Masculinity and Femininity: A multidimensional approach

Álvaro Menéndez-Aller[1*], Pelayo Montes-Álvarez[2], Álvaro Postigo[2], Covadonga González-Nuevo[2], Jaime García-Fernández[2], Marcelino Cuesta[2], and Eduardo García-Cueto[2]

1 Instituto de Investigación Sanitaria del Principado de Asturias, Oviedo (Spain)
2 Universidad de Oviedo, Asturias (Spain)

Título: Masculinidad y Feminidad: Una visión multidimensional.
Resumen: La investigación de la Masculinidad y la Feminidad posee una larga trayectoria. A pesar de ello, aún se continúa discutiendo sobre la naturaleza de estos conceptos. En el presente estudio, se presenta el desarrollo de la Escala de Roles de Género de Oviedo (ERGO). Se empleó una muestra de 612 participantes procedentes de la población general española ($M_{años}$ = 34.2; $DT_{años}$ = 15.9). Se estudió la dimensionalidad, los índices de discriminación, la fiabilidad y las evidencias de validez divergente y convergente del instrumento. Además, se estudiaron diferencias en rasgos generales (modelo Big Five) y específicos de personalidad en función del sexo, y se realizó un ANCOVA controlando las variables de Masculinidad y Feminidad. Se observó un buen ajuste a una estructura multidimensional de tres factores, con alfas de Cronbach indicando una fiabilidad buena (Socioemocional = .75; Comparación = .81; Agresividad = .77) y adecuadas evidencias de validez. Se observaron diferencias en función del sexo en varios rasgos de personalidad, pero, al controlar la Masculinidad y Feminidad, las diferencias desaparecieron. El ERGO es una prueba válida y fiable para el estudio de los roles de género. Se discute la implicación de una aproximación multidimensional de la Masculinidad y Feminidad.
Palabras clave: Masculinidad. Feminidad. Sexo. Roles de género. Big Five. BEPE. ERGO.

Abstract: The research on Masculinity and Femininity has a long history. Despite this, there is still discussion about the nature of these concepts. In the present study, the development of the Oviedo Gender Roles Scale (GRSO) is presented. A sample of 612 participants belonging to general Spanish population is used (M_{years} = 34.2; SD_{years} = 15.9). The dimensionality, discrimination indices, reliability and evidence of convergent and divergent validity of the instrument were studied. In addition, differences in general (Big Five model) and specific personality traits based on sex were measured and an ANCOVA was performed controlling the variables of Masculinity and Femininity. A good fit to a three-factor multidimensional structure, with Cronbach's alphas indicating good reliability (Socioemotional = .75; Comparison = .81; Aggressiveness = .77) and adequate evidence of validity were observed. Differences based on sex were observed in various personality traits, but when controlling for Masculinity and Femininity, such differences disappeared. ERGO is a reliable and valid test for the study of gender roles. The implication of a multidimensional approach on Masculinity and Femininity is discussed.
Keywords: Masculinity. Femininity. Sex. Gender Roles. Big Five. BEPE. ERGO.

Introduction

In 1955, John Money laid the basis for the distinction between sex and gender in his research with intersex people (Money et al., 1955). Since then, it has become one of the most widely-used and influential topics in scientific discussion; one example of which can be seen in the progressive replacement of the term sex by the term gender, fundamentally in the humanities and social sciences (Haig, 2004). Currently, the *American Psychological Association* (APA) defines sex and gender as two different variables (VandenBos, 2015). In this definition, sex refers to the physical and biological traits that differentiate men and women. In contrast, gender refers to behavioral, social, and cultural aspects that are considered as belonging to men and women, in other words, Masculinity and Femininity.

Research into Masculinity and Femininity has also changed over time. In the early stages, the two variables were placed at either end of a single bipolar dimension (Fernández et al., 2007; López-Sáez & García-Dauder, 2020); a low score in Masculinity would mean a high score in Femininity and vice versa. The instruments used to measure this dichotomy employed items where statistically significant differences had

been previously found between men and women; for example, liking baseball was a predominant item for measuring Masculinity as statistically significant differences had always been found, with men scoring higher (Fernández, 2011; López-Sáez & García-Dauder, 2020). During the 20th century, various scientific studies found that the bipolar model was unable to satisfactorily explain the complexity of Masculinity and Femininity (Fernández, 2011). However, it was not until the work by Constantinople (1973) that Masculinity and Femininity began to be considered as two independent dimensions (Mateo & Fernández, 1991; López-Sáez & García-Dauder, 2020). From that point on, instruments to measure Masculinity and Femininity began to evaluate the extent to which a person behaved in accordance with the prevailing masculine and feminine gender roles in a given culture (Ward, 2000; López-Sáez & García-Dauder, 2020).

