MORALITY STEREOTYPE CONTENT SCALE (MSCS): RASCH ANALYSIS AND EVIDENCE OF VALIDITY

PABLO SAYANS-JIMÉNEZ, ISABEL CUADRADO GUIRADO, ANDREA BLANC MOLINA, JORGE ORDÓÑEZ-CARRASCO, ANTONIO ROJAS TEJADA

University of Almeria, Department of Psychology and Centre for Migration Studies and Intercultural Relations, Almería

ABSTRACT
The aim of this work was to develop a new scale, the Morality Stereotype Content Scale (MSCS), and to analyze its psychometric properties. The MSCS contains both positive and negative morality-specific attributes. MSCS utility stands out for its concision and its translation ease. The study was developed with the collaboration of a sample of 550 Spaniards (keeping with the composition, in age and gender, of the Spanish population). Stereotypes toward the Spanish Roma people were assessed. Polytomous Rasch model (rating scale model) was employed. The results showed MSCS good psychometric properties that support its use for monitoring of intergroup perceptions and the assessment of the effectiveness of such interventions aimed to reduce negative intergroup stereotypes, particularly in the Spanish and European context.

key words: stereotypes, morality, scale, stereotype content, item response theory

kľívová slova: stereotypy, morálka, škála, obsah stereotypu, teorie odpovědí na položku

Psychosocial literature is in general agreement that the features that may be used to evaluate a social target share to a greater or lesser degree, two types of content: 1) Content related to the ability or possibilities of the social target evaluated to achieve their goals, and 2) content related to beneficial or harmful intentions from or toward the social target (e.g., Abele & Wojciszke, 2014; Brambilla & Leach, 2014; Judd et al., 2005). This content has been linked to two dimensions that have received different names (e.g., agency and communion, self-profitability and other-profitability, competence and warmth), but show strong conceptual equivalence (Wojciszke & Abele, 2008). In intergroup relations and stereotypes, competence and warmth are terms used very often (Fiske et al., 1999; Fiske et al., 2002).

Group features representing these two kinds of contents have shown to be highly correlated (e.g., Fiske et al., 1999; Eckes, 2002; Judd et al., 2005; Lin et al., 2005; López-Rodríguez, Cuadrado & Navas, 2013; Sayans-Jiménez, Rojas, & Cuadrado, 2017). These correlations are mainly due to the shared evaluative content between

Došlo: 16. 6. 2017; A. R. T., Facultad de Psicología, Universidad de Almería, 04120 Almería, España, UE; e-mail: arojas@ual.es
This study was funded by the Andalusian Ministry of Innovation, Science and Business (Spain) under Excellence Research Project (Reference P09-SEJ-4657) and by the Spanish Ministry of Economy and Competitiveness (Research Project Reference PSI2011-22731).

366 / Předběžné studie
competence and warmth (Abele & Wojciszke, 2014; Sayans-Jiménez, Cuadrado, & Rojas, 2017), being the second dimension the one that has been demonstrated to have a predominant weight in making evaluations of other people or groups (Cuddy, Fiske, & Glick, 2008; Sayans-Jiménez et al., 2017; Wojciszke, 2005; Wojciszke, Bazinska, & Jaworski, 1998). That makes possible to use warmth features to capture group evaluations in a straight way. However, it is possible that the construct responsible of the higher evaluative weight of the warmth dimension do not correspond with its whole content. Many studies have shown that warmth consists of two differentiated contents, morality, and sociability (Leach, Bilali, & Pagliaro, 2015; Leach, Ellemers, & Barreto, 2007), or that social warmth and morality are separable dimensions (Goodwin, Piazza, & Rozin, 2014). The sociability dimension is related to cooperation, reciprocity, and developing relationships with others (Brambilla et al., 2011; Brambilla et al., 2012), whereas the morality dimension has been defined as “the perceived appropriateness of the behaviors of social targets” (Brambilla et al., 2013, p. 812).

