TOWARDS A MEASUREMENT OF GENDER

María Pilar SÁNCHEZ-LÓPEZ, Fátima SÁNCHEZ, Raquel RIVAS, Isabel CUELLAR-FLORES, Juan Francisco DÍAZ-MORALES

Abstract: *Gender is the result of an evolutionary process, a social category, which indicates the non-existence of traits or types of conduct which "a priori" might belong to one sex and by means of which social expectations and norms are interiorised. It also refers to the individual's psychological feeling of being a male or female, including the social and psychological modes of behaviour. The aim of this study is to adapt the gender questionnaires (CFNI-CMNI) to the Spanish population. If the psychometric results are acceptable, the intention is to continue expanding the sample of participants and evaluating their reliability and validity. The ultimate aim is to use the instruments in future researches on gender and health.*

Keywords: *Gender, measurment of gender, gender questionnaire*

Introduction

Our interest in psychometric evaluation of the sex-gender system ((Rubin, 1975) stems from the importance it has in explaining a wide range of manifestations: social conduct and personality (Cardenal and Fierro, 2001; Costa, Terracciano and McCrae, 2001; Hyde, 1984; Maccoby and Jacklin, (1974), activities and interests (Huston, 1983) and identities or self-perceptions (Chodorow, 1978; Fernández, 1996). The role played in health has also been pointed out (Mahalik, Lagan and Morrison, 2006; Martínez, 2005; Sánchez-López, 2003) with the relationship between both and the importance it has for preventing and treating problems (Brown, 1986; Gilbert and Scher, 1999; Brooks and Good, 2001).

Whereas sex is defined as the biological characteristics (chromosomic, gonadal, hormonal, cerebral and genital dimorphism, etc), gender is the result of an evolutionary process, a social category, which indicates the non-existence of traits or types of conduct which "a priori" might intrinsically belong to one sex and by means of which social expectations and norms are interiorised. It also refers to the individual's psychological feeling of being a male or female, including the social and psychological modes of behaviour designed by society as masculine or feminine (Barberá, 2004; Fernandez, 1996, 1998; Hare-Mustin and Maracek, 1990; Laqueur, 1990; Maccoby, 1990).

Thus, several authors of either sex defend the importance of the gender concept and insist on the need to look at those aspects related with it when explaining differences between men and women (Bem, 1981; Chodorow, 1978; Kagan, 1964). However, how can people's gender be evaluated? Its measurement has traditionally been dealt with by means of the evaluation of Masculinity/Femininity, conceived as the extent, more or less to which people close to the social stereotypes related to the behaviour of each of the sexes.

Gender-measuring instruments

In mid twentieth century there appeared several instruments for the measurement of masculinity and femininity, aiming to measure the psychological differences between men and women (Gough, 1952; Hataway and McKinley, 1943; Strong, 1936; Terman and Miles, 1936). Each of these scales sprang from different concepts, so comparing them is not easy. What can be affirmed is that their contents were very heterogeneous and what they were seeking was to discriminate between men and women with no explanatory theory as a base (Fernández, 1998).

Constantinople (1973) delivered a harsh criticism of the psychometric characteristics of these scales because they started out from the idea that masculinity and femininity are the two opposite poles of just one dimension. This author posits that they are necessary measurements of these two variables.

At the end of the twentieth century new scales of masculinity and femininity appeared, Bem's instrument, called *Bem Sex Role Inventory* (BSRI), was the most innovatory, since it was the first one to use independent measurements for these variables, where masculinity is not the opposite of femininity. This author proposes the concept of "androgyny" to designate subjects who present high scores in both categories, so that these persons would be shown to be assertive or flexible, instrumental or expressive depending on what the situation required.

Later, Spence, Helmreich and Stapp (1974, 1975) devised a questionnaire which measured masculinity and femininity as orthogonal dimensions. The test is known as PAQ *(Personal Attributes Questionnaire)* and showed once again that these variables are independent. Berzins, Willings and Wetter (1978) developed the ANDRO scale (*Personality Research Form)* which includes social and intellectual aspects on autonomy and orientation to achievement in the masculinity scale and contemplates aspects related to emotionality, subordination and upbringing in the femininity scale. Heilbrun's inventory (1976), *Adjective Check List* (ACL) also evaluates these dimensions through positive and negative adjectives. The scale produced by Balcom (1976) measures gender identity on two scales: MSC (masculinity) and FMN (Femininity). It is based on cultural stereotypes, not on personality traits as was the case in previous instruments. This explains the high correlation between Baucom's scale and Bem's inventory. Another measuring instrument for these variables is the *Sex Role Behaviour* (Orlofsky, 1981) which evaluates the interests and conduct of men and women in their roles.

All previous scales use as support the equivalent concepts of instrumentality-expressiveness (Parsons and Bales, 1955), agency-communion (Bakan, 1966) and self-assertive-integrative tendencies (Koestler, 1967, 1978) identifying masculinity and femininity with each of the terms of those independent dimensions. The resulting outline is

a quadruple typology (people can be androgynous, masculine, feminine or undifferentiated, depending on whether they score more or less on one of the two scales, in both or in neither), and by this means the aim was to evaluate that which is socially determined as desirable for each of the sexes (Fernández, 1998). Nevertheless, the criticisms made nowadays of these *new scales* centre precisely upon social desirability, since if masculinity and femininity are concepts conditioned by historical evolution and social construction, it does not seem apt to fix a framework as rigid as that of instrumentality-expressivity (Auster and Ohm, 2000; Cook, 1985, 1987; Fernández, Quiroga, del Olmo and Rodriguez, 2007; Marsh and Myers, 1986; Taylor and Hall, 1982; Woodhill and Samuels, 2003). As pointed out by Robinson, Shaver and Wrightsman (1991), gender, being a social concept, is sensitive to variations and changes that may take place in each society.

From all of this the need to have available a valid, reliable instrument which is representative of the present day can be deduced. Such an instrument can be incorporated into research and thus a more profound study can be carried out into its relationship with other variables. This is the purpose for which the instruments of Mahalik (2003, 2005), "*Conformity to Masculinity Norms Inventory*" (CMNI) and "*Conformity to Femininity Norms Inventory*" (CFNI) are designed. Both measurements offer a series of advantages with regard to their predecessors: they are updated instruments, since they have been created quite recently and incorporate a multidimensional perspective of gender. In this way they work with the constructs of "femininity" and "masculinity" not as homogeneous entities, but as multiples (there would be different "femininities" and "masculinities" with which people would identify us in different degrees). This latter is the other of the limitations of the global concepts used in the scales of the 70s, since both through its own theoretical definition and from the results shown by some research at the present time, they are multidimensional concepts (Fernández et al., 2007; García-Vega, Fernández García and Rico Fernández, 2005; Mahalik, Locke, Ludlow, Diemer, Scott, Gottfried and Freitas, 2003; Thompson and Pleck, 1995). Moreover, the instruments designed by Mahalik et al. (2003, 2005) use the term "femininity/masculinity norms" which, from our point of view, are quite adequate, insofar as they clarify that one is dealing with femininity and masculinity understood as a set of guidelines of conduct, attitudes and beliefs of a prescriptive nature and that in no measure do they reflect the natural psychological derivation of sexual dimorphism.

Preparation of the original instruments

Devising the original inventories followed a scrupulous empirical methodology, directed in a rational way. The instruments were put together at two different times, but in both cases the same phases were followed. The first step was to identify the dominant feminine and masculine norms in United States culture, that is, predominantly heterosexual and white. For this purpose the bibliography on traditional feminine and masculine norms was revised, and in the case of the CFNI a small group of women was chosen to create a discussion group in which they were asked to identify the messages they had perceived as to how women should think, feel and act, with 100 different messages received. Finally, in two mixed focus groups of male and female doctorate and Master's students there were discussions on the examples identified as masculine norms in one case and feminine in the other, categories were refined and items were devised.

