive to others' needs	0.68	Makes decisions easily	0.33
Sympathetic	0.56	Defends own beliefs	0.29
Cronbach coefficient alpha			
0.89		0.74	
Variance explained			
69%		31%	
KMO measure for sampling adequacy		0.77	
P-value of Barlett's test of sphericity		<.0001	
Inter-factors correlation		0.09	

Please cite this article in press as: Vafaei, A., et al., The validity of the 12-item Bem Sex Role Inventory in older Spanish population: An examination of the androgyny model. Arch. Gerontol. Geriatr. (2014), http://dx.doi.org/10.1016/j.archger.2014.05.012

Table 3
Associations between biological sex and gender roles.

Gender role	Males	Females	*P*-value[a]
Androgyny by geometric mean			
Mean (SD)	32.3 (4.28)	32.3 (4.30)	0.96
High androgyny, *n* (%)	29 (49)	30 (50)	0.86
Low androgyny, *n* (%)	31 (51)	30 (50)	
Androgyny by t ratio differences			
Mean (SD)	13.8 (11.2)	16.2 (10.9)	0.25
High androgyny, *n* (%)	31 (52)	29 (48)	0.72
Low androgyny, *n* (%)	29 (48)	31 (52)	
Four gender roles			
Undifferentiated, *n* (%)	13 (52)	12 (48)	0.95
Masculine, *n* (%)	17 (53)	15 (47)	
Feminine, *n* (%)	15 (47)	17 (53)	
Androgynous, *n* (%)	15 (48)	16 (52)	

[a] From *t*-test for comparing means and Chi-square for comparing frequencies

half the population were androgynous and half not androgynous with no differences between men and women. Classifying the participants into four traditional gender role groups also showed no significant association between gender roles and biological sex ($p = 0.37$).

3.4. Comparison between three different strategies of measuring androgyny

Different measurement strategies classified participants into gender roles differently. The similarity between geometric mean and *t*-ratio measurements of androgyny was that both classified masculine individuals as androgenic and participants with prominent feminine characteristics as non-androgenic (Table 4). As expected, the *t*-ratio measure was unable to distinguish between undifferentiated and androgenic individuals, 23% of people categorized as androgenic had been classified as undifferentiated by traditional four gender role groups classification (Table 4).

3.5. Relationships between gender roles and indicators of health

Overall, regardless of the gender role measurement method, androgynous older adults reported better mobility and physical health as they less frequency rated their health as poor and reported fewer difficulties in climbing stairs and walking. However, for these two indicators the differences were not significant except for mobility limitation when gender roles were classified into four traditional groups ($p = 0.003$). There were more inconsistencies between measurement methods for depression. Surprisingly, higher androgyny when measured by *t*-ratio was significantly associated with more depressive symptoms ($p = 0.011$). This can be explained by two other findings: compared to geometric mean, *t*-ratio classified more masculine individual as androgynous (Table 4) and in this population people with masculine gender role (measured by the traditional way) suffered from d more depressive symptoms (Table 5). When we categorized the population into four gender role groups, significantly fewer androgenic individuals were depressed compared to those who scored low on both femininity and masculinity scores (6% vs. 32%, $p = 0.003$). Although not significant ($p = 0.54$), consistent with the androgyny model, androgynous individuals reported fewer depression symptoms when androgyny was measured by the geometric mean.

4. Discussion

Our main aim for this study was to examine the Bem androgyny model in a Spanish community-dwelling older population. The Bem androgyny model suggests people who are both feminine and masculine (androgenic) are more adaptive and therefore more likely to have better mental and physical health. Acknowledgment of the fact that gender role is a determinant of health requires proper measurement. One issue of multi-item measurement tools is that their validity in context-specific (Streiner & Norman, 2008), thus the need for re-validation when are being used in new populations. The original 60 item BSRI and various shorter versions have been used mostly in younger western populations (Alain, 1987; Fontayne et al., 2000; Schneider-Düker & Kohler, 1988; Wilcox & Francis, 1997) with very few validation studies in older adult (Carver et al., 2013; Gale-Ross et al., 2009) or Spanish population (Mateo & Fernández, 1991). Our population had a distinctive culture and age group, therefore, we re-validated the BSRI and showed that consistent with similar studies in older adults (Carver et al., 2013; Gale-Ross et al., 2009) it remains a reliable and valid tool for measuring gender roles in Spanish older adults. We adopted three approaches for measurement of androgyny and despite some inconsistencies across these approaches, we were able to provide some support for the existing theory that implies androgenic people are healthier and those with undifferentiated gender role are the least healthy. Contrary to gender roles, biological sex was not associated with any of the health indicators, suggesting that gender roles have an effect on health that is independent of biological sex.

Measurement of gender roles by different standard approaches and comparing the results has exposed several issues in studying gender and health. It has been suggested that gender roles measurement and classification methods impact the gender role attributed to an individual (Lenney, 1991; Sedney, 1981). We showed that depending on the employed method people were classified into different groups (Table 4). Dichotomizing individuals into androgenic and non-androgenic orientations also resulted in two problems: (1) each method classifies masculine-dominant and feminine-dominant people as androgynous and not-androgynous differentially, (2) when measurement is solely based on differences between femininity and masculinity scores, people with undifferentiated gender roles will also be labeled as androgynous because both of their scores are close (thus androgen by definition) but low. It seems that traditional classification of

Table 4
Comparison between different androgyny measurement approaches.

Traditional four gender role group	Measurement by geometric mean		Measurement by *t* ratio		Total
	High androgyny (*n*, %)	Low androgyny (*n*, %)	High androgyny (*n*, %)	Low androgyny (*n*, %)	
Undifferentiated	0 (0)	25 (41)	14 (23)	11 (18)	25
Masculine	22 (37)	10 (16.5)	29 (48.5)	3 (5)	32
Feminine	6 (10)	26 (42.5)	0 (0)	32 (53.5)	32
Androgynous	31 (53)	0 (0)	17 (28.5)	14 (23.5)	31
Total	59	61	61	59	

Please cite this article in press as: Vafaei, A., et al., The validity of the 12-item Bem Sex Role Inventory in older Spanish population: An examination of the androgyny model. Arch. Gerontol. Geriatr. (2014), http://dx.doi.org/10.1016/j.archger.2014.05.012

Table 5
Frequency of indicators of physical health, mobility status, and mental health across androgyny groups and four gender role groups.

Gender role	Total (%)	Number (%) with poor SRH[a]	Number (%)mobility difficulty climbing/walking	Number (%) depressed
Androgyny by geometric mean				
High androgyny, *n* (%)	59 (49)	15 (25)	16 (27)	10 (17)
Low androgyny, *n* (%)	61 (51)	22 (36)	26 (43)	13 (21)
Chi-square *P*-value		0.21	0.075	0.54
Androgyny by t ratio differences				
High androgyny, *n* (%)	60 (50)	15 (25)	17 (28)	17 (28)
Low androgyny, *n* (%)	60 (50)	22 (36)	25 (42)	6 (10)
Chi-square *P*-value		0.17	0.13	0.01
Four gender role groups				
Undifferentiated	25 (21)	12 (48)	15 (60)	8 (32)
Masculine	32 (27)	8 (25)	8 (25)	11 (34)
Feminine	32 (27)	11 (34)	14 (44)	2 (6)
Androgynous	31 (26)	6 (19)	5 (16)	2 (6)
P-value[b]		0.11	0.003	0.002

[a] Self-rated health.
[b] From Chi-square tests comparing distributions of health indicators across four gender roles groups.

gender roles into four groups can better disentangle the particular health effects of each type of gender role.

Limitations of this study warrant comment. We used a convenient sample of older adults attending three primary care clinics and invited all those coming into the clinics to the study. Though it is not a random sample, but unrepresentativeness will not be a big issue in this study. There is a National Health System in Spain with universal accessibility to a family doctor in the neighborhood clinics which covers more than 97% of the population (López-Valcárcel, Librero, Sanfélix-Gimeno, & Peiró, 2011), therefore we can assume the participants as good representatives of community-dwelling older adults in Aragon. However, this sample only included seniors healthy enough to go to clinics and excludes those with advanced diseases and mobility issues which receive care through national home care programs. Relatively high missing data especially for masculine items is another limitation of this study. To make the factor analysis models more stable, we imputed for these missing data, but cannot rule out the possibility of unreliable models due to missing data. Finally, our study might have been underpowered for detection of a true effect of gender roles considering including a fairly healthy population with low frequency of occurrence of health outcomes.

5. Conclusion

In conclusion, the shorter 12-item version of BSRI is a valid scale for measurement of gender roles in Spanish older adults. Grouping people into four traditional gender role groups of 'undifferentiated', 'masculine', 'feminine', and 'androgynous' better reflects psychological androgyny and is more capable of differentiating the impact of various gender roles on health outcomes, and is suggested to be used in epidemiological studies. Consistent with most studies of the androgyny model, in our population of community-dwelling older adults, androgynous gender role indicated by having both masculine and feminine traits was associated with better physical and mental health.

Conflict of interest statement

None from all authors.

Acknowledgments

This study was partially funded by the International Mobility in Aging Study (IMIAS) project: Canadian Institutes of Health Research (CIHR), Institute of Aging, New Emerging Team; gender differences in immobility; reference number AAM 108751.

References

Alain, M. (1987). A french version of the Bem Sex-Role Inventory. *Psychological Reports, 61*(2), 673–674.
Bassoff, E. S., & Glass, G. V. (1982). The relationship between sex roles and mental health: A meta-analysis of twenty-six studies. *Counseling Psychologist, 10*, 105–112.
Behar, R., De la Barrera, M., & Michelotti, J. (2001). Gender identity and eating disorders. *Revista Médica de Chile, 129*(9), 1003–1011.
Bem, S. L. (1974). The measurement of psychological androgyny. *Journal of Consulting and Clinical Psychology, 42*, 155–162.
Bem, S. L. (1977). On the utility of alternative procedures for assessing psychological androgyny. *Journal of Consulting and Clinical Psychology, 45*(2), 196–205. (research support, US Gov't., Non-PHS).
Berzins, J. I., Welling, M. A., & Wetter, R. E. (1978). A new measure of psychological androgyny based on the personality research form. *Journal of Consulting and Clinical Psychology, 46*(1), 126.
Blakely, T. A., Lochner, K., & Kawachi, I. (2002). Metropolitan area income inequality and self-rated health—A multi-level study. *Social Science & Medicine, 54*(1), 65–77.
Blanchard-Fields, F., Suhrer-Roussel, L., & Hertzog, C. (1994). A confirmatory factor analysis of the Bem Sex Role Inventory: Old questions, new answers. *Sex Roles, 30*(5–6), 423–457.
Bledsoe, J. C. (1983). Factorial validity of the Bem Sex-Role Inventory. *Perceptual and Motor Skills, 56*(1), 55–58.
Bryan, L., Coleman, M., & Ganong, L. (1981). Geometric mean as a continuous measure of androgyny. *Psychological Reports, 48*(3), 691–694.
Campbell, T., Gillaspy, J. A., & Thompson, B. (1997). The factor structure of the Bem Sex-Role Inventory (BSRI): Confirmatory analysis of long and short forms. *Educational and Psychological Measurement, 57*(1), 118–124.
Carver, L. F., Vafaei, A., Guerra, R., Freire, A., & Phillips, S. P. (2013). Gender differences: Examination of the 12-item Bem Sex Role Inventory (BSRI-12) in an older Brazilian population. *PLoS ONE, 8*(10), e76356.
Cattell, R. B. (1966). The scree test for the number of factors. *Multivariate Behavioral Research, 1*(2), 245–276.
Cerny, B. A. K., & Henry, F. (1977). *A study of a measure of sampling adequacy for factor-analytic correlation matrices* (Vol. 12, pp. 43–47). Hillsdale, NJ: Lawrence Erlbaum.
Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika, 16*(3), 297–334.
Downey, A. (1984). The relationship of sex-role orientation to self-perceived health status in middle-aged males. *Sex Roles, 11*(3–4), 211–225. http://dx.doi.org/10.1007/BF00287515
Doyal, L. (2000). Gender equity in health: Debates and dilemmas. *Social Science & Medicine, 51*(6), 931–939.
Fernandez, J., & Coello, M. T. (2010). Do the BSRI and PAQ really measure masculinity and femininity? *Spanish Journal of Psychology, 13*(2), 1000–1009.
Fontayne, P., Sarrazin, P., & Famose, J.-P. (2000). The Bem Sex-Role Inventory: Validation of a short version for French teenagers. *European Review of Applied Psychology, Revue Européenne de Psychologie Appliquée50*(4), 405–416.
Gale-Ross, R., Baird, A., & Towson, S. (2009). Gender role, life satisfaction, and wellness: Androgyny in a southwestern Ontario sample. *Canadian Journal on Aging, 28*(2), 135–146.
Hatcher, L. (1994). *A step-by-step approach to using the SAS system for factor analysis and structural equation modeling*. Cary, NC: SAS Institute Inc.
Heilbrun, A. B. (1981). Gender differences in the functional linkage between androgyny, social cognition, and competence. *Journal of Personality and Social Psychology, 41*(6), 1106.
Hernandez, J. A. E. (2009). Reevaluating the Bem Sex-Role Inventory. *Estudos de Psicologia (Campinas), 26*(1), 73–83.
Hoffman, D. M., & Fidell, L. S. (1979). Characteristics of androgynous, undifferentiated, masculine, and feminine middle-class women. *Sex Roles, 5*(6), 765–781.
Hunt, K., Lewars, H., Emslie, C., & Batty, G. D. (2007). Decreased risk of death from coronary heart disease amongst men with higher 'femininity'scores: A general population cohort study. *International Journal of Epidemiology, 36*(3), 612–620.
Huselid, R. F., & Cooper, M. L. (1992). Gender roles as mediators of sex differences in adolescent alcohol use and abuse. *Journal of Health and Social Behavior, 33*, 348–362.
Idler, E. L., & Benyamini, Y. (1997). Self-rated health and mortality: A review of twenty-seven community studies. *Journal of Health and Social Behavior, 38*, 21–37.
Idler, E. L., & Kasl, S. V. (1995). Self-ratings of health: Do they also predict change in functional ability? *Journals of Gerontology Series B: Psychological Sciences and Social Sciences, 50*(6), S344–S353.
Johnson, J. L., Greaves, L., & Repta, R. (2007). *Better science with sex and gender: A primer for health research*. Vancouver: Women's Health Research Network.
Kaiser, H. F. (1960). The application of electronic computers to factor analysis. *Educational and Psychological Measurement, 20*, 141–151.
Kalin, R. (1979). Method for scoring androgyny as a continuous variable. *Psychological Reports, 44*(3c), 1205L 1206.

Please cite this article in press as: Vafaei, A., et al., The validity of the 12-item Bem Sex Role Inventory in older Spanish population: An examination of the androgyny model. Arch. Gerontol. Geriatr. (2014), http://dx.doi.org/10.1016/j.archger.2014.05.012

Katsurada, E., & Sugihara, Y. (1999). A preliminary validation of the Bem Sex Role Inventory in Japanese culture. *Journal of Cross-Cultural Psychology, 30*(5), 641–645.
Krames, L., England, R., & Flett, G. (1988). The role of masculinity and femininity in depression and social satisfaction in elderly females. *Sex Roles, 19*(11–12), 713–721. http://dx.doi.org/10.1007/BF00288987
Lee, C. (2002). *The psychology of men's health*. Buckingham: Open University Press.
Lefkowitz, E. S., & Zeldow, P. B. (2006). Masculinity and femininity predict optimal mental health: A belated test of the androgyny hypothesis. *Journal of Personality Assessment, 87*(1), 95–101.
Lenney, E. (1991). Sex role. The measurement of masculinity, femininity and androgyny. In L. S. Robinson, & Wrightsman (Eds.), *Measures of personality and social psychological attitudes* (Vol. 1, pp. 573–660). San Diego, CA: Academic Press.
López-Valcárcel, B. G., Librero, J., Sanfélix-Gimeno, G., & Peiró, S. (2011). Are prescribing doctors sensitive to the price that their patients have to pay in the Spanish National Health System? *BMC Health Services Research, 11*(1), 333.
Macdonald, J., Gibbs, C. D., & Oliffe, L. J. (2004). Promoting men's health. In R. Moodie & A. Hulme (Eds.), *Hands-on health promotion*. Melbourne: I.P. Communications.
Månsdotter, A., Lundin, A., Falkstedt, D., & Hemmingsson, T. (2009). The association between masculinity rank and mortality patterns: A prospective study based on the Swedish 1969 conscript cohort. *Journal of Epidemiology and Community Health, 63*(5), 408–413.
Mateo, M., & Fernández, J. (1991). La dimensionalidad de los conceptos de masculinidad y feminidad. *Investigaciones Psicológicas, 9*, 95–116.
Nagi, S. Z. (1976). An epidemiology of disability among adults in the United States. *The Milbank Memorial Fund Quarterly. Health and Society, 54*, 439–467.
Özkan, T., & Lajunen, T. (2005). Masculinity, femininity, and the Bem sex role inventory in Turkey. *Sex Roles, 52*(1-2), 103–110.
Radloff, L. S. (1977). The CES-D Scale: A Self-Report Depression Scale for research in the general population. *Applied Psychological Measurement, 1*(3), 385–401.
Robison, J., Gruman, C., Gaztambide, S., & Blank, K. (2002). Screening for depression in middle-aged and older Puerto Rican primary care patients. *Journals of Gerontology Series A: Biological Sciences and Medical Sciences, 57*(5), M308–M314.
Sanchez-Santos, M., Zunzunegui, M., Otero-Puime, A., Cañas, R., & Casado-Collado, A. (2011). Self-rated health and mortality risk in relation to gender and education: A time-dependent covariate analysis. *European Journal of Ageing, 8*(4), 281–289.
Schmitz (2010). Sex, gender, and the brain—Biological determinism versus sociocultural constructivism. In I. Klinge (Ed.), *Sex and gender in biomedicine: Theories, methodologies, results* (pp. 57–72). Göttingen: University of Goettingen.
Schneider-Düker, M., & Kohler, A. (1988). Assessment of sex roles: Results of a German version of the Bem Sex-Role Inventory. *Diagnostica, 34*(3), 256–270.