Despite that, in the last twenty years, the bidimensional concept of Masculinity and Femininity has been criticized, particularly in relation to the low percentage of variance it explains—under 50% (Fernández, 2011). Choi and Fuqua (2003) reviewed 23 studies related to the explanatory analysis of the factorial structure of the *Bem Sex Role Inventory* (Bem, 1974), one of the most widely-used instruments for studying gender roles. They found that, in most of the studies, the test demonstrated better fit to a multidimensional structure, which explained more than 50% of the variance. This structure is generally composed of one factor related to Femininity and two or more factors related to Masculinity. The au-

thors of those studies indicate that Masculinity and Femininity may be multifaceted constructs, such that a multidimensional conception would allow the study of the dimensions making up the two variables. This conclusion has been supported by subsequent research (Choi et al., 2006; Fernández et al., 2007).

In parallel to the aforementioned changes in the study into gender roles, it has also been found that they vary according to the socio-cultural context at the time (García-Cueto et al., 2015; López-Sáez & García-Dauder, 2020). A meta-analysis by Moya and Moya-Garófano (2021) examined the development of gender roles in Spain comparing social views of what was ascribed to men and to women in 1985 to responses from 2018. They found a trend towards equality, but noted that they still found differences in terms of social perception of what was considered to belong to each sex. Those results were consistent with findings from Andrade (2016), who found statistically significant differences between men and women in their attitudes towards gender roles. She also found that these differences were smaller in adolescents and young adults than in older adults, in line with the trend towards equality noted by Moya and Moya-Garófano (2021). Both results indicate that, despite the changes in society, there are still differences between men and women with respect to behaviors and attitudes (Eagly et al., 2020).

Researchers continue to find differences in personality according to sex, specifically in the variables from the Big Five model (Costa & McCrae, 1985). Women have been found to score more highly in Neuroticism and Agreeableness, while men have been found to score more highly in Extraversion, Responsibility, and Openness to Experience (Furnham & Treglown, 2021; Pedrosa et al., 2010; Schmitt et al., 2017; Smith et al., 2019). Mac Giolla and Kajonius (2018) examined whether these differences in personality depended on the country's gender equality policies. They assessed a sample from 22 countries grouped according to the Global Gender Gap Index—an evaluation of gender equality in a country calculated using political, economic, health, and educational factors (World Economic Forum, 2020). The authors found that the higher the index, and therefore the greater the gender equality, the greater the differences in personality between men and women.

Differences between men and women have also been found in more specific personality traits. Postigo et al. (2021) examined the enterprising personality of 1170 subjects using the Battery for Evaluating Enterprising Personality (BEPE; Cuesta et al., 2018) and found statistically significant differences by sex in Stress Tolerance and Risk-taking, with men scoring higher than women. No differences were found between men and women in Internal Locus of Control or Achievement Motivation. Internal Locus of Control is defined as the extent to which a person believes that they are responsible for the consequences of their actions (Cuesta et al., 2018). Achievement Motivation is the desire to achieve ever-higher goals and objectives (Cuesta et al., 2018). Despite the authors not finding sex related differences in the

two variables, previous studies have found differences between men and women in similar concepts. Araújo et al. (2019) examined students' academic expectations in their first year of university, finding that women exhibited much more pessimistic expectations about their achievements than men. Ferradás et al. (2018) examined the differences between men and women with regard to self-handicapping (the tendency to believe that academic achievement will be hindered by external factors). They found that women, to a greater extent than men, thought that their future achievements would not depend on their effort. For this reason it is useful to continue looking at whether men and women exhibit differences in the variables of Internal Locus of Control and Achievement Motivation.

It is also important to examine whether sex-related differences are actually influenced by scores in Masculinity and Femininity. Pedrosa et al. (2010) examined this in Verbal Fluency, Spatial Orientation, Neuroticism, and Abstract Reasoning. They found differences by sex in the variables, with men scoring higher in Spatial Orientation and Abstract Reasoning, and women scoring higher in Verbal Fluency and Neuroticism. However, the differences disappeared once scores in Masculinity and Femininity were controlled for. According to the authors, gender roles may be encouraging sex-related differences in cognitive and personality variables.

As we have seen, the study of gender roles has changed over recent years. In addition, what is socially considered as 'belonging' to men and women has also changed dramatically over the last hundred years. This means there needs to be a new instrument to measure Masculinity and Femininity which looks not only at current gender roles, but also at their multidimensional nature. That instrument will also allow us to understand whether some of the differences found between men and women can be explained by scores in Masculinity and Femininity.

The main objective of this present study is to create a Masculinity and Femininity scale. Starting from this overall objective, other, more specific objectives will be pursued, such as examining the psychometric properties of the test (dimensionality, item discrimination indices, reliability indices), as well as examining evidence of divergent and convergent validity between test scales. Finally, the study examines the influence of Masculinity and Femininity on sex-related differences in personality variables: Agreeableness, Openness to Experience, Internal Locus of Control, Achievement Motivation, Stress Tolerance, and Risk-taking.

