Running Head: COUNTER-STEREOTYPES

Counter-stereotypes Reduce Emotional Intergroup Bias by Eliciting Surprise in the Face of Unexpected Category Combinations

Francesca Prati[1], Richard J. Crisp[2] & Monica Rubini[1]

[1]Department of Psychology, Alma Mater Studiorum University of Bologna, Bologna, Italy

[2]Aston Business School, Aston University, Birmingham, UK

Author note

Correspondence regarding this article should be sent to Francesca Prati at the University of Bologna, Department of Psychology, 5 Berti Pichat, 40126, Italy; phone number: 0039 051 2091826; email: francesca.prati@unibo.it

Counter-stereotypes Reduce Emotional Intergroup Bias by Eliciting Surprise in the Face of Unexpected Category Combinations

A person who belongs to a surprising combination of social categories can be described as a counter-stereotype. Research has shown that being encouraged to think about counter-stereotypes leads to a lesser reliance on stereotypes in favor of bespoke impressions (Hastie, Schroeder & Weber, 1990; Hutter & Crisp, 2005; 2006; Kunda, Miller & Claire, 1990). However, when we think, we also feel, and it is notable that the affective outcomes of counter-stereotype exposure have received less attention than cognitive outcomes. The current paper aims to fill this gap by examining the impact on emotional intergroup bias following exposure to counter-stereotypes. Specifically, we examined whether and how exposure to counter-stereotypes can reduce the application of biased, stereotyped emotions and prejudice in the form of dehumanization.

Stereotypes and Stereotyped Emotions

Recent research has shown that cognitive representations of close and distant groups can be highly predictive of emotions and behaviors that individuals adopt towards these groups. We refer in particular to the evidence accumulated for the *Stereotype Content Model* (SCM; Fiske Cuddy, Glick & Xu, 2002). Fiske et al. (2002) showed that social groups are distinguished along the two dimensions of competence and warmth, resulting in four combinations that evoke unique patterns of intergroup emotions and corresponding behaviours. Groups perceived as highly warm and competent (e.g., middle class, students) elicit positive emotions, such as admiration and pride, whereas groups stereotyped as low in warmth and incompetent (e.g., immigrants, homeless) elicit extremely negative emotions, such as contempt. Groups stereotyped as low on competence but high on warmth, such as women and elderly people, are cross-culturally viewed as harmless but pathetic, likely eliciting the specific emotion of pity, but not a willingness to work with them (Cuddy, Norton, & Fiske, 2005). In contrast, groups stereotyped as high in competence and low in warmth, including men and rich people, are generally viewed as worthy of respect, elicit cooperation, but they are not liked and evoke the emotion of envy.

The SCM shows how emotional experiences can be an outcome of stereotypical judgments and a predictor of biased intergroup behaviour. However, until now applications of this model have arguably provided a rather static representation of intergroup cognition, emotion and behaviour. People do not always rely on only single categorizations (specific groups are not always pitied, respected or disliked); rather, they often rely on multiple categorizations when forming impressions of others (Vanbeselaere, 1987). Accordingly, increasing the number of categories which simultaneously define an outgroup target can reduce intergroup bias (Crisp & Hewstone, 2007; Roccas & Brewer, 2002). It is therefore plausible that emotions, like social judgments, can be experienced in a less biased fashion when more than one categorical criterion defining individuals' group membership is salient.

Counter-stereotypes and Countering Bias

A particular type of multiple categorization has proved to be an important way of reducing intergroup bias: *counter-stereotype exposure*. Counter-stereotypes are characterized by the conjunction of two non-overlapping (i.e., conflicting) categories. Thus, when more than one source of information about a target is simultaneously salient, the task for the social perceiver is to integrate them to form a holistic impression that somehow integrates the conflicting information. In other words, to form an impression of someone described by a surprising combination of social categories (e.g., an Oxford-educated bricklayer) one must inhibit the use of the conflicting stereotypes typically associated with the constituent categories (Hastie et al., 1990; Kunda et al., 1990). Hutter and Crisp (2005) showed that after thinking about counter-stereotypic conjunctions (e.g., a female mechanic), individuals were more surprised and generated fewer stereotypic attributes to describe the target in comparison to stereotypic conjunctions (e.g., a female nurse). Correspondingly, when the target cannot be fitted into existing categories (e.g., a gay priest, a male midwife), this grabs perceivers' attention and they shift out of a stereotypic thinking mode as they try to figure out how to classify the target (see Hutter, Crisp, Humphreys, Waters, & Moffitt, 2009).

Drawing on these findings, in their Categorization-Processing-Adaptation-Generalization (CPAG) model, Crisp and Turner (2011) argued that the process of stereotypic inconsