Is it advisable to include negative attributes to assess the stereotype content?
Yes, but only in the morality dimension.

Dr. Pablo Sayans-Jiménez[1]

Dr. Antonio José Rojas Tejada[2]

Dra. Isabel Cuadrado Guirado[2]

University of Almería, Spain[1,2]

Centre for the Study of Migrations and Intercultural Relations (CEMyRI)[2]

Word account : 4540

Abstract

Competence, morality and sociability dimensions have shown to be essential to measure stereotypes. Theoretically, the attributes associated with the negative pole of morality are more reliable and have shown to have higher evaluative weight. However, the current researches usually employ only positive attributes to measure each dimension. Since the advantages of the inclusion of negative morality are clear it would be interesting to know about the effects of the inclusion of such type of attributes (i.e., it is good or bad for the measurement). The purpose of this study is to examine if the addition of negative items makes possible to improve the stereotype content measures. This study compares the differences between scales with various compositions of positive and negative items of stereotypes to predict three related variables: anger, fear and a semantic differential of evaluation. The study was carried out with a sample of 550 Spaniards. The data found highlights the importance of using attributes of the negative pole of morality in studying stereotypes. Their use was able to explain the intergroup emotional responses and the semantic differential more efficiently.

Keywords: stereotypes, competence, morality, sociability, negative attributes

Is it advisable to include negative attributes to assess the stereotype content? Yes, but only in the morality dimension

Social objects, and specifically people, have attributes that enable a privileged perception of them compared to non-social objects (Fiske & Taylor, 2013; Ortiz, 1993). These features are referred to qualities shared by social objects and make possible their representation. The characteristics or attributes associated with a group or any of its members are called stereotypes (Stangor, 2009) and they can be accessed directly by self-report tests.

Many studies have emphasized the importance of certain dimensions in the perception of non-familiar social objects (e.g., people). These studies have shown evidence backing the use of attributes associated with the competence, morality and sociability dimensions (competence and warmth according to some models) for the study of intergroup relations in different countries and cultures (Brambilla & Leach, 2014; Brambilla, Rusconi, Sacchi, & Cherubini, 2011; Brambilla, Sacchi, Rusconi, Cherubini, &Yzerbyt, 2012; Fiske, Cuddy, Glick, & Xu, 2002; Fiske, Xu, Cuddy, & Glick, 1999; Goodwin, Piazza, & Rozin, 2014; Leach, Ellemers, & Barreto, 2007, López-Rodríguez, Cuadrado, & Navas, 2013). However, although the use of negative items in the morality dimension is empirically and theoretically recommended, as it will be detailed bellow, in the studies mentioned just before the use of negative items is neglected. For this reason, this research is aimed to explore if there are theoretical reasons that support or contravene the inclusion of negative items when the stereotype content is explored. Furthermore this work will examine if the addition of negative items should be avoided (as it is normally done) or if this addition makes possible to improve the stereotype content measures.

One of the most applied models when stereotype content is measured is the stereotype content model (SCM; Fiske et al., 1999, 2002). The SCM presents

competence (i.e., capability, competence, safety, skill) and warmth (i.e., friendship, integrity, sincerity, warmth) as universal, functionally relevant dimensions of social judgment (Cuddy, Fiske, & Glick, 2008; Cuddy et al., 2009). However, recent work has demonstrated that warmth dimension encompasses two distinct aspects: sociability (i.e., friendliness, sympathy) and morality (i.e., honesty, sincerity; Brambilla & Leach, 2014; Brambilla et al., 2012; Goodwin et al., 2014; Leach et al., 2007).

The competence dimension contains attributes referring to the target's ability in achieving goals (Cuddy et al., 2009; Fiske et al., 1999, 2002). Regarding morality dimension, Gray, Young, and Waytz (2012) show how "morality can be understood through the lens of interpersonal harm—the combination of intention and suffering" (p. 110). In order to make operative the definition, morality dimension could be delimited as the target's moral goals (i.e., positive or negative intentions) and the benefits or harms that such goals can cause people surrounding the target (including the observer). This definition includes concerns regarding morality psychology and person perception studies (see Haidt & Kesebir, 2010; Wojciszke, 2005). The items in the sociability dimension are attributes related to cooperation, reciprocity and developing relationships with others (Brambilla et al., 2011, 2012; Goodwin et al., 2014). These three dimensions may be considered the Universal Dimensions of Social Perception (UDSP).

Among these dimensions, the importance of morality is widely recognized (Brambilla et al., 2011, 2012; Brambilla, Hewstone, & Colucci, 2013; Brambilla & Leach, 2014; Goodwin et al., 2014, Gray et al., 2012; Leach et al., 2007). The information related to this dimension is perceived to be more reliable and objective (Goodwin & Darley, 2012). Van Bavel, Packer, Haas, and Cunningham

(2012) emphasize the importance of moral content in evaluation. When a person makes a moral evaluation, the perception of the target is influenced by "moral intuitions" which determines the evaluation process (Van Bavel et al., 2012). This effect facilitates that moral evaluations become faster, more extreme and more closely associated with universal norms (Van Bavel et al., 2012).

The UDSP offer diagnostic cues for impression formation. People use implicational schemas to relate specific points along an attribute continuum with the production of behaviors (Reeder & Brewer, 1979). Furthermore, the way in which the points along the continuum and the behavior are related can also be under the influence of the specific characteristics of the dimension of judgment being used, the observer's expectations and the context demands (Reeder & Brewer, 1979). In the case of morality and competence dimensions, they have different structures that entail certain degree of asymmetry between their poles. The attributes of the positive pole of competence and the negative attributes of the morality dimension have more diagnostic capability than their opposites (Skowronski & Carlston, 1987). Skowronski and Carlston (1987) affirm that differences between these two dimensions are due to external factors (i.e., not related to competence or morality) that are assumed as affecting behaviors in these domains. Social desirability or conformism could explain positive moral behaviors without involving high levels of morality. The opposite happens with competence; it results difficult to any person to perform successfully if he/she lacks of ability. However, the failure in the performance could be explained by mean of fatigue, lack of motivation or other external motives (Skowronski & Carlston, 1987). Regarding the valence of the attributes, Rothbart and Park (1986) demonstrated that negative traits are easy to confirm and difficult to disconfirm.