Method

Participants

The initial sample was made up of 711 participants, however 99 were removed as they did not meet all of the inclusion criteria (they failed to properly answer all of the attentional control questions in the questionnaire). This meant

that the final sample consisted of 612 participants (55.4% women) from all over Spain. The participants' ages ranged from 18 to 83 years old ($M = 34.2$; $SD = 15.9$).

Instruments

Oviedo Gender Roles Scale (ERGO)

This is a self-report created *ad hoc* for this study to evaluate Masculinity and Femininity. Masculinity is defined as the attitudes and behaviors belonging to men, Femininity is those belonging to women.

The following steps were followed to create the instrument. First, 30 people were asked individually to list behaviors and attitudes of men and of women, as understood in society. The participants were from the general Spanish population, without any requirements for their participation other than being over 18 years old. This produced a list of 41 behaviors and attitudes belonging to women, and 36 belonging to men. The items were then reworded so that they made no explicit reference to sex, only the behaviors and attitudes themselves. For example, "Women remove body hair" was changed to "I remove body hair". The next step was to create a survey to assess whether each of the behaviors and attitudes were really considered as belonging to men (Masculinity) or women (Femininity) or were neutral. This survey was completed by 128 qualified psychologists. The items that were assessed as neutral (where more than half of respondents agreed) were eliminated, as were the items which were not considered to be only belonging to men or to women by at least 75% of the respondents. This produced 48 items, 23 for Masculinity and 25 for Femininity. Subsequently, a team of three experts in psychometry reviewed the two subscales, following the recommendations for constructing gender role scales proposed in Baber and Tucker (2006). Based on those recommendations, one item in the Masculinity scale and three items in the Femininity scale did not use completely neutral language and were removed. This meant that the final scale comprised 44 Likert-type items with five response options, where 1 means "completely disagree" with the statement in the item, and 5 means "completely agree".

Agreeableness

The subscale of the NEO Five Factor Inventory (NEO-FFI; Costa & McCrae, 1985), specifically, the Spanish adaptation by Cordero et al. (2008). This is defined as the qualities of a person related to altruism, sympathy, and care of others. The scale is made up of 12 Likert-type items ($α = .83$ in the manual) with five response options, where 1 means "completely disagree" with the statement in the item, and 5 means "completely agree".

Openness to experience

The subscale of the NEO Five Factor Inventory (NEO-FFI; Costa & McCrae, 1985), specifically, the Spanish adaptation by Cordero et al. (2008). This refers to the characteristics of a person who is interested in the new, the original, and the artistic. The scale is made up of 12 Likert-type items ($α = .83$ in the manual) with five response options, where 1 means "completely disagree" with the statement in the item, and 5 means "completely agree".

Internal locus of control

Subscale from the Battery for the Evaluation of Enterprising Personality (BEPE) (Cuesta et al., 2018). It refers to the causal attribution that one is responsible for the consequences of one's own behavior. The scale is made up of 10 Likert-type items ($α = .85$ in the original article) with five response options, where 1 means "completely disagree" with the statement in the item, and 5 means "completely agree".

Achievement motivation

Subscale from the Battery for the Evaluation of Enterprising Personality (BEPE) (Cuesta et al., 2018). It is defined as the desire to achieve excellence, i.e., achieve and improve objectives. The scale is made up of 10 Likert-type items ($α = .86$ in the original article) with five response options, where 1 means "completely disagree" with the statement in the item, and 5 means "completely agree".

Stress tolerance

Subscale from the Battery for the Evaluation of Enterprising Personality (BEPE) (Cuesta et al., 2018). It refers to resistance to perceiving external stimuli as stressors thanks to proper use of coping strategies. The scale is made up of 10 Likert-type items ($α = .84$ in the original article) with five response options, where 1 means "completely disagree" with the statement in the item, and 5 means "completely agree".

Risk-taking

Subscale from the Battery for the Evaluation of Enterprising Personality (BEPE) (Cuesta et al., 2018). It refers to a person's tendency and predisposition to accepting a certain amount of insecurity that will allow them to achieve a goal that has more benefits than negative consequences. The scale is made up of 10 Likert-type items ($α = .87$ in the original article) with five response options, where 1 means "completely disagree" with the statement in the item, and 5 means "completely agree".

Attention scale

A 9-item scale was used to assess how much attention subjects paid to the statements and instructions and to ensure that responses were not given randomly. Each item had five response options, from 1 to 5, and subjects were asked to choose a specific response ("Please mark option 1").

Procedure

The tests were applied using an online form. The items from the different questionnaires were randomly shuffled, together with the items from the attention control scale, with the condition that no two concurrent items would belong to the same dimension. The sample for the study was obtained through snowball sampling using various social networks.

Before completing the tests, the participants were asked to give their informed consent. Their anonymity was ensured and confidentiality was maintained in compliance with current data protection laws and professional standards (Official College of Psychologists of the Principality of Asturias, 2015; Organic Law 3/2018, of December 5, on the Protection of Personal Data and guarantee of digital rights).