The final categories were 12 feminine norms and 12 masculine norms, respectively. Each category included 12 items, reflecting a continuum of conformity which each norm could have; moreover, they described cognitive, behavioural and affective elements. 144 initial items were applied in constructing the CFNI to 733 females at the university, In the case of the CMNI, the 144 items were first submitted for judgement to three graduate students and acceptable indices of agreement were obtained. Subsequently, they were applied to 753 men, all of them university people. After applying the relevant factorial analyses the authors reached a solution of 8 factors in the CFNI and 11 in the CMNI, by means of main axes and oblique rotation.

In the CFNI items were selected with loadings higher than or equal to (.4) on one factor and with no loadings above (.3) in any other factor. This gave a matrix of 84 items. 7 of the 8 factors coincided with the rationally devised categories and only one of them combined two of them. As for the relationship between factors, only two pairs of them (Nice in relationships-Caring for children and Thinness-Invest in appearance) showed correlations equal to or more than (.3) .The internal consistency for the total scale is α =.88 and that for the subscale go from α =.77 the least to α =.92 the highest. External validity was checked by the Bem Sex role Inventory (BSRI- Bem, 1974, 1981) and the Feminist Identity Composite (FIC- Fischer et al., 2000). Significant, positive relations were found with the BSRI Femininity subscale and the FIC Passive Acceptance, and a negative one for the BSRI Masculinity subscale. Amongst the CFNI subscales not all had significant correlation with the other instruments, thus confirming the predictions of the authors insofar as the CFNI also evaluates other characteristics not included in the BSRI and FIC. As far as the time stability of the instrument is concerned, the test-retest coefficients in a 2-3 week period were .94 for total score and .87 average for the subscales (.83 the lowest for Domestic and.95 for Caring for children).

In the CMNI only 1 out of 55 pairs of correlations between factors was higher than (.3) (Power over women-Playboy). As in the CFNI, the items were chosen with loadings higher than or equal to (.4) in a factor and with no loadings higher than or equal to (.3) in any other factor, the result being a matrix with 94 items. Each factor was labelled with the same names as the rationally proposed categories. As far as their internal consistency is concerned, the total scale has an α =.94 and the subscales have values from α =.72 the least to α =.91 the highest. Time stability is also satisfactory in a period of 2-3 weeks, since the test-retest coefficients were .95 for the total scale and .76 on average for the subscales (with.51 being the lowest for Pursuit of Status and .96 the highest for Disdain for Homosexuals). The inventory was externally validated by other measures of masculinity, specifically with the Brannon Masculinity Scale Short Form (BMS- Brannon and Juni, 1984), the Gender Role Conflict Scale (GRCS- O'Neill et al., 1986) and the Masculine Gender Role Stress Scale (MGRS- Eisler and Skidmore, 1987). The total CMNI had a significant correlation with the total scores of the three scales, and most of the subscales designed to evaluate similar characteristics also presented significant relationships.

Aim

The aim of this study is to adapt the gender questionnaires (CFNI-CMNI) to the Spanish population. If the psychometric results are acceptable, the intention is to con-

tinue expanding the sample of participants and evaluating their reliability and validity. The ultimate aim is to use the instruments in future researches on gender and health.

Method

Participants

The CFNI was applied to 706 women from several Spanish regions, between 17 and 86 years of age (mean age = 32.62; D.T. =12.17) 75% of these participants have received higher education ,15% middle-school studies and 10% basic studies .What is more, 63.7% are working, 5% unemployed , 25% are students, 3.6% are retired and 1.7% work in the home.