In the context of the stereotype content, morality and sociability contents have shown to perform different roles in generating evaluations (Brambilla et al., 2011, 2012, 2013). Specifically, content related to morality has been shown to be more important for generating evaluations of both people and groups (Brambilla & Leach, 2014; Goodwin et al., 2014). It has further been shown to be more related to other psychosocial variables, such as acculturation preferences and prosocial tendencies (López-Rodriguez & Zagefka, 2015), behavioral intentions (Brambilla et al., 2013), perception of threat (Brambilla et al., 2012), or individual and group identity (Brambilla et al., 2013; Goodwin et al., 2014).

Morality content importance is probably mainly due to its function (i.e., inform about possible harm or benefit). When people think about the morality of an object, their “thoughts” are faster, more extreme and are more strongly associated to universal prescriptions (i.e., thoughts assumed to be shared by everyone; Van Bavel et al., 2012). Furthermore, morality-related information is perceived to be more reliable and objective (Goodwin & Darley, 2012). Moreover, the capability of morality stereotypes for capturing relevant group-related evaluations can be strengthened if both positive and negative morality features are included in the same assessing scale (see Sayans-Jiménez, Rojas, et al., 2017).

The role of the morality dimension is broadened when its content refers to the negative evaluative pole (Baumeister et al., 2001; Martijn et al., 2012; Skowronski & Carlston, 1987). 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). These attributes referring to morality inform about the potential harm to the individual (Phalet & Poppe, 1997; Wojciszke et al., 1998). The negative evaluative pole of morality attracts more attention and has more evaluative weight than the positive pole (Goodwin & Darley, 2012). Furthermore, immoral acts encompass negative effects and the violation of norms (Gray, Young, & Waytz, 2012). Attributes associated with the malice of negative moral actions (e.g., stealing, cheating) are perceived to be more reliable and objective than those referring to the goodness of positive moral actions (e.g., donate money; Goodwin & Darley, 2012; Rothbart & Park, 1986; Skowronski & Carlston, 1987). Additionally, when trying to predict emotional reactions and global evaluations based on the stereotype content the best way to measure morality content is to use positive and negative attributes of morality (Sayans-Jiménez, Rojas et al., 2017). Although the predictive advantage of the negative attributes of morality has been highlighted in different researches, these negative attributes are not commonly used when working with the
basic dimensions of social categorization, using either the morality or warmth dimension (Brambilla et al., 2011, 2012; Cuddy et al., 2009; Fiske et al., 1999, 2002; López-Rodríguez, Cuadrado, & Navas, 2013).

In view of the importance of measuring the content of the morality dimension, the purpose of this paper is to show the psychometric properties of a new scale for measuring the stereotype content of the morality dimension associated with outgroups or people belonging to them, the Morality Stereotype Content Scale (MSCS). Positive and negative features which only reflect the specific content associated with morality (e.g., trustworthy, sincere) were used for the scale, avoiding attributes that could share content referring to both morality and sociability (e.g., attributes related to kindness, benevolence, lovingness; Goodwin et al., 2014). The MSCS also includes both attributes related more to “virtue” than to moral (honest or trustworthy) and attributes reflecting intentions and possible results of interaction with outgroup people (e.g., harmful or aggressive).

The MSCS has been designed to be a quick scale capable of capture the most relevant content related to the moral representation of the social target. Despite people can hold different stereotypes to different groups and that context-specific stereotypes are better predictor of group evaluations (Kleinpenning & Hagendoorn, 1991), it would be hard to generate group-specific scales for each group settled in our countries, and psychometrically inadequate (i.e., it would not be possible to compare directly or easily evaluations of different groups with different scales). We suggest that the MSCS composition (it captures the content of morality stereotypes directly related to other-profitable features; see Abele & Wojciszke, 2014) will allow applying it as a useful tool in intergroup context were the assessment of the perceived morality of one or more (out)groups would be required (either for assessing the effectiveness of interventions or for making quick diagnoses on intergroup perceptions). Furthermore, the MSCS is a measurement tool extremely easy to translate owing to the universality of its items.