Data analysis

In order to examine the dimensionality of ERGO, an exploratory factorial analysis was performed. Prior to that, the KMO index and Bartlett's statistic were calculated in order to confirm the suitability of the data for factorial analysis. Once that was confirmed, the factorial analysis was performed. An input matrix of polychoric correlations was used. The extraction method was robust weighted least squares with oblique rotation (Promin), as the dimensions of the test were understood to correlate with each other (Ferrando & Lorenzo-Seva, 2014). To determine the appropriate number of factors to extract, optimal implementation of parallel analysis was used (Timmerman & Lorenzo-Seva, 2011). Iteratively, items which demonstrated factor loading greater than 0.3 in more than one factor were eliminated, given that items with mixed loading do not contribute to the definition of a specific factor (Fabrigar & Wegener, 2012). Two indices of fit were used to confirm the fit to the data (Hu & Bentler, 1999): CFI, which should be greater than .900, and RMSR, which should be less than 0.080 in order for the fit to be considered good (Hoyle, 2012). Finally, the percentage of the total variance explained by the factors was calculated.

The discrimination index for the items in ERGO was calculated using the partial item-test correlation coefficient, removing the influence of the first. Any item with an index below 0.3 was removed (Hernández et al., 2016). The reliability of the instruments was estimated using Cronbach's alpha coefficient.

The AVE coefficient was used to assess evidence of divergent and convergent validity between the test subscales (Fornell & Larcker, 1981; Hair et al., 2009). When the AVE index of a scale has a value greater than 0.5, it can be considered evidence of convergent validity between scales. Also, when the squared correlation between two scales is less than the AVE indices of both scales, it can be taken as evidence of divergent validity between scales for both.

Differences in ERGO scores and personality variables as a function of sex was analyzed using successive ANOVAS, applying Bonferroni correction. Effect size was calculated using partial eta squared, with values between .010 and .039 considered small, .040 to .110 considered moderate, and between .111 and .200 considered large (Lenhard & Lenhard, 2016). ANOVAs were chosen rather than the Student t test in order to be able to compare the results with the subsequent analysis of covariance. To examine the influence of Masculinity and Femininity on those variables in which a sex-based difference was observed, successive ANCOVAs were performed applying Bonferroni correction, controlling for the scores of one or more of the ERGO scales.

The data were analyzed using IBM SPSS (Version 24) and FACTOR software (Version 10.10.02) (Lorenzo-Seva & Ferrando, 2020), using 95% confidence levels in each analysis.

Results

First, the suitability of the data for factorial analysis was checked. Both Bartlett's statistic and the KMO indicated that the data was suitable (Bartlett's statistic = $p < 0.01$; KMO = .79). The Parallel Analysis suggested a factor structure with three dimensions. On analyzing which items loaded on each of the factors, the structure was confirmed to be consistent with the multidimensional theory (Choi & Fuqua, 2003). A total of 28 items were removed iteratively (Fabrigar & Wegener, 2012) based on the criteria for item elimination (loading on more than one factor). The ERGO was reduced to 16 items, and the exploratory factorial analysis indicated a good fit to a three-dimensional structure (CFI = 0.961; RMSR = 0.041) (Hoyle, 2012), explaining 54.54% of the variance.

Table 1 presents the factor loadings. The items that load on the Socio-emotional factor refer to interpersonal relationships and emotional expression; in the Comparison factor, they refer to comparison between the sexes and the division of responsibilities in a couple; and in the Aggressiveness factor, they refer to enjoying verbal and visual aggression.

Table 1
Factor loadings of the ERGO's items

Item	Social-emotional	Comparison	Aggressiveness
Soy una persona delicada	.546		
Soy sensible	.807		
Siento que dependo emocionalmente de los demás	.487		
Me gusta cuidar de los demás	.498		
Me dejo llevar por mis emociones	.631		
Soy una persona romántica	.594		
Siento que debo ser yo quien invite en una relación		.745	
Siento que debo ser yo quien proteja a mi pareja		.705	
Habitualmente cedo el paso a las personas de otro sexo		.648	
Como más que las personas del otro sexo		.477	
Siento que debo ser yo quien de los primeros pasos en una relación		.678	
Si tuviera una pareja, (o con mi pareja actual) creo que debería ser yo quien se ocupe de hacer los trabajos que impliquen el uso de la fuerza		.754	
Suelo decir tacos			.893
Juego a videojuegos			.515
Disfruto observando contenidos violentos			.473
Utilizo insultos de forma amistosa			.960

It is worth noting that during the process of creating the scale, the items making up the Socio-emotional factor were mostly assessed by the experts as belonging to traditionally feminine roles (Femininity). The items making up the Comparison and Aggressiveness factors were mostly assessed as belonging to traditionally masculine roles (Masculinity).

Table 2 gives the correlation coefficients between factors. Note that the Comparison and Aggressiveness factors (related to Masculinity) correlate positively with each other and negatively with the Socio-emotional scale (related to Femininity).