The CMNI was completed by 213 men between the ages of 18 and 79 (mean age =33.62 D.T.12.45). The participants came from various Autonomous Spanish Communities. 64.7% of them had university studies, 24.2% secondary studies and 11% primary studies. Moreover, 68%had a job, 7.4% were out of work, 23.2% are students and 1.4% are retired.

Instruments

CFNI: Conformity to Femininity Norms inventory

The Conformity to Femininity Norms Inventory (CFNI) contains 84 items answered on a 4-point Likert scale (0 = completely disagree, 1 = disagreement, 2 = agreement, 3 = total agreement). The statements have been designed to measure attitudes, beliefs and behaviour associated with the feminine gender roles, both traditional and non-traditional. For example, statements deal with topics such as the aspect, how to look after others, sexuality and relationships. They are grouped into eight female prescriptions: Nice in relationships, Thinness, Modesty, Domestic, Caring for Children, Romantic relationships, Sexual Fidelity and Invest in Appearance. The questionnaire takes 10–15 minutes to apply. Participants are requested to indicate to what extent they are personally in agreement or disagreement with each statement.

CMNI: Conformity to Masculinity Norms Inventory

The Conformity to Masculinity Norms (CMNI) contains 94 items answered on a 4-point Likert scale (0 = completely disagree, 1 = disagreement, 2 = agreement, 3 = total agree). The statements have been devised to measure attitudes, beliefs and behaviour reflecting conformity or non-conformity with eleven messages associated with masculine gender roles: Winning, Emotional Control, Risk-Taking, Violence, Power over women, Dominant, Playboy, Self-Reliance, Primacy of work, Disdain for homosexuals and Pursuit of Status. The time needed to fill in the questionnaire and give instructions to the subjects is the same as for the CFNI.

Procedure

Adaptation of instruments to the Spanish population

To adapt the original inventory to the Spanish population an action protocol has been devised, taking on board the recommendations of international regulations (Guide-

lines of the ITC, International Test Commission) and national ones (model following the norms of test evaluations (CET), Test Commission of the Official College of Psychologists), the basic steps of which are the following:

1. Translation by a bilingual woman psychologist, with experience in translating psychological evaluation instruments, and with whom the group has previously worked on several occasions.

2. Application to a small group of psychologists who are experts in psychological evaluation and the adapting of instruments to Spanish reality. Evaluation is made of the linguistic, cultural and psychological adaptation of the instrument. Interjudge agreement enables the second experimental version to be devised.

3. Application to a small group from the ordinary population, asking them o indicate possible difficulties in both linguistic and cultural comprehension.

As a result of the process, small modifications were made in order to adapt the test to Spanish idiosyncrasy and culture. Finally our research team began the applications to a broad group of people from the Spanish territory.

With these data the next stage was the statistical analysis, for which the SPSS version statistical passage has been used .In calculating reliability as internal consistency Cronbach's alpha coefficient has been used.

Results

CFNI

The results of the reliability analysis with the CFNI indicate that the consistency index for the total scale is α =.85 (Table 1) Most of the subscales have acceptable α values, between .70 (Modesty) and .82 (Thinness and Sexual Fidelity), except for two subscales with values of .66 in Romantic relationships and.38 in Nice in relationships. The indices corresponding to the US population were higher than for all the subscales and for the total scale.

Table 1. Alpha reliability indices of CFNI subscales

	American Population [a] α	Spanish population A
1. Nice in relationships	.84	.38
2. Thinness	.90	.82
3. Modesty	.82	.70
4. Domestic	.84	.72
5. Caring for children	.92	.84
6. Romantic relationships	.77	.66
7. Sexual fidelity	.85	.82
8. Invest in appearance	.82	.74
TOTAL	.88	.85

a: Values extracted from Mahalik et al. (2005)

CMNI

The reliability index for the total CMNI is α = .92(Table 2). In the subscales the range of scores is from .70 (Dominant and Pursuit of Status) and .89 (Disdain for homosexuals), all of which are at acceptable levels. Only in the Risktaking subscale were the index values for internal consistency higher among the US population.