Given the serious impact that negative stereotypes can exert on minority ethnic groups in their access to resources (e.g., Williams & Rucker, 2000), threatening their self-images (e.g., Crocker & Garcia, 2015) and also in their mental health and wellbeing (e.g., Missinne & Bracke, 2012; Veling, Hoek, & Mackenbach, 2008) we have chosen to show the MSCS psychometric properties when applied to the most relevant (and the poorest) non-migrant ethnic minority group in Europe, Roma people (Sordé, Flecha, & Mireea Alexiu, 2013). The prejudice against them is extended along the whole continent (García-Ramírez, Escobar-Ballesta, & Lizana, 2015) perpetuating social inequity and discriminatory behaviors toward them (Escobar-Ballesta et al., 2018). This unfair situation is extended also to the Spanish Roma population (for a concise description see Urbíola et al., 2014). Consequently, we propose the MSCS as a tool essential for planning and assessing interventions aimed to reduce this prejudice and inequity that oppress the Roma Population. MSCS psychometric properties will be analyzed using Rasch’s rating scale model (RSM, Andrich, 1978; Masters & Wright 1984) indicated for scales with identical rating scales for all their items. The RSM, a polytomous Rasch model, represents in a unidimensional continuum of measurement the conjoint measure of persons and items. Morality items scaling shows “how morally relevant different traits are, which should distinguish the extent to which such traits drive person perception processes, particularly those related to impression formation” (Goodwin et al., 2014, p. 149). The use of this measurement model will also allow to get a greater amount of information about the psychometric properties of the MSCS, as the reliability of the scale in every point of the morality continuum or the adequacy of the response categories of the items.
Finally, evidence of validity is shown based on the relationship with other constructs. On one hand, the relationship of MSCS scores with a Semantic Differential of Evaluation (SDE) of the same outgroup is estimated. The two measurements are expected to be highly correlated. Finally, to test the invariance of the MSCS parameters, cross-validation test was conducted with two randomly selected subsets of the sample.

METHOD

Participants

Five hundred and fifty people residing in Spain, 280 women and 270 men, participated. Sample selection was by quota sampling by age and sex in keeping with the composition of the Spanish population (none of them belonged to the Spanish Roma group). In sample, women were 50.09% and the rest men. The age intervals 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

Morality Stereotype Content Scale (MSCS). The attributes that make up the scale are the result of an item selection process that studied the psychometric properties for: 1) The positive items (five items) that are used most often when the morality dimension is employed (e.g., Brambilla et al., 2011, 2012; Cuddy et al., 2008; Fiske et al., 1999, 2002; López-Rodríguez et al., 2013), and 2) New items found (five negative items) from a review of the literature and experts’ suggestions. The appropriateness of these attributes to predict emotional reactions and global evaluation together with positive items of competence and sociability can be seen in Sayans-Jiménez et al. (2017). The items used were: sincere, honest, respectful, trustworthy, reliable, malicious, treacherous, aggressive, false, and harmful. The scale measured how non-Roma people represent people of the Spanish Roma group over the morality dimension. Ingroup and outgroup social identities were made salient at the beginning of the questionnaire the first by tagging the person who answered the scales as “not Roma”.

The instructions contained an important difference with respect to the instructions used in measuring stereotypes in the context of intergroup relations. Quinn, Mason, and Macrae (2009) showed that functionality of social categorization diminishes when information on the individual identity of the attitude target is available. Therefore, the purpose of the instructions was for the respondent not to concentrate on concrete characteristics of outgroup people whom they might know personally (about whom they have acquired individuating information through interaction). To do that, the raters were asked to imagine the composition of a big group of non-familiar Spanish Roma people (more than 300 people) and to decide how many of them possess the listed attributes. Seven-choice Likert-type (none, almost none, little, half, many, almost all and all) item response scales were used. Higher scores imply that the outgroup target is more morally qualified.

Semantic Differential of Evaluation (SDE). A Semantic Differential scale with seven-item (see Díaz-Guerrero & Salas, 1975) was used. All the items (with seven-point responses) referred to the evaluation dimension. The pairs of adjectives used were: Sweet-Bitter, Transparent-Opaque, Light-Dark, Perfect-Imperfect, Whole-Broken, Tasty-Unpleasant, Innocuous-Poisonous. The order and direction of the items were randomized to control for method effects. After application, the items were recoded for an easier interpretation. The SDE scores were calculated by averaging the
empirical responses on each item following the instructions of Diaz-Guerrero and Salas (1975). Higher scores imply more positive evaluations.