Table 2
Interfactorial correlation matrix

	Social-emotional	Comparison	Aggressiveness
Social-emotional	1	-.220	-.284
Comparison		1	.456
Aggressiveness			1

In terms of discrimination, all of the items exhibited indices greater than .30, ranging from .630 to .705 (Socio-emotional), from .721 to .771 (Comparison), and from .571 to .703 (Aggressiveness).

Table 3
Cronbach's alpha of the used instruments

Instruments	Cronbach's alpha
Social-emotional	.75
Comparison	.81
Aggressiveness	.77
Agreeableness	.89
Openness to Experience	.89
Internal Locus of Control	.94
Achievement Motivation	.95
Stress Tolerance	.91
Risk-taking	.96

Table 3 shows the reliability coefficients for the ERGO subscales. All of the coefficients indicate good reliability (over .70) or excellent reliability (over .80) according to the revised standards from Hernández et al. (2016). The lowest alpha coefficient was for the Socio-emotional scale ($\alpha = .75$).

To assess divergent and convergent validity, the AVE coefficient (Fornell & Larcker, 1981; Hair et al., 2009) was calculated for the Socio-emotional (AVE = 0.36), Comparison (AVE = 0.43), and Aggressiveness (AVE = 0.48) scales. Although the values were close to 0.5, none were greater. The squared correlations between the ERGO scales were lower than the AVE indices for each scale, which indicates evidence in favor of divergent validity between the test scales.

To analyze the differences in ERGO scores and personality variables as a function of sex, successive ANOVAs were performed applying Bonferroni correction. There is considered to be a statistically significant difference when p is found to be less than .005.

In the ERGO subscales, statistically significant differences were found in the Socio-emotional scale ($p < .001$), with a moderate effect size ($\eta^2 = .108$), as well as in the Comparison ($p < .001$) and Aggressiveness ($p < .001$) scales, with large effect sizes ($\eta^2 = .324$ and $\eta^2 = .147$, respectively) (Lenhard & Lenhard, 2016). Looking at the means, women scored higher in the Socio-emotional scale, whereas men scored higher in the Comparison and Aggressiveness scales.

With regard to the personality variables, statistically significant differences were found in the scales of Stress Tolerance ($p < .001$), with a moderate effect size ($\eta^2 = .067$), along with the Agreeableness ($p < .001$) and Achievement Motivation ($p < .001$) scales, with small effect sizes ($\eta^2 = .035$ and $\eta^2 = .020$, respectively). Looking at the means, women scored higher in Agreeableness and Achievement Motivation, while men scored higher in Stress Tolerance.

The next step was to examine the influence of Masculinity and Femininity on the differences found in Agreeableness, Stress Tolerance, and Achievement Motivation as a function of sex. To do that, one or more of the ERGO scales were controlled for via successive ANCOVAs applying Bonferroni correction.

There is considered to be a statistically significant difference when *p* is found to be less than .016.

The results show that the differences between men and women in Agreeableness disappeared when the scores in Socio-emotional, Comparison, and Aggressiveness factors were controlled for together (*p* = .130). The sex-based differences in Achievement Motivation disappeared when the scores in the Socio-emotional and Aggressiveness scales were controlled for together (*p* = .023). The same occurred with Stress Tolerance (*p* = .460) once the scores from the Socio-emotional and Comparison scores were controlled for together.

Discussion

The distinction between sex and gender proposed following John Money's work with intersex people (López-Sáez & García-Dauder, 2020; Money et al., 1995) allowed a much deeper analysis of the differences between men and women. It was possible to see that differences which initially were thought to be a product of sex were due to socio-cultural factors, such as Masculinity and Femininity (Pedrosa et al., 2010). The study of these factors has also developed over time from that of a continuum to a bi-dimensional, and even multidimensional idea (Constantinople, 1973; Fernández et al., 2007; Fernández et al., 2014; Fernández, 2011).

The present study produced an instrument to measure the current prevailing gender roles and also examined whether these roles influenced differences in personality that previous researchers have found to be a function of sex.

In terms of psychometric properties, the factorial analysis indicates that ERGO has a good fit to a three-dimensional structure, made up of Socio-emotional, Comparison, and Aggressiveness scales. According to the results of the AVE index, there is suitable evidence of divergent validity between the scales, but there was no evidence of convergent validity. In any case, it is possible to conclude that ERGO is a reliable, valid instrument for the study of gender roles.

If we look at the content of the dimensions making up the test (see Table 1), the Socio-emotional factor can be defined as "behaviors and attitudes related to high emotiveness, warmth, and interpersonal empathy". The items refer to a set of qualities that are closely linked to traditionally feminine gender roles, emotional dependence, care of others, and emotional sensitivity (Ceballo-Fontes & Oramas, 2015). There were statistically significant differences in scores in this scale by sex, with women scoring higher, and a moderate effect size. Hence, the Socio-emotional factor can be considered to be related to Femininity.