The opposite is true for positive traits. The negative pole of morality attracts more attention and has more evaluative weight than the positive pole (Goodwin & Darley, 2012). Attributes associated with badness in negative moral action (e.g., theft, cheating) are perceived as more reliable and objective than those refer to the goodness of positive moral actions (e.g., donate money; Goodwin & Darley, 2012). According to Brambilla and Leach (2014), "when people search for the most diagnostic information available about a person, they search for negative information about that person's morality" (p. 400). However, in spite of predictive advantages associated with the negative attributes of morality, their use is uncommon in current studies and models working with the basic dimensions of social categorization, using either morality or warmth dimension (e.g., Brambilla et al., 2011, 2012; Cuddy et al., 2009; Fiske et al., 1999, 2002; López-Rodríguez et al., 2013).

Due to the importance of negative attributes and the possible asymmetry between poles of each dimension, it would be interesting to find out whether by including negative attributes the measurement of stereotypes could be improved. The purpose of this study is to find out whether modification of the content of UDSP (including negative pole attributes) in the measurement of stereotypes shows empirical evidences backing their theoretical importance (i.e., according to the theory the UDSP scales should include, at least, positive competence items and negative morality items). The established hypothesis of this research is mainly focused in the improvement of the morality measurement by the addition of negative items. Evidences supporting the inclusion of negative items of morality are presupposed. Therefore, we expect that the models that include negative items of morality will show better psychometric properties that those that

do not include them. However it is difficult to assume the same hypothesis for competence and sociability items. On one side, the competence attributes stand out for their diagnostic capability when they are referred to positive traits (Reeder & Brewer, 1979; Skowronski & Carlston, 1987). Consequently, we cannot predict improvements in the measurement models when negative items of competence will be added. On the other side, regarding sociability dimension it is necessary to stop for a while in its definition. Sociability is related to cooperation, reciprocity and developing relationships with others. Therefore, the opposite pole should be referred to features of people that imply an impediment to the collaboration but without generating harms o damages toward others (i.e., if the attributes imply harms or damages they are then referred to the morality dimension). Items of the negative pole of sociability could generate an undesirable overlapping between morality and sociability. Furthermore they could activate representations not related exclusively with the cooperativeness of the social object. For these reasons it is difficult to expect any psychometrical improvement related to addition of the attributes of the negative pole of sociability. As a conclusion, we can anticipate that, according to the empirical and theoretical evidences, the best way to improve the measurement of the stereotype content is to add only negative items to the currently used measures. Finally, regarding the expected relations between UDSP and the related variables, we predict that positive UDSP will be positively correlated with SD and negatively with the anger and fear scales.

To carry out this study, negative items have been added to the usual set of positive UDSP attributes employed to measure stereotype content. Five structural models have been designed to choose the best content to measure the UDSP. First of all, the internal structure of each model will be tested, after that the capability

of the UDSP scales to predict global evaluations and intergroup Emotional
Reactions (ER) will be compared. The variables related to stereotype
measurement will be a scale measuring the evaluation dimension of the Semantic
Differential (SD) and two scales measuring intergroup ERs. The stereotypes
measured using the UDSP are determinant in generating outgroup evaluations.
Kervyn, Fiske, and Yzerbyt (2013) have demonstrated the direct relationship of
competence and warmth with the dimensions explored by the SD (evaluation,
potency and activity; Osgood, Suci, & Tannenbaum, 1957). Furthermore, the
content of the UDSP stereotypes will be related to intergroup ER and behavioral
tendencies. Specifically, the UDSP will be related with the intergroup emotions of
anger and fear because they are the main emotions that have been related to
threatening behaviors from the outgroup target (i.e., emotions will reflect if the
participants expect threatening behaviors from the outgroup target; see Mackie,
Devos and Smith, 2000).

Method

Sample

Five hundred fifty people, 280 women and 270 men, participated in this study.
Participants were selected using an accidental approach in the province of
Almeria, Spain. Sample selection was by quota sampling by age and sex in
keeping with the composition of the Spanish population (none of them belonged
to gypsy ethnic group). The sex quotas were 50.09% for women and the rest men.
The age intervals of the quotas were 35% for ages 18 to 35, 36% for 36 to 55, and
29% for 56 and over. The mean age of men was 45.97 years ($SD = 17.53$) and of
women 46.57 years ($SD = 17.97$). **Instruments**

All the instruments were designed for this study and applied in Spanish. Most of the items used have shown evidence of validity in different intergroup and intercultural contexts. All the items were congeneric for every UDSP and every related variable.

Stereotype scales with the UDSP. Five combinations of scales were designed for measuring stereotypes using the UDSP. All of them were built after the application of a pool of items including positive and negative items of the three UDSP. It is important to highlight that the structure and the definition of the UDSP determine the election of the negative items of these dimensions. The election of the negative items has been made looking for the opposite attributes in each dimension.

The first combination (Model 1) is made up of three stereotype scales, competence, sociability and morality; all of them include only positive items. The second one (Model 2) includes both positive and negative items of the three UDSP. The third one (Model 3) consists of three stereotype scales, competence, sociability and morality, but in this case the competence dimension contains both positive and negative items. Positive UDSP items and negative morality items were employed to the fourth combination of scales (Model 4). Finally, positive UDSP items and negative sociability items form the fifth combination (Model 5). There are not crossed loading between in the UDSP indicators and their factor loadings and uniqueness can vary freely. The attributes that make up each scale (see Table 1) are the result of an item selection process that studied the psychometric properties for: 1) the items that are used most often when UDSP are employed (e.g., Brambilla et al., 2011, 2012; Fiske et al., 1999, 2002; Leach et al.,

2007; López-Rodríguez et al., 2013), and 2) new items found by reviewing the literature and the UDSP definitions.

Each of the scales measured how non-gypsy people represent people of the gypsy ethnic group over the UDSP. Social identity of ingroup and outgroup was made salient tagging the person who answer the scales as "not gypsy". The combinations of stereotype scales contained the same instructions (see Appendix). The UDSP items were mixed and randomized. The item response scales used were seven-choice Likert-type (*none, almost none, little, half, many, almost all and all*, rated from 1 to 7). The higher the value of the answer corresponds the greater the perceived association between the target and the trait (i.e., morality, competence or sociability).