Procedure
One questionnaire was delivered individually by trained staff. The questionnaire included attributes of different contents randomized in order to prevent method effects. Subsequently, the questionnaire included other measurements of psychosocial variables such as SDE. There was no time limit. Anonymity and confidentiality of participants were guaranteed. 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
Application of the RSM. First, as a preliminary step before RSM application, scale dimensionality was examined by optimal implementation of parallel analysis (PA; Horn, 1965) using the O’Connor macro for SPSS (O’Connor, 2000). Second, the scale’s psychometric properties were estimated by RSM (Wright & Masters, 1982), a polytomous Rasch model. Data fit to the model were examined by residual analysis. The mean square residuals and standardized mean square residuals for items and persons (MNSQ and ZEMP, respectively) used were: Infit and Outfit (Linacre, 2002). Infit is the fit index of expected to observed values sensitive to behavior of items near the score of the persons evaluated. Outfit is sensitive to the behavior of items far from the score of persons evaluated. Since the MNSQ is more affected by the sample size than the ZEMP (Smith, Schumacker, & Bush, 1998), they had more weight for exploring item fit. ZEMP values considered appropriate are those from -2 to +2 (Linacre, 2002). The position of persons and items was estimated (and of response categories) on the morality continuum using the RSM. Finally, the separation values of both items and persons were estimated as reliability indicators. For persons, separation values below 2.00 show that the scale is unable to distinguish between at least two statistically different strata in the sample (e.g., persons who associate the target outgroup with high morality vs. persons who associate it with low morality). Items under 3.00 show that the sample used does not enable us to confirm organization into a hierarchy of items based on their moral relevance (Linacre, 2014a). The analyses were done with WINSTEPS software version 3.63.2 (Linacre, 2014b).

Evidence of validity. To study evidence of validity in relation to other variables, the correlation between the MSCS and the SDE scores was calculated. The invariance of item parameters was also checked. The sample was divided in half at random and the parameters of the items estimated separately. Then they were checked by linear regression to see if the estimates of the parameters remained stable in each half. The values expected in a perfect fit would be 1.00 for the correlation between both datasets, 0.00 for the coordinate at the origin and 1.00 for the slope of the line. The 95% control lines based on the SEs for each of the item pairs are plotted on the regression graph, following Bond (2003).

RESULTS
Dimensionality
The results of optimal implementation of parallel analysis (PA) support the one-dimensional assumption for MSCS. The first factor explains 58.24% of the common variance, while the following possible factor would only explain 7.89% of the common variance.

370 / Předběžné studie
Application of the RSM

Fit of the rating scale model. The results of statistical fit are shown in Table 1. Two persons were excluded from the analysis because of their extreme scores. The RSM applied is able to explain 62.20% of the empirical variance in the MSCS (compared to 6.50% of the variance explained in the first test). The mean infit and outfit MNSQ statistics for persons are both 1.00 and the ZEMP is -0.20. For items, the mean infit MNSQ is 0.99 and the mean ZEMP is -0.10, and for the outfit the mean MNSQ is 1.00 and the ZEMP -0.10. The ZEMP statistics for both internal fit (infit) and external fit (outfit) of persons and items are indicators of adequate fit. These results also provide evidence for the unidimensionality of the MSCS. Fit of each of the items shows adequate values for all items (Table 1). The results of the item-total correlations show optimal values.