The Comparison scale is defined as "social beliefs about the differences between the sexes". The content aligns, in large part, with gentlemanly behavior and so called benevolent sexism (Bria et al, 2020), which is related to a condescending posture towards women, seeing them as weaker and in need of help. The Aggressiveness scale covers "behavior and attitudes related to the enjoyment of violence and

verbal aggression". Both the Comparison and Aggressiveness scales deal with qualities related to traditionally masculine roles, a view of women as inferior and a tendency to enjoy violence (Baquerín, 2017). There were statistically significant differences, men scoring higher in both factors, with large effect sizes in both cases. Hence, both factors can be concluded to be related to Masculinity.

In short, ERGO is a multidimensional instrument that, in a similar way to the results from Choi and Fuqua (2003), has one factor related to Femininity and two factors related to Masculinity. The results of the present study agree with previous research that advocates for multidimensional study of Masculinity and Femininity (Choi et al., 2006; Fernández et al., 2014; Fernández, 2011).

Looking at the differences in personality variables as a function of sex, there were statistically significant differences in Stress Tolerance, with a moderate effect size, and in Achievement Motivation, with a small effect size. According to the mean scores, men scored higher in Stress Tolerance and women scored higher in Achievement Motivation. The sex-related differences in Stress Tolerance are in line with the findings from Postigo et al. (2021).

The present study also found statistically significant sex-related differences in Agreeableness, with a small effect size (η^2 = .035), but found no differences in Openness to Experience. This second finding differs from previous studies (Pedrosa et al., 2010; Schmitt et al., 2017; Smith et al., 2019).

Nonetheless, the sex-based differences disappeared once the scores in the ERGO scales were controlled for. When the scores in the Socio-emotional, Comparison, and Aggressiveness scales were controlled for, the sex-based differences in Agreeableness disappeared. Similarly, once the scores in Socio-emotional and Comparison scales were controlled for together, the sex-based differences in Stress Tolerances disappeared, and once the differences in Aggressiveness and Socio-Emotional scales were controlled for together, the sex-based differences in Achievement Motivation disappeared.

These results are similar to the findings from Pedrosa et al. (2010), as they show how the differences between men and women are largely controlled by scores in gender roles. The multidimensional approach allows us to understand which aspects of the two concepts, or to put it another way, which gender roles, may be influencing the difference.

Through ERGO, future studies may look more deeply into the influence of gender roles on a wide variety of psychological aspects, such as self-esteem and body image (Agam et al., 2015), ambition (Lopez-Zafra et al., 2021), or the acceptance of benevolent sexism in educational contexts (Bonilla-Algovia, 2021). In addition, future studies may investigate the extent to which Masculinity and Femininity influence variables where differences have been seen between men and women. It is important to underscore that ERGO is a first step in the investigation of the multiple facets making up Masculinity and Femininity, future research may also produce new instruments with more factors.

Finally, it is important to note the limitations of the present study. The main limitation is due to the imbalanced sample, as it was mainly young people and university students. It is also important to state that the study did not examine any concordance between participants' sex and gender identities, hence it is not possible to determine whether gender identity may have played an important role in the results.

Conclusions

The following conclusions in relation to the objectives established initially can be drawn from the study. ERGO demonstrated a good fit to a three-dimensional structure. The relia-

bility of the ERGO scales was good (α > .70 in all three scales). The discrimination indices of the ERGO scale items were over .3. The ERGO scales exhibited adequate evidence of divergent validity between scales but did not show evidence of convergent validity. When the scores in the ERGO scales were controlled for, the sex-based differences in Achievement Motivation, Stress Tolerance, and Agreeableness changed.

References

Agam, R., Tamir, S., y Golan, M. (2015). Gender differences in respect to self-esteem and body image as well as response to adolescent's school-based prevention programs. *Journal of Psychology & Clinical Psychiatry, 2*(5), 1-7. https://doi.org/10.15406/jpcpy.2015.02.00092

Andrade, C. (2016). Adaptation and factorial validation of the attitudes toward gender roles scale. *Paideía, 26*(63), 7-14. https://doi.org/10.1590/1982-43272663201602

Araújo, A. M., Assis-Gomes, C. M., Almeida, L. S., y Núñez, J. C. (2019). A latent profile analysis of first-year university students' academic expectations. *Annals of Psychology, 35*(1), 58-67. https://doi.org/10.6018/analesps.35.1.299351

Baber, K. M., y Tucker, C. J. (2006). The Social Roles Questionnaire: A new approach to measuring. *Sex Roles, 54*, 459–467. https://doi.org/10.1007/s11199-006-9018-y

Baquerín, I. S. (2017). *Masculinities, sexuality and gender* [Masculinidades, sexualidad y género] [Acta]. Congreso Internacional de la Red española de Filosofía. Madrid, España

Bem, S. L. (1974). The measurement of psychological androgyny. *Journal of Consulting and Clinical Psychology, 42*(2), 155–162. https://doi.org/10.1037/h0036215