ER scales. Two scales were designed based on items used by Mackie et al. (2000). The item response scales used were seven-choice Likert-type (*none, almost none, little, half, many, almost all and all*, rated from 1 to 7). The instructions for the ER are similar to those for stereotypes and may be found in Appendix. The items of both scales were mixed and randomized.

Anger ER scale. This dimension is intended to measure how much people of the gypsy ethnic group are associated with an anger ER. These reactions usually appear when there is a relative absence and when the person perceives that he has strong collective support. It is comprised of the following items: Rage, fury, anger, irritation and frustration.

Fear ER scale. This dimension is intended to measure how much people of the gypsy ethnic group are associated with the fear ER. These reactions usually appear when there is a relative absence, but this time, when the person perceives

that he has weak collective support. It is comprised of the following items: Fear, panic and vulnerability.

Semantic Differential. A seven-item SD with a seven-point response scale validated in Spanish (Díaz-Guerrero & Salas, 1975) was used. All the items referred to the evaluation dimension. The pairs of adjectives used were: Sweet-Bitter, Transparent-Opaque, Light-Dark, Perfect-Imperfect, Whole-Broken, Tasty-Unpleasant, and Innocuous-Poisonous (the item scores were ranged from 1 to 7). The order and the direction of the items were randomized to control method effects (acquiescence and item wording effects—positive/negative). After the application, the items were recoded so that they could be interpreted more easily. Higher scores entail more positive evaluations. The instructions for the SD can be found in Appendix.

Procedure

One questionnaire with the whole pool of items was administered individually by staff trained for the purpose. The items were grouped and randomized after their own instructions (stereotypes, ER and SD). There was no time limit (approx. time necessary 20 to 30 minutes). The trained staff read out loud a text specifying that all the data would be handled anonymously and in a global and statistic way. This text also indicated that the participation was voluntary and that their collaboration could be stopped at any time. The trained staff confirmed that all the participants were over 18 and they were participating voluntarily. This procedure was approved by the Human Research Bioethical Committee of the University of Almeria.

Data analysis

A model with only positive items (the currently used) will be contrasted with a model including positive and negative attributes in each dimension. At the same time, these previous models will also be compared with three models with positive items in the three dimensions and only negative attributes in one each UDSP (competence, morality or sociability). The objective of these three last models is to test if the possible improvement of the predictive capability of the stereotype content is due to the inclusion of negative items in all UDSP or whether is due to the inclusion of negative item only in one of the three dimensions (as it is predicted by the theory).

The descriptive statistics of all the indicators were analyzed. Confirmatory Factor Analysis (CFA) and Structural Equation Models (SEM) approaches were used to test construct relationships, reliability and adequacy of the factorial structure of the five combinations of stereotype scales, the ER scales and the SD scale. The Maximum Likelihood (ML) method was used to estimate the parameters. The latent factors metric was assigned by fixing variance to 1.00 and mean to 0.00 in each UDSP and related variable. Cross-loadings were fixed to .00 and correlations between UDSP were estimated freely. The data set contained 472 complete cases. Full Information Maximum Likelihood was used to estimate missing values. Fit to the models was checked using the Chi-square test, the Tucker-Lewis Index (TLI), the Comparative Fit Index (CFI) and the Root Mean Square Error Approximation (RMSEA) and its 90% confidence interval (90% CI). A RMSEA of about .06 or less and TLI and CFI of about .95 or higher (Hu & Bentler, 1999) are indicative of adequate fit. A RMSEA equal or lower than .05 and TLI higher or equal to 0.97 indicate good fit (Schermelleh-Engel,

Moosbrugger, & Müller, 2003). Analyses were performed using SPSS v19.0 (Ibm Corp., 2010) and MPlus v7.0 (Muthén & Muthén, 1998-2011).

Predictive capability of the UDSP scales combinations. Five structural models were created to compare the predictive capability of each combination of UDSP scales (models 1 to 5). The SD and both ERs were regressed on each UDSP in each model (see Figure 1). Fit indicators, regression coefficients and coefficients of determination were compared among models in order to choose the best combination of UDSP scales. The description of the measurement model was made, *a posteriori*, only over the selected model in order not to provide redundant information.

—INSERT FIGURE 1 HERE—

Internal structure and reliability of the selected UDSP combination. Once the best combination of UDSP scales was chosen its internal structure was analyzed. Additionally, a cross-validation technique was performed between two randomly selected halves of the sample to test the measurement invariance of the UDSP scales (measurement model) and also the invariance of its predictive advantages (structural model). Strict invariance between the two halves of the sample was tested (i.e., equal form, factor loadings, intercepts and indicator residual variances). Furthermore, latent factor means were also fixed among subsamples when the strict invariance was tested. CFI difference test was performed to test the invariance between models (model names: measurement-free and structural-free in contrast to measurement-invariant and structural-invariant). To test the structural invariance the structural-free model was represented freeing correlations and regression coefficients between both subsamples, whereas the structural-invariant model equalized these parameters between both subsamples,

14

the rest of parameters remained fixed among both models. CFI differences lower than .01 indicates that the models have practically no significant differences in fit (Cheung & Rensvold, 2002). The reliability of the observed measures was estimated using composite reliability for congeneric measures (Raykov, 1997), omega. The reliability was estimated in the invariant model for each dimension in the UDSP scales and for each related variables. The 95% confidence intervals were estimated using MPlus (based on the application of the delta method; Raykov & Marcoulides, 2004). Omega coefficient of each scale reflects the reliable variance accounted by all the factors/constructs that underlie a scale score (McDonald, 1999).

Results

The descriptive statistics were calculated for all the items (see Table 1). No item showed extreme skewness or kurtosis, so parameters could be estimated by the maximum likelihood method (Brown, 2006).

—INSERT TABLE 1 HERE—

Predictive capability of the UDSP scales combinations

Fit statistics are shown in Table 2. The chi-squared test showed lack of fit with data in all models. According to other less restrictive indicators, models 1, 3 and 4 (i.e., only positive items, positive items and negative items of competence, and positive items and negative items of morality respectively) showed adequate fit indicators, whereas model 5 (i.e., positive items and negative items of sociability) showed acceptable/adequate fit indicators. The worst fit indicators were found in model 2 (i.e., positive and negative items in the three UDSP); according the CFI indicator this model should be discarded.