<table>
<thead>
<tr>
<th></th>
<th>INFIT</th>
<th>OUTFIT</th>
<th>Item-total correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$M$</td>
<td>Error</td>
<td>MNSQ</td>
</tr>
<tr>
<td>Person</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$M$</td>
<td>-0.79</td>
<td>0.42</td>
<td>1.00</td>
</tr>
<tr>
<td>$SD$</td>
<td>1.37</td>
<td>0.05</td>
<td>0.88</td>
</tr>
<tr>
<td>Max.</td>
<td>3.84</td>
<td>0.50</td>
<td>0.44</td>
</tr>
<tr>
<td>Min.</td>
<td>-6.57</td>
<td>1.82</td>
<td>1.00</td>
</tr>
<tr>
<td>Items</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$M$</td>
<td>0.00</td>
<td>0.05</td>
<td>0.99</td>
</tr>
<tr>
<td>$SD$</td>
<td>0.33</td>
<td>0.00</td>
<td>0.11</td>
</tr>
<tr>
<td>Max.</td>
<td>0.60</td>
<td>0.05</td>
<td>0.98</td>
</tr>
<tr>
<td>Min.</td>
<td>-0.50</td>
<td>0.05</td>
<td>0.94</td>
</tr>
<tr>
<td>Items</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trustworthy</td>
<td>0.60</td>
<td>0.05</td>
<td>0.98</td>
</tr>
<tr>
<td>[De Confianza]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reliable [Formales]</td>
<td>0.41</td>
<td>0.05</td>
<td>1.03</td>
</tr>
<tr>
<td>Sincere [Sinceras]</td>
<td>0.29</td>
<td>0.05</td>
<td>1.01</td>
</tr>
<tr>
<td>Treacherous [Traicioneras]</td>
<td>0.03</td>
<td>0.05</td>
<td>0.95</td>
</tr>
<tr>
<td>Honest [Honetas]</td>
<td>-0.02</td>
<td>0.05</td>
<td>0.80</td>
</tr>
<tr>
<td>False [Falsas]</td>
<td>-0.05</td>
<td>0.05</td>
<td>1.12</td>
</tr>
<tr>
<td>Respectful [Respetuosas]</td>
<td>-0.21</td>
<td>0.05</td>
<td>1.18</td>
</tr>
<tr>
<td>Malicious [Malintencionadas]</td>
<td>-0.25</td>
<td>0.05</td>
<td>0.82</td>
</tr>
<tr>
<td>Aggressive [Agresivas]</td>
<td>-0.31</td>
<td>0.05</td>
<td>1.08</td>
</tr>
<tr>
<td>Harmful [Dañinas]</td>
<td>-0.50</td>
<td>0.05</td>
<td>0.94</td>
</tr>
<tr>
<td>Thresholds</td>
<td>$\delta_1$</td>
<td>$\delta_2$</td>
<td>$\delta_3$</td>
</tr>
<tr>
<td></td>
<td>-3.02</td>
<td>-2.4</td>
<td>-0.42</td>
</tr>
</tbody>
</table>

Note. The items are ordered according to their moral relevance (items are not shown in the administration order).
**Item calibration and person measurement.** Calibration of items shows the position of items on the continuum. Table 1 shows how the scale items have values from -0.5 for the *harmful* item to +0.60 logits for the *trustworthy* item, with a standard error of measurement of 0.05. Table 1 also shows calibration of the threshold parameters that delimit the location of the answer choices and informs on adequacy.

The combined calibration of persons and items may be observed graphically on the persons and items map (Figure 1), which also shows the distribution of persons based on their measurement. The items are relatively close together. This shows that they measure the construct with similar intensity, but are not redundant since the content of each item is different. As shown in Figure 1, most of the items on the positive pole are able to measure more of the construct.

The mean person score is slightly above the item mean. However, as observed in Figure 1, the items and item responses, are able to cover the whole range of the continuum on which the persons assessed are located satisfactorily.

**Response category characteristic curves.** The study of the characteristic curves of the response categories (RCCC) makes it possible to find out whether each category functions adequately. The vertical axis shows the probability of response to each category and the horizontal axis shows the locations of each category on the continuum. As shown in Figure 2, all the response categories are the most likely at some point of the continuum. This shows that the seven response categories proposed are ordered as theoretically proposed and function adequately.