Bonilla-Algovia, E. (2021). Acceptance of ambivalent sexism in trainee teachers in Spain and Latin America countries. *Annals of Psychology, 37*(2), 253-264. https://doi.org/10.6018/analesps.441791

Bria, P., Gómez Yepes, T., Etchezahar, E., y Ungaretti, J. (2020). Chivalry as a mediator between authoritarianism and gender roles. *Psicología, Conocimiento y Sociedad, 10*(3), 34-47. https://dx.doi.org/10.26864/pcs.v10.n3.3

Ceballo-Fontes, M., y Oramas, M. J. (2015). Traditional gender roles and type C personality in breast-cancer diagnosed women [Roles de género tradicionales y personalidad tipo c en mujeres con diagnóstico de cáncer de mama]. *Ciencia ergo sum, 23*(3), 229-238.

Choi, N., Fuqua, D.R., y Newman, J.L. (2006). Hierarchical confirmatory factor analysis of the Bem Sex Role Inventory. *Educational and Psychological Measurement, 67*, 818-832. https://doi.org/10.1177/0013164406299106

Choi, N., y Fuqua, D. R. (2003). The structure of the Bem Sex Role Inventory: A summary report of 23 validation studies. *Educational and Psychological Measurement, 63*(5), 872–887. https://doi.org/10.1177/0013164403258235

Constantinople, A. (1973). Masculinity-femininity: An exception to a famous dictum? *Psychological Bulletin, 80*(5), 389-407. https://doi.org/10.1177/0959-353505057611

Cordero, A., Pamos, A., y Seisdedos, N. (2008). *NEO PI-R, Personality Inventory NEO Revised* [NEO PI-R, Inventario de Personalidad NEO Revisado]. TEA Ediciones.

Costa, P. T., y McCrae, R. R. (1985). The NEO personality inventory. *Journal of Career Assessment, 3*(2), 123-139.

Cuesta, M., Suárez-Álvarez, J., Lozano, L., García-Cueto, E., y Muñiz, J. (2018). Assessment of eight entrepreneurial personality dimensions: Va-

lidity evidence of the BEPE battery. *Frontiers in Psychology, 9*, 1-10. https://doi.org/10.3389/fpsyg.2018.02352.

Eagly, A. H., Nater, C., Miller, D. I., Kaufmann, M., y Sczesny, S. (2020). Gender stereotypes have changed: A cross-temporal meta-analysis of U.S. public opinion polls from 1946 to 2018. *American Psychologist, 75*(3), 301-315. https://doi.org/10.1037/amp0000494

Fabrigar, L. R., & Wegener, D. T. (2012). *Exploratory factor analysis.* Oxford University Press.

Fernández, J. (2011). A century of masculinity and femininity research: A critical review [Un siglo de investigaciones sobre masculinidad y feminidad: Una revisión crítica]. *Psicothema, 23*(2), 167-172.

Fernández, J., Quiroga, A., Escorial, S., y Privado, J. (2014). Explicit and implicit assessment of gender roles. *Psicothema, 26*(2), 244-251. https://doi.org/10.7334/psicothema2013.219

Fernández, J., Quiroga, M. A., del Olmo, I., y Rodríguez, A. (2007). Masculinity and femininity scales: state of the art [Escalas de masculinidad y feminidad: Estado actual de la cuestión]. *Psicothema, 19*(3), 357-365.

Ferradás, M.-del-M., Freire, C., Rodríguez-Martínez, S., y Piñeiro-Aguín, I. (2018). Profiles of self-handicapping and self-esteem, and its relationship with achievement goals. *Annals of Psychology, 34*(3), 545-554. https://doi.org/10.6018/analesps.34.3.319781

Ferrando, P. J., y Lorenzo-Seva, U. (2014). Exploratory Item Factor Analysis: Some additional considerations. *Anales de Psicología, 30*(3), 1170-1175. https://doi.org/10.6018/analesps.30.3.199991

Fornell, C., y Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research 18*, 30–50. https://doi.org/10.2307/3151312

Furnham, A., y Treglown, L. (2021). Sex differences in personality scores on six scales: Many significant, but mostly small, differences. *Current Psychology*, 1-11. https://doi.org/10.1007/s12144-021-01675-x

García-Cueto, E., Rodríguez-Díaz, F. J., Bringas-Molleda, C., López-Cepero, J., Paíno-Quesada, S., y Rodríguez-Franco, L. (2015). Development of the gender role attitudes scale (GRAS) amongst young spanish people. *International Journal of Clinical and Health Psychology, 15*(1), 61-68. https://doi.org/10.1016/j.ijchp.2014.10.004

Haig, D. (2004). The inexorable rise of gender and the decline of sex: Social change in academic titles, 1945–2001. *Archives of Sexual Behavior, 33*(2), 87-86. https://doi.org/10.1023/B:ASEB.0000014323.56281.0d

Hair, J. F., Black, W. C., Babin, B. J., y Anderson, R. E. (2009). *Multivariate Data Analysis* (7th ed.). Prentice Hall.