—INSERT TABLE 2 HERE—

Correlations, regression coefficients and coefficients of determination of each model can be seen in Table 3. The coefficients of determination of each variable in each model were compared to the same coefficients in the model 1 (using the fisher r-to-z) to check if the addition of negative items improved the predictive capability of the UDSP over the related variables. Models 2 and 4 showed an overall improvement of the predictive capability, however fit indicators of model 2 suggested that this model could not be appropriate to represent the UDSP. In consonance with the theoretical expectancies, the addition of negative items of morality (and not of competence or sociability) is the best way to improve the content of the UDSP scales.

—INSERT TABLE 3 HERE—

Internal structure and reliability of the selected UDSP combination

The model with all the positive items of the UDSP and the negative items of morality (model 4) was selected as the best theoretical and empirical option to assess the stereotype content of the UDSP. To test the stability of the estimated parameters, in the measurement model and in the structural model, the sample was randomly divided in two halves.

First of all, the invariance of the measurement model was contrasted (i.e., measurement-free in contrast to measurement-invariant models). Fit indicators of both models can be seen in Table 4. Both models, measurement-free and measurement-invariant, showed acceptable/adequate fit indicators and there were not CFI differences between models, therefore measurement invariance can be assumed. These results show the stability of the estimated parameters in two big

subsamples. Standardized factor loadings and correlations between UDSP can be seen in the representation of the structural model that is going to be analyzed below (see Figure 2). Competence factor loadings were moderate $M = .61$, range = [.65, .55]. Morality factor loadings were high, unsigned $M = .73$, unsigned range = [.65, .82]. Sociability factor loadings were high, with the exception of *open* indicator, $M = .64$, range = [.37, .76]. The reliability of the observed scores was estimated in the invariant model. Omega estimations [95% CI] were 70 [.66, .73] for competence dimension, .92 [.91, .93] for morality, and .72 [.68, .75] for sociability. Competence and sociability scales showed reliability estimations slightly below .80, the recommended cutoff point (Raykov, 1997).

—INSERT TABLE 4 HERE—

—INSERT FIGURE 2 HERE—

Once the invariance of the measurement model was confirmed the parameter's invariance of the structural model was tested. CFI differences were estimated between the structural-free model (freed correlations and regression coefficients between both subsamples) and the structural-invariant model (fixed correlations and regression coefficients between both subsamples). Fit indicators for both models can be seen in Table 4. Both models showed acceptable/adequate fit indicators. Since there were not CFI differences between models structural invariance can be assumed. Consequently it can be stated that the relationship between the UDSP in model 4 (including all the positive UDSP items and only the negative items of morality) and the related variables is invariant in both big subsamples. The coefficients of determination in the invariant model were .39 for anger, .29 for fear, and .73 for the SD (all $p < .001$). These results support the stability of the advantages of including negative morality items when the

stereotype content is measured. Furthermore standardized factor loadings of the UDSP in model 4 and of the related variables, correlations and regression coefficients can be seen in Figure 2. Omega estimations [95% CI] for the related variables were: .94 [.93, .94] for anger scale, .84 [.83, .86] for fear and .80 [.78, .82] for the SD of evaluation. The SD showed unfavorable reliability estimations of its scores, however this results do not affect the relations among the latent constructs (free of measurement error).

Discussion

In consonance with the importance of the negative pole of morality (Baumeister, Bratslavsky, & Finkenauer, 2001; Goodwin & Darley, 2012; Skowronski & Carlston, 1987), the results of this study showed the higher theoretical and empirical support for the inclusion of negative morality items in the UDSP scales. The model 4 showed the best combination of fit indicators and regression coefficients in contrast to the rest of the models, including the model 1 (the currently used). Furthermore, immorality indicators showed higher standardized factor loadings than morality indicators. Evidences of validity based on the relationship with other variables showed that the addition of negative items of morality allow to explain a bigger amount of variance than when only positive items are used. These results have shown to be stable between the two big randomly selected subsamples. This study is only a first step in the exploration of the advantages of using negative items of morality. Although the present research included a big sample it could be recommendable to explore our hypothesis with other outgroups.

Regardless, this study has proven the greater importance of the negative pole of morality to explore the stereotype content. The importance of this pole is due to both the properties of the schema used in the impression formation process (Reeder & Brewer, 1979; Skowronski & Carlston, 1987), and the higher psychological activation produced by the perception of negative stimuli (instead of positive; Cunningham, Raye, & Jonhson, 2004).

As it was expected, negative items of competence showed worst psychometric properties than positives. These results can be related to the properties of the schema used (Reeder & Brewer, 1979), but also with their higher stability. The positive attributes of competence demand more trait-inconsistent behavioral instances to be disconfirmed (Tausch, Kenworthy, & Hewstone, 2007). In this sense, the attributes of the sociability dimension have shown to work in a similar way that those referred to competence.

At this point it is necessary to clarify that the sociability attributes are empirically (but also theoretically) separated from the morality dimension. As it is shown in several researches sociability and morality attributes not only belong to different latent constructs (Brambilla & Leach, 2014; Brambilla et al., 2012; Goodwin et al., 2014; Leach et al., 2007), they play indeed different social functions in the social judgment process (Landy, Piazza, & Goodwin, 2016). Initially, in this research, the attributes of sociability were defined as attributes related to cooperation, reciprocity and developing relationships. However, it was not established if these attributes were representing intentions or abilities. For this reason it was impossible to determine, *a priori*, whether the diagnostic asymmetry of the sociability poles were going to be similar to competence or to morality. Fortunately, the recent work of Landy et al. (2016) contributes to clarify the

definition of the sociability dimension and provides an explanation that matches with the obtained results in the present study. According to these authors, the attributes of sociability, in the same way that the attributes of competence, provide information related to a skill of the observed person. Specifically, the dimension of sociability informs about the "ability to attract the social support, needed to carry out their intentions, whether those intentions are helpful or harmful" (Landy et al., 2006, p. 2). For this motive it stands to reason that competence and sociability dimensions work in a similar way. Therefore positive attributes of sociability would require more information to be disconfirmed in contrast to the negative attributes.