Table 2. Alpha reliability indices of CMNI subscales

Alpha reliability indices of CMNI subscales	American population[b] α	Spanish population α
1. Winning	.88	.81
2. Emotional Control	.91	.86
3. Risk-Taking	.82	.83
4. Violence	.84	.76
5. Power over women	.87	.75
6. Dominant	.73	.70
7. Playboy	.88	.85
8. Self-Reliance	.85	.81
9. Primacy of Work	.76	.72
10. Disdain for homosexuals	.90	.89
11. Pursuit of Status	.72	.70
TOTAL	.94	.92

b: Values extracted from Mahalik et al. (2003)

Discussion and conclusions

The process of adapting the two instruments to the Spanish population has, as we have already mentioned, followed the steps recommended by national and international norms. This enables an improvement to be made in the adaptation to our medium and ensures that the process carried out is the correct one.

When making a statistical check of the reliability of the scales through the Cronbach alpha for internal consistency ,which is one of the most used in psychometrics (Ledesma, Molina and Valero, 2002), we found that the scales´ values for the Spanish sample were very similar to the American sample, with minimum values of α =.7. Nevertheless, as is to be expected, when dealing with an adapted instrument, the indices are always lower in the Spanish samples than in the American ones, but the difference is small in most subscales, so we can claim that, in general, they present adequate results.

We have shown that the internal consistency of the CMNI is higher, in general, than that of the CFNI, probably because the former has a greater number of items (Ledesma, Molina and Valero, 2002). Thus, the values obtained in the CMNI do not pose any problem, all of them being higher than or equal to .7. However, in the CFNI the "Romantic relationships" and "Nice in relationships" scales present some weaknesses, particularly in the case of the latter. The "Romantic relationships" subscale presents the

lowest α index of all in the American population, an aspect which must be borne in mind in our adaptation. Despite all that, such a low level of internal consistency in "Nice in relationships" shows that the items in this subscale may be measuring several different aspects and, therefore, will be an object of future studies.

The values of the internal consistency indices presented in this work suggest that the adaptations carried out in the CFNI and CMNI may be good instruments for measuring conformity with male and female social norms in the Spanish population (once the "Nice in Relationships subscale has been corrected), and consequently, useful for evaluating the "gender) variable in our medium. These findings enable us to continue the study of the psychometric properties of both inventories and to extend the application sample.

Among the limitations of our study is the fact that the number of participants in the CMN is limited, an aspect which will be resolved as the process of adaptation continues. Another of the limitations is that in the CFNI the "Nice in relationships" scale shows some weaknesses and, therefore, will have to be studied in the future. Finally, Cronbach's alpha index is necessary but not sufficient to guarantee the psychometric properties and the validity of both instruments, so we will continue to study them.

Among future prospects is to be found the fact that if the two instruments continue to display their reliability and validity, the adaptation made for the Spanish population will allow us to use these instruments to evaluate the gender variable in our research, and apply the measurement to study relationships between health and gender. These constitute the final aim of our research group (EPSY, UCM).

K MĚŘENÍ GENDERU

Souhrn: Gender je výsledkem evolučního procesu, sociální kategorie, která označuje způsob chování, který by „a priori" měl patřit jednomu pohlaví, a prostřednictvím kterého jsou interiorizovány sociální očekávání a normy. Také označuje individuální psychologické cítětí být mužem nebo ženou, zahrnující sociální a psychologické způsoby chování. Cílem této studie je adaptovat gendrový dotazník (CFNI-CMNI) na španělskou populaie. Jestliže budou psychometrické výsledky akceptovány, naším záměrem bude rozšířit výzkumný soubor a evaluovat reliabilitu a validitu dotazníku. Konečným cílem je použít tyto nástroje v dalších výzkumech zaměřujících se na gender a zdraví.

Klíčová slova: Gender, diagnostika genderu, genderový dotazník