**Measurement of accuracy.** To find out the accuracy of the estimated scores on the construct, the test information function and distribution of standard errors of measurement were used without distinction (Figure 3). These functions (one is the reciprocal of the other) make it possible to find out the error that affects a certain level of the construct and at what point on the continuum the test is the most informative. The separation index between persons was 2.96. This shows that the MSCS is sufficiently sensitive to discriminate between three statistically different strata in the set of persons as a function of the morality they associate with the outgroup target. This value is equivalent to a Cronbach’s alpha of .90. Furthermore, the separation of the items was 6.23, which corresponds to an item reliability of .97. Both the item and person separation indices back adequacy of the MSCS to the sample employed and vice versa.

**Evidence of validity**

As evidence of validity based on its relationship with other constructs, the correlation between the MSCS scores and the SDE evaluation dimension (both measures using the RSM) gives an adequate Pearson’s correlation coefficient \( r = .71, p < .001 \). This means there is favorable evidence of validity for the interpretation of MSCS scores.

To test the stability of the MSCS parameters, the parameters of the items were estimated using two random subsets of the total sample (Subset 1 and Subset 2), and a regression was performed between the parameters of both sets. The result of the regression between the two subsamples may be seen in Figure 4. Both sets of parameters of the items yield a very high Pearson’s correlation coefficient \( r = .98, p < .001 \), practically equal to 0 on the ordinate at the origin and 1.13 \( p < .001 \) for the slope of the regression line. It can be seen graphically how all the points remain within the 95% control lines based on the SEs for each of the item pairs, so it can be assumed that the MSCS parameters are invariant for the sample with which they are estimated.
Figure 1. Item and person map (including the largest and the lowest category location)

Note. Each “#” represents four people and each “.” from one to three. (i) = inverted item.
Figure 2. Category characteristic curves
Figure 3. Standard error and information function of MSCS
Figure 4. Item difficulty invariance (95% control lines based on the SEs for each of the item pairs)
MSCS: Morality Stereotype Content Scale
DISCUSSION

The importance of morality in forming impressions about other people or groups has been widely recognized (Brambilla & Leach, 2014; Goodwin et al., 2014). It has been shown to be more useful in the study of intergroup relationships than the dimensions of competence and sociability (Brambilla et al., 2011, 2012, 2013). The purpose of this study was to analyze the psychometric properties of the MSCS, a scale designed to measure the perceived morality of an outgroup based on the stereotype content (i.e., to what extent a specific outgroup is associated with goals or positive or negative results for the person or group assessed or the group they belong to).

The psychometric properties of the MSCS have been provided and its usefulness for carrying out measurements of morality of a group or its members has been confirmed. In the first place, the scale dimensionality analysis shows that most of the common variance captured by the set of items used corresponds to a single dimension. The unidimensionality of the MSCS parameters is backed by the fit of data to the model.

All the items employed show adequate fit. The stronger diagnostic nature and heavier evaluative weight of the negative attributes of the morality dimension (Goodwin & Darley, 2012; Rothbart & Park, 1986; Skowronski & Carlston, 1987) are not reflected by the position of the items in Figure 2. These results could be reproducing the greater influence of social desirability when negative outgroup attributes are expressed. The content and the relevance of the attributes of the negative pole when people make moral judgments of outgroups (Brambilla et al., 2011, 2012, 2013) are indispensable to the measurement of morality.

With respect to the response categories, the results of this study show that the scale with seven response categories is adequate for capturing all the possible answer choices over the morality continuum. These results coincide with those that emphasize that seven-point rating scales are more accurate for the measurement of attitudes and beliefs (Krosnick, Judd, & Wittenbrink, 2005).

The separation indexes show that precision is adequate for both persons and items. Notably these results suggest that the MSCS is able to distinguishing between three statistically different strata within the sample. The MSCS information function enables us to affirm that this scale is able to measure all the persons in the sample (i.e., regardless of the person position on the continuum) with an acceptable error of measurement. Reliability of both persons and items scores was good. Finally, although the invariance property of the item parameters is intrinsic to the Rasch RSM, this study has shown empirically that this property is complied with.