Finally, it is difficult to understand why immorality items are not employed more frequently to measure the stereotype content of morality, moreover when their inclusion is thoroughly supported (e.g., Brambilla & Leach, 2014; Goodwin & Darley, 2012; Rothbart & Park, 1986). It has to be said that the addition of negative items of morality can be taken as a double diagnostic advantage. On one hand, the fact that the most salient traits of the social object are negative provides information with greater evaluative weight than the positive items. On the other hand, it is important to keep in mind that when we apply test/scales we are dealing (and scaling) with the raters' behaviors. For this reason, the higher influence of social desirability should not be seen as an inconvenient; indeed it can be taken as a resource to make better distinctions between the people who evaluate the social object as moral and the people who evaluate it as immoral. The execution of negative behaviors, as answering in a non-social desirable manner, offers great diagnostic information. That is, the representation

of the social object should be negative enough to motivate the rater/s to answer in a socially undesirable way.

In addition, the importance of an adequate measurement of the morality dimension it is highly recommend when intergroup relations are studied. The importance of the morality dimension on its own when people form overall impressions has been highlighted in several studies (e.g., Brambilla & Leach, 2014; Brambilla et al., 2012; Goodwin et al., 2014; Landy et al., 2016; Leach et al., 2007). But now there also exists evidence of its moderating effect between the competence/sociability attributes and the formation of overall impressions (i.e., the contribution of the competence and sociability traits depends on a target's morality; Landy et al., 2016). Therefore, given the importance of the morality dimension its content should not be neglected. In conclusion, this article has presented results backing an improvement in the scales constructed with the items traditionally employed in the study of the stereotype content. The addition of items of the negative pole of morality provides more theoretical coherence and has been backed by empirical data when positive morality items are also presented.

References

Baumeister, R. F., Bratslavsky, E., Finkenauer, C., & Vohs, K. D. (2001). Bad is

    stronger than good. *Review of General Psychology, 5*, 323-370. doi:

    [10.1037/1089-2680.5.4.323](10.1037/1089-2680.5.4.323)

Brambilla, M., & Leach, C. W. (2014). On the importance of being moral: The

    distinctive role of morality in social judgment. *Social Cognition, 32*, 397-

    408. doi: 101521soco2014324397

Brambilla, M., Hewstone, M., & Colucci, F. P. (2013). Enhancing moral virtues:

    Increased perceived outgroup morality as a mediator of intergroup contact

    effects. *Group Processes & Intergroup Relations, 16*, 648-657. doi:

    [10.1177/1368430212471737](10.1177/1368430212471737)

Brambilla, M., Rusconi, P., Sacchi, S., & Cherubini, P. (2011). Looking for

    honesty: The primary role of morality (vs. sociability and competence) in

    information gathering. *European Journal of Social Psychology, 41*, 135-

    143. doi: 10.1002/ejsp.744

Brambilla, M., Sacchi, S., Rusconi, P., Cherubini, P., & Yzerbyt, V. Y. (2012).

    You Want to Give a Good Impression? Be Honest! Moral Traits Dominate

    Group Impression Formation. *British Journal of Social Psychology, 51*,

    149-166. [doi: 10.1111/j.2044-8309.2010.02011.x](doi: 10.1111/j.2044-8309.2010.02011.x)

Brown, T. A. (2006). *Confirmatory factor analysis for applied research*. New

    York, NY: Guilford Press.

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating Goodness-of-Fit Indexes

    for Testing Measurement Invariance. *Structural Equation Modeling*, *9*,

    233–255. doi: 10.1207/s15328007sem0902_5

Cuddy, A. J. C., Fiske, S. T., & Glick, P. (2008). Warmth and competence as

    universal dimensions of social perception: The Stereotype Content Model

    and the BIAS Map. In M. P. Zanna (Ed.), *Advances in Experimental Social*

    Psychology (pp. 61-149). San Diego, CA: Academic Press.

Cuddy, A. J. C., Fiske, S. T., Kwan, V. S. Y., Glick, P., Demoulin, S., Leyens, J.

    P., …. Ziegler, R. (2009). Stereotype content model across cultures:

    Towards universal similarities and some differences. *British Journal of*

    Social Psychology, *48,* 1-33. doi: 10.1348/014466608X314935

Cunningham, W. A., Raye, C. L., & Johnson, M. K. (2004). Implicit and explicit

    evaluation: FMRI correlates of valence, emotional intensity, and control in

    the processing of attitudes. *Journal of Cognitive Neuroscience, 16*, 1717-

    1729. doi: 10.1162/0898929042947919

Díaz-Guerrero, R., & Salas M. (1975). *El diferencial semántico del idioma*

    español [The semantic differential of Spanish]. México: Trillas.

Fiske, S. T., & Taylor, S. E. (2013). *Social cognition: From brains to culture*.

    London: Sage.

Fiske, S. T., Cuddy, A. J., Glick, P., & Xu, J. (2002). A model of (often mixed)

    stereotype content: Competence and warmth respectively follow from

    perceived status and competition. *Journal of Personality and Social*

    Psychology, *82*, 878-902. doi: 10.1037/0022-3514.82.6.878

Fiske, S. T., Xu, J., Cuddy, A. C., & Glick, P. (1999). (Dis)respecting versus

    (dis)liking: Status and interdependence predict ambivalent stereotypes of

    competence and warmth. *Journal of Social Issues*, *55*, 473-491. doi:

    10.1111/0022-4537.00128

Gray, K., Young, L., & Waytz, A. (2012). Mind Perception Is the Essence of

    Morality. *Psychological Inquiry*, *23*, 101–124.

    doi:10.1080/1047840x.2012.651387

Goodwin, G. P., & Darley, J. M. (2012). Why are some moral beliefs perceived to

    be more objective than others? *Journal of Experimental Social*

    Psychology, 48, 250-256. doi: 10.1016/j.jesp.2011.08.006

Goodwin, G. P., Piazza, J., & Rozin, P. (2014). Moral character predominates in

    person perception and evaluation. *Journal of Personality and Social*

    Psychology, 106, 148-168. doi: 10.1037/a0034726

Haidt, J., & Kesebir, S. (2010). Morality. In S. Fiske, D. Gilbert, & G. Lindzey

    (Eds.) *Handbook of Social Psychology* (5th ed., pp. 798-832)*. Hobeken,*

    NJ: Wiley.