Hernández, A., Ponsoda, V., Muñiz, J., Prieto, G., y Elosua, P. (2016). Review of the model to assess the quality of the tests used in Spain [Revisión del modelo para evaluar la calidad de los tests utilizados en España]. *Papeles del Psicologo, 37*(3), 192–197.

Hoyle, R. H. (Ed.). (2012). *Handbook of structural equation modeling.* Guilford press.

Hu, L.-t., y 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

Lenhard, W., y Lenhard, A. (2016). *Calculation of effect sizes*. Psychometrica. https://www.psychometrica.de/effect_size.html.

López-Sáez, M. A., y García-Dauder, D. (2020). Masculinity/femininity tests as psychological technologies for gender control [Los test de masculinidad/feminidad como tecnologías psicológicas de control de género]. *Athenea Digital*, *20*(2), e2521. https://doi.org/10.5565/rev/athenea.2521

Lopez-Zafra, E., Sánchez-Álvarez, N., y Carmona-Cobo, I. (2021). Is ambition a gendered issue? Students´ vs employees´ antecedents of Ambition about Leadership. *Annals of Psychology*, *37*(2), 352-360. https://doi.org/10.6018/analesps.435341

Lorenzo-Seva, U., y Ferrando, P. J. (2020). *FACTOR* (Versión 10.10.03) [Computer software]. Universitat Rovira i Virgili. https://psico.fcep.urv.es/utilitats/factor/Download.html

Mac Giolla, E., y Kajonius, P. J. (2018). Sex differences in personality are larger in gender equal countries: Replicating and extending a surprising finding. *International Journal of Psychology*, *54*(6), 705–711. https://doi.org/10.1002/ijop.12529

Mateo, M. A., y Fernández, J. (1991). The dimensionality of masculinity and femininity concepts [La dimensionalidad de los conceptos de masculinidad y feminidad]. *Investigaciones Psicológicas, 9*, 95-116.

Money, J., Hampson, J. G., y Hampson, J. L. (1955). An examination of some basic sexual concepts: The evidence of human hermaphroditism. *Bulletin of the Johns Hopkins Hospital*, *97*(4), 301–319.

Moya, M., y Moya-Garófano, A. (2021). Evolution of gender stereotypes in Spain: From 1985 to 2018. *Psicothema, 33*(1), 53-59. https://doi.org/10.7334/psicothema2020.328

Official College of Psychologists of the Principality of Asturias (2015). *Ethical Code of the Official College of Psychologists of the Principality of Asturias* [Código deontológico del Colegio Oficial de Psicólogos del Principado de Asturias] http://www.cop-asturias.org/@ASThiar_JUN/uploads/deonto/Codigo Deontologico-COPPA-2015.pdf

Organic Law 3/2018, of December 5, on the Protection of Personal Data and guarantee of digital rights. (2018). *Boletín Oficial del Estado*, 294, section I, December 6, 2018, 119788 to 119857. https://www.boe.es/eli/es/lo/2018/12/05/3/dof/spa/pdf

Pedrosa, I., Suárez-Álvarez, J., Pérez-Sánchez, B., y García-Cueto, E. (2010). Effect of sex and masculinity-femininity on some cognitive and personality aspects [Efecto del sexo y de la masculinidad-feminidad sobre algunos aspectos cognitivos y de personalidad]. Revista de psicología general y aplicada, 63(1-2), 23-32.

Postigo, Á., García-Cueto, E., Muñiz, J., González-Nuevo, C., y Cuesta, M. (2021). Measurement invariance of entrepreneurial personality in relation to sex, age, and self-employment. *Current Psychology*. https://doi.org/10.1007/s12144-021-01685-9

Schmitt, D. P., Long, A. E., McPhearson, A., O'brien, K., Remmert, B., y Shah, S. H. (2017). Personality and gender differences in global perspective. *International Journal of Psychology*, *52*, 45-56. http://dx.doi.org/10.1002/ijop.12265

Smith, M., Sherry, S., Vidovic, V., Saklofske, D., Stoeber, J., y Benoit, A. (2019). Perfectionism and the five-factor model of personality: A meta-analytic review. *Personality and Social Psychology Review*, *23*(4), 367–390. https://doi.org/10.1177/1088868318814973.

Timmerman, M. E., y Lorenzo-Seva, U. (2011). Dimensionality assessment of ordered polytomous items with parallel analysis. *Psychological Methods, 16*(2), 209–220. https://doi.org/10.1037/a0023353

VandenBos, G. R. (Ed.). (2015). *APA dictionary of psychology* (2nd ed.). American Psychological Association. https://doi.org/10.1037/14646-000

Ward, C. A. (2000). Models and measurements of psychological androgyny: A cross-cultural extension of theory and research. *Sex Roles: A Journal of Research*, *43*(8), 529–552. https://doi.org/10.1023/A:1007171500798

World Economic Forum. (2020). *Global Gender Gap* (Report 2020).