The importance of the morality dimension for generating global evaluations (Wojciszke, Bazinska, & Jaworski, 1998) is reflected in the relationship between the scores on the MSCS and the SDE. This relationship is high and in the expected direction. This evidence guarantees the usefulness of the content of the morality dimension for making predictions on global evaluations of different social targets.

To sum up, the MSCS, in contrast to other morality-related scales like the Moral Foundations Questionnaire (Graham et al., 2011), the Moral Identity Scale (Aquino & Reed, 2002), or the Affective Morality Index (Cimbro & McIntosh, 2003), is set completely within stereotyping framework. Additionally, it is a quick scale, including positive and negative morality attributes, with a simple factor structure and solid psychometric properties that allows to scale both items and people in a moral continuum in an aseptic way (i.e., without implying further constructs and only focused on the stereotype content). Therefore, and owing that morality content of stereotypes have shown to be decisive in whole variety of intergroup processes (e.g., Brambilla et al.,
Since policies, campaigns, or programs destined for reducing prejudice are not always effective (e.g., Vrij & Smith, 1999), it is necessary to provide these programs with measurement tools that facilitate the monitoring of intergroup perceptions and the assessment of the effectiveness of such interventions aimed to reduce negative intergroup stereotypes. The MSCS is a measurement tool easy to translate so that it would work in different countries with different languages. A good illustration of a context where this kind of scale could be useful is in the European Union. The (supposed) free movement of European citizen all along its territory has accelerated the creation of new migrant groups (i.e., similar migrant groups can get to different countries) and, equally important, several different countries have to unite their efforts to reduce their historical discrimination toward the same minority groups (e.g., Roma people).

Owing to the fact that (theoretically) the European Union is really concerned about Roma discrimination (Escobar-Ballesta et al., 2018; European Commission, 2011) we consider that the MSCS is a relevant tool for this context. Specifically, we think that the MSCS could be useful in the Spain, for the relevance of the Roma group in this country (Sordé et al., 2013) and because the MSCS have provided evidences, in a different sample, of MSCS capability for differentiating discrimination profiles toward Spanish Roma (Sayans-Jíménez, Órdoñez-Carrasco, & Rojas, 2017). Particularly, the MSCS would be useful for assessing interventions aimed to reduce prejudice against Roma People (e.g., Urbiola et al., 2014) or for studying the relationship among morality stereotype content with other psychosocial variables such as multicultural ideology (e.g., Urbiola et al., 2018).

In conclusion, the MSCS is a scale with psychometric properties that guarantees an adequate measurement of morality associated with social targets. Future studies should test the usefulness of the MSCS when the social target is related to the interpersonal context. It would also be recommendable to compile new evidence of validity with behavioral criteria (e.g., other constructs relevant to the study of intergroup relations: perception of threat, acculturation strategies or emotional reactions).

Supplementary studies should address the development of similar scales referred to the competence and sociability content. Following Sayans-Jíménez et al. (2017) results, those scales should only include positive attributes, in contrast to the MSCS. Likewise, the MSCS could be used to explore how the morality content affects the desirability of competence and sociability attributes (e.g., Landy, Piazza, & Goodwin, 2016).

REFERENCES


Krosnick, J. A., Judd, C. M., & Wittenbrink, B. (2005). The measurement of


**SOUHRN**

Škála obsahu mravních stereotypů (MSCS): Raschova analýza a doklad validity

Cílem bylo vyvinout novou škálu, Škálu obsahů mravních stereotypů (MSCS) a analyzovat její psychometrické vlastnosti. škála obsahuje jak pozitivní, tak negativní morálně specifické atributy. Účinnost škály je dána její koncizností a snadnou jejího překladu. Validizační studie byla realizována s 550 Španěly při zachování poměrů složení španělské populace (zastoupení věku a pohlaví). Byly zjištěny stereotypní vůči španělským Romům. Byl použit polytomický Raschův model (model posuzovací škály). Výsledky ukazaly dobré psychometrické vlastnosti škály, které ji doporučuji pro monitorování meziskupinového vnímání a hodnocení efektivity intervencí zaměřených na redukci negativních meziskupinových stereotypů, zvláště ve španělském a evropském kontextu.