IBM Corp. Released 2010. *IBM SPSS Statistics for Windows, Version 19.0*.

    Armonk, NY: IBM Corp.

Kervyn, N., Fiske, S. T., & Yzerbyt, V. Y. (2013). Integrating the stereotype

    content model (warmth and competence) and the Osgood semantic

    differential (evaluation, potency, and activity). *European Journal of Social*

    Psychology, 43, 673-681. doi: 10.1002/ejsp.1978

Landy, J. F., Piazza, J., & Goodwin, G. P. (2016). When Its Bad to Be Friendly

    and Smart: The Desirability of Sociability and Competence Depends on

    Morality. *Personality and Social Psychology Bulletin, 42*, 1272–1290.

    doi:10.1177/0146167216655984

Leach, C. W., Ellemers, N., & Barreto, M. (2007). Group virtue: The importance of morality (vs. competence and sociability) in the positive evaluation of in-groups. *Journal of Personality and Social Psychology, 93*, 234-249. doi: 10.1037/0022-3514.93.2.234

López-Rodríguez, L., Cuadrado, I., & Navas, M. (2013). Aplicación extendida del Modelo del Contenido de los Estereotipos (MCE) hacia tres grupos de inmigrantes en España [Extended application of the Stereotype Content Model (SCM) toward three immigrant groups in Spain]. *Estudios de Psicología, 34*, 197-208. doi: 10.1174/021093913806751375

Mackie, D. M., Devos, T., & Smith, E. R. (2000). Intergroup emotions: Explaining offensive action tendencies in an intergroup context. *Journal of Personality and Social Psychology, 79*, 602-616. doi: 10.1037/0022-3514.79.4.602

McDonald, R. P. (1999). *Test theory: A unified treatment*. Mahwah, NJ: Erlbaum.

Muthén, L. K., & Muthén, B. O. (1998-2011). *MPlus User's Guide* (Sixth ed.). Los Angeles, CA: Muthén & Muthén.

Ortiz, M. J. (1993). Conocimiento social: génesis y naturaleza de la representación de las figuras de apego [Social knowledge: genesis and nature of attachment figures' representation]. *Infancia y aprendizaje, 61*, 55-72. Retrieved from: http://dialnet.unirioja.es/descarga/articulo/48421.pdf

Osgood, C. E., Suci, G. J., & Tannenbaum, P. H. (1957). *The measurement of meaning.* Urbana IL: University of Illinois Press.

Raykov, T. (1997). Estimation of composite reliability for congeneric measures. *Applied Psychological Measurement, 21*, 173-184. doi: 10.1177/01466216970212006

Raykov, T., & Marcoulides, G. A. (2004). Using the Delta Method for Approximate Interval Estimation of Parameter Functions in SEM. *Structural Equation Modeling: A Multidisciplinary Journal*, *11*, 621–637. doi: 10.1207/s15328007sem1104_7

Reeder, G. D., & Brewer, M. B. (1979). A schematic model of dispositional attribution in interpersonal perception. *Psychological Review, 86*, 61-79. doi: http://dx.doi.org/10.1037/0033-295X.86.1.61

Rothbart, M., & Park, B. (1986). On the confirmability and disconfirmability of trait concepts. *Journal of Personality and Social Psychology, 50*, 131-142. doi: 10.1037//0022-3514.50.1.131

Schermelleh-Engel, K., Moosbrugger, H., & Müller, H. (2003). Evaluating the fit of structural equation models: Tests of significance and descriptive goodness-of-fit measures. *MPR-online*, *8*, 23-74.

Skowronski, J. J., & Carlston, D. E. (1987). Social judgment and social memory: The role of cue diagnosticity in negativity, positivity, and extremity biases. *Journal of Personality and Social Psychology, 52*, 689-699. doi: 10.1037/0022-3514.52.4.689

Stangor, C. (2009). The study of stereotyping, prejudice, and discrimination within social psychology: A quick history of theory and research. In T. D. Nelson (Ed.), *Handbook of prejudice, stereotyping, and discrimination* (pp. 1-22). New York, NY: Psychology Press.

Tausch, N., Kenworthy, J. B., & Hewstone, M. (2007). The confirmability and

    disconfirmability of trait concepts revisited: Does content matter? *Journal*

    of Personality and Social Psychology, 92, 542–556. doi:10.1037/0022-

    3514.92.3.542

Van Bavel, J. J., Packer, D. J., Haas, I. J., & Cunningham, W. A. (2012). The

    Importance of Moral Construal: Moral versus Non-Moral Construal Elicits

    Faster, More Extreme, Universal Evaluations of the Same Actions. *PLoS*

    ONE, 7: e48693. doi: 10.1371/journal.pone.0048693

Wojciszke, B. (2005). Morality and competence in person- and self-perception.

    European Review of Social Psychology, *16*, 155-188. doi:

    10.1080/10463280500229619

Table 1
Descriptive statistics for all the items

Items	M	SD	A	Sk	Items	M	SD	A	Sk
Competence					**Anger ER**				
Intelligent [Inteligentes]	4.18	1.24	-0.08	-0.09	Rage [Ira]	3.16	1.67	0.56	-0.41
Creative [Creativas]	3.88	1.31	-0.14	-0.30	Fury [Rabia]	3.26	1.74	0.43	-0.71
Skillful [Habilidosas]	4.50	1.19	-0.41	0.21	Anger [Enfado]	3.41	1.65	0.24	-0.78
Smart [Listas]	3.94	1.25	0.20	-0.25	Irritation [Irritación]	3.73	1.69	0.07	-0.77
Able [Capaces]	3.51	1.11	0.35	0.46	Frustration [Frustración]	3.24	1.64	0.38	-0.64
Hopeless [Torpes]*	4.38	1.25	-0.49	0.11	**Fear ER**				
Ignorant[Ignorantes]*	4.82	1.11	-0.60	0.70	Fear [Temor]	3.53	1.58	0.24	-0.61
Uncultured[Incultas]*	3.72	1.31	0.16	-0.49	Panic [Pánico]	3.29	1.58	0.53	-0.38
Not very bright [Con pocas luces]*	4.18	1.24	-0.08	-0.09	Vulnerability [Vulnerabilidad]	3.55	1.48	0.26	-0.46
Morality					**Semantic Differential**				
Sincere [Sinceras]	3.26	1.33	0.10	-0.51	Sweet-Bitter [Dulces-amargas]	3.74	1.33	0.04	-0.47
Honest [Honestas]	3.47	1.19	-0.09	-0.24	Transparent-Opaque [Transparentes-opacas]	3.48	1.30	0.13	-0.15
Respectful [Respetuosas]	3.61	1.29	0.11	-0.28	Light-Dark [Claras- Oscuras]	3.50	1.15	0.00	0.42
Trustworthy [De Confianza]	3.04	1.24	0.25	0.13	Perfect-Imperfect [Perfectas-Imperfectas]	3.14	1.35	0.46	0.04
Reliable [Formales]	3.18	1.17	0.32	0.39	Whole-Broken [Enteras-Rotas]	3.11	1.10	-0.13	-0.34
Malicious [Malintencionadas]*	4.35	1.32	0.14	-0.55	Tastey-Unpleasant [Sabrosas-Desagradables]	4.11	1.16	-0.11	0.79
Treacherous [Traicioneras]*	4.56	1.44	-0.01	-0.79	Innocuous –Poisonous [Inocuas-Venenosas]	3.46	1.25	0.01	-0.13
Aggressive [Agresivas]*	4.32	1.23	0.12	-0.34					
False [Falsas]*	4.50	1.37	-0.20	-0.40					
Harmful [Dañinas]*	4.18	1.31	0.25	-0.33					
Sociability									
Likeable [Simpáticas]	4.39	1.17	-0.45	0.24					
Affectionate [Cariñosas]	4.34	1.20	-0.39	0.22					
Open [Abiertas]	4.16	1.52	0.03	-0.88					
Helpful [Serviciales]	4.36	1.23	-0.24	-0.02					
Friendly [Amistosas]	2.62	1.03	0.46	0.92					
Shy [Tímidas]*	2.39	1.15	1.11	2.04					
Lonely [Solitarias]*	3.46	1.31	0.60	0.23					
Cold [Frías]*	3.45	1.47	0.47	-0.30					
Reticent[Reservadas]*	4.39	1.17	-0.45	0.24					

Note. M = Mean; *SD* = Standard Deviation; *Sk* = Skewness; *A* = Asymmetry.
*Negative items of the UDSP.

Table 2

Fit statistics for models 1 to 5

	$\chi2$	Df	RMSEA [90% CI]	CFI	TLI
Model 1	638.29	335	.04 [.04, .05]	.96	.95
Model 2	1946.01	764	.05 [.05, .06]	.89	.88
Model 3	928.2	449	.04 [.04, .05]	.94	.93
Model 4	1041.13	480	.05 [.04, .05]	.94	.93
Model 5	1106.27	449	.05 [.05, .06]	.92	.91

Note. $\chi2$ = Chi–square test $p < .001$; Df = Degrees of freedom;

RMSEA = Root Mean Square Error of Approximation; 90% CI =

90% Confidence Interval of the RMSEA; CFI = Comparative Fit

Index; TLI = Tucker–Lewis Index.

Table 3

Correlations, regression coefficients and coefficients of determination of the models with the UDSP combinations and the related variables

		Correlation coefficients				Regression coefficients		
		Fear	SD	Morality	Sociability	Anger	Fear	SD
Model 1 (All positive items)	Anger	.58**	-.35**					
	Fear		-.22**					
	Competence			.56**	.71**	-.02	-.10	.23**
	Morality				.65**	-.49**	-.48**	.67**
	Sociability					-.07	.11	.01
	R²					**.30**	**.22**	**.71**
Model 2 (All positive and negative items)	Anger	.52**	-.24**					
	Fear		-.11					
	Competence			-.60**	.64**	-.02	-.10	.28**
	Morality				-.64**	.59**	.54**	-.61**
	Sociability					-.05	.10	.07
	R²					**.39**	**.29**	**.74**
Model 3 (All positive items and negative competence items)	Anger	.57**	-.34**					
	Fear		-.19**					
	Competence			.58**	.64**	-.10	-.17*	.31**
	Morality				.65**	-.46**	-.45**	.64**
	Sociability					-.04	.13	.00
	R²					**.31**	**.24**	**.74**
Model 4 All positive items and negative morality items)	Anger	.53**	-.24**					
	Fear		-.13					
	Competence			-.52**	.71**	-.00	-.08	.24**
	Morality				-.60**	.60**	.57**	-.67**
	Sociability					-.04	.13	.04
	R²					**.39**	**.29**	**.73**
Model 5 (All positive items and negative sociability items)	Anger	.58**	-.35**					
	Fear		-.22**					
	Competence			.56**	.70**	-.01	-.07	.23**
	Morality				.67**	-.47**	-.47**	.68**
	Sociability					-.10	.06	.02
	R²					**.30**	**.22**	**.71**

Note. SD = semantic differential; R^2 = coefficient of determination. * $p < .05$; ** $p < .001$.

Table 4

Fit statistics for model 4 and invariance test

		χ2	Df	ΔDf	RMSEA [90% CI]	CFI	ΔCFI	TLI
Measurement	Free	593.11	279		.06 [.06, .07]	.93		.92
model	Invariant	633.13	318	39	.06 [.05, .07]	.93	0	.93
Structural	Free	1755.11	1059		.05 [.05, .05]	.93		.93
model	Invariant	1769.71	1074	15	.05 [.04, .05]	.93	0	.93

Note. χ2 = Chi–square test, *p* < .001; Df = Degrees of freedom; ΔDf = Degrees of freedom

differences between the model and the free model; RMSEA = Root Mean Square Error of

Approximation; 90% CI = 90% Confidence Interval of the RMSEA; CFI = Comparative Fit Index;