Sedney, M. (1981). Comments on median split procedures for scoring androgyny measures. *Sex Roles, 7*(2), 217–222.
Seymour-Smith, S., Wetherell, M., & Phoenix, A. (2002). 'My wife ordered me to come!': A discursive analysis of doctors' and nurses' accounts of men's use of general practitioners *Journal of Health Psychology, 7*(3), 253–267.
Shifren, K., & Bauserman, R. (1996). The relationship between instrumental and expressive traits, health behaviors, and perceived physical health. *Sex Roles, 34*(11-12), 841–864. http://dx.doi.org/10.1007/BF01544319
Sinnott, J. D., & Shifren, K. (2001). Gender and aging: Gender differences and gender roles. In J. E. Birren & K. W. Schaie (Eds.), *Handbook of the psychology of aging* (5th ed., pp. 454–476). San Diego: Academic Press.
Spence, J. T., Helmreich, R. L., & Stapp, J. (1974). *The personal attributes questionnaire: a measure of sex role stereotypes and masculinity-femininity*. Washington: Journal Supplement Abstract Service, American Psychological Association.
Strahan, R. F. (1981). Remarks on scoring androgyny as a single continuous variable. *Psychological Reports, 49*(3), 887–890.
Streiner, D. L., & Norman, G. R. (2008). *Health Measurement Scales: A practical guide to their development and use*. Oxford: OUP.
Subramanian, S., Kawachi, I., & Kennedy, B. P. (2001). Does the state you live in make a difference? Multilevel analysis of self-rated health in the US. *Social Science & Medicine, 53*(1), 9–19.
Sugihara, Y., & Katsurada, E. (2000). Gender-role personality traits in Japanese culture. *Psychology of Women Quarterly, 24*(4), 309–318.
Twenge, J. (1997). Changes in masculine and feminine traits over time: A meta-analysis. *Sex Roles, 36*, 305–326. http://dx.doi.org/10.1007/BF02766650
Vega, V. C. (2007). Adaptación Argentina de un inventario para medir identidad de rol de género1. *Revista Latinoamericana de Psicología, 39*(3).
Ward, C., & Sethi, R. R. (1986). Cross-cultural validation of the Bem Sex Role Inventory Malaysian and South Indian research. *Journal of Cross-Cultural Psychology, 17*(3), 300–314.
Weiten, W. (1997). *Psychology: Themes and variations, briefer version* (3rd ed.). Pacific Grove, CA: Brooks/Cole Publishing Company.
Wilcox, C., & Francis, L. J. (1997). Beyond gender stereotyping: Examining the validity of the Bem Sex-Role Inventory among 16- to 19-year old females in England. *Personality and Individual Differences, 23*(1), 9–13.
Wilson, D., McMaster, J., Greenspan, R., Mboyi, L., Ncube, T., & Sibanda, B. (1990). Cross-cultural validation of the Bem sex role inventory in Zimbabwe. *Personality and Individual Differences, 11*(7), 651–656.
Windle, M., & Sinnott, J. D. (1985). A psychometric study of the Bem Sex Role Inventory with an older adult sample. *Journal of Gerontology, 40*(3), 336–343.
Zhang, J., Norvilitis, J. M., & Jin, S. (2001). Measuring gender orientation with the Bem Sex Role Inventory in Chinese culture. *Sex Roles, 44*(3-4), 237–251.
Zimmermann, F., Sieverding, M., & Müller, S. M. (2011). Gender-related traits as predictors of alcohol use in male German and Spanish university students. *Sex Roles, 64*(5–6), 394–404.

Please cite this article in press as: Vafaei, A., et al., The validity of the 12-item Bem Sex Role Inventory in older Spanish population: An examination of the androgyny model. Arch. Gerontol. Geriatr. (2014), http://dx.doi.org/10.1016/j.archger.2014.05.012