ΔCFI = CFI differences between the model and the free model.TLI = Tucker–Lewis Index.
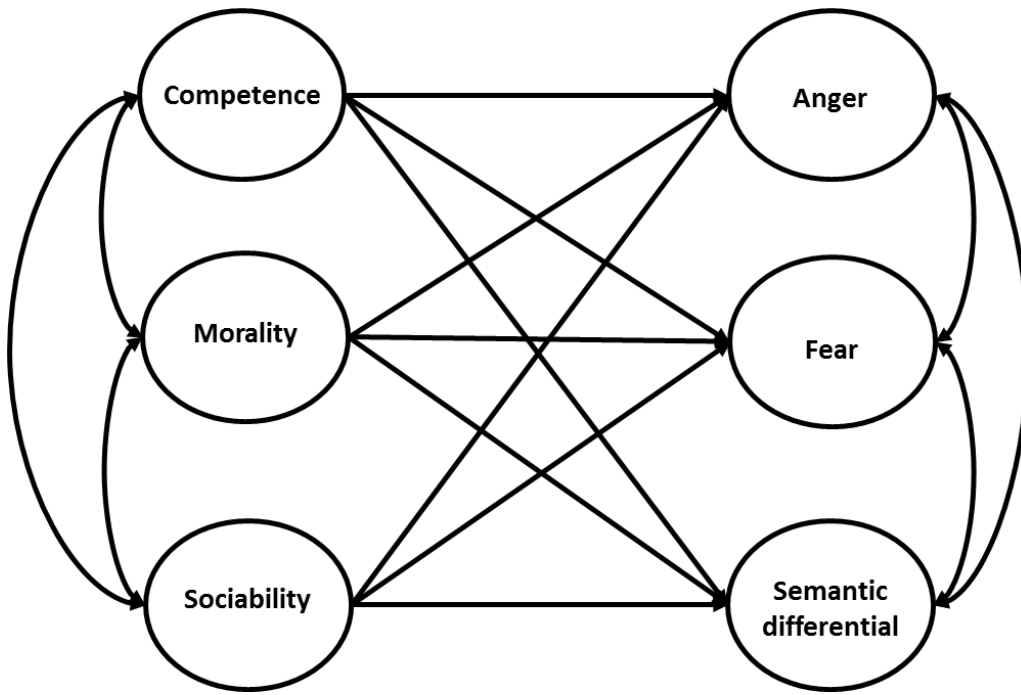
Fig. 1. Structural model to test the predictive capability of the UDSP
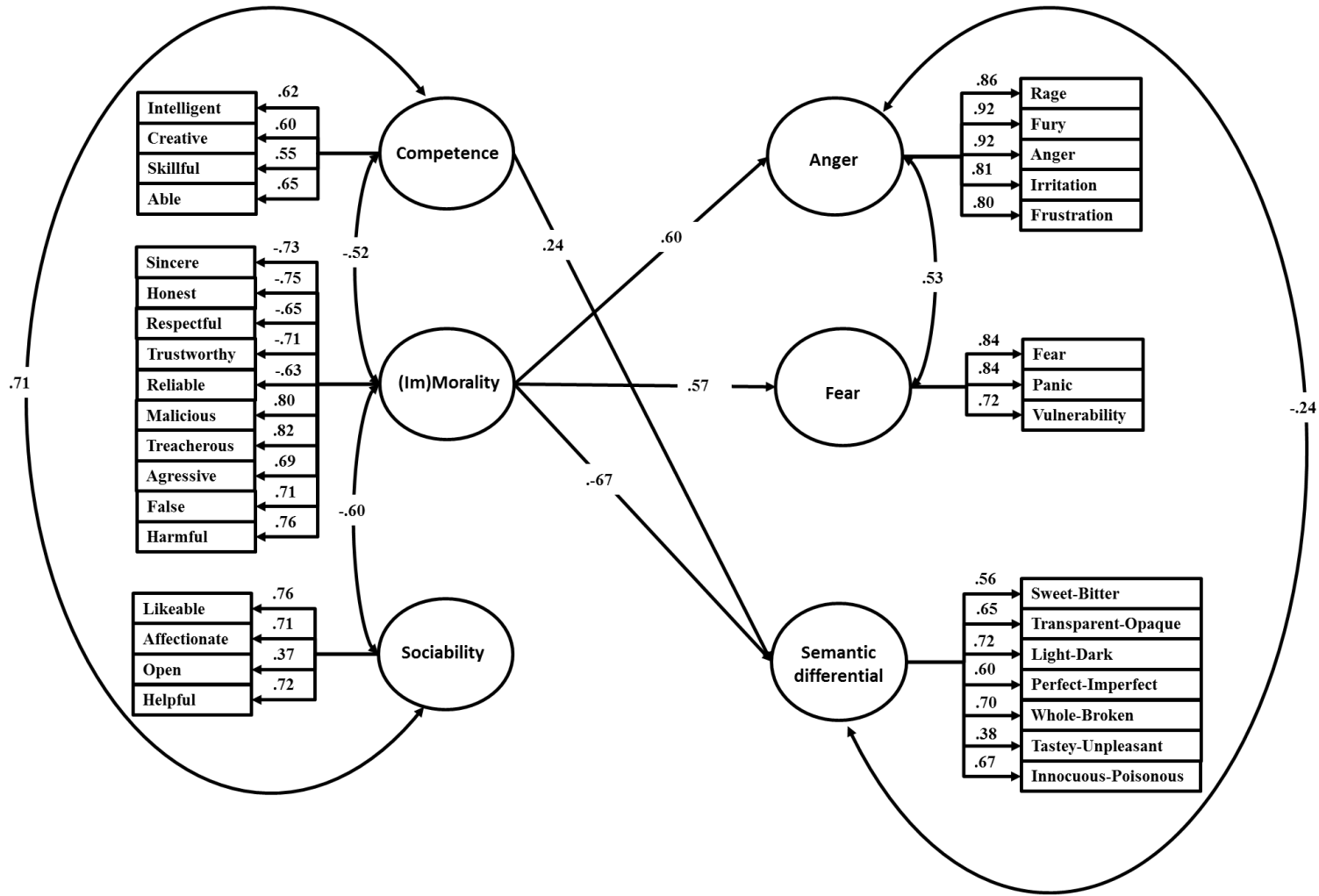scales combinations.

Fig. 2. Structural model including positive UDSP items and only negative morality items, invariant between the two randomly selected halves of the whole sample. Only significant standardized parameters are shown (p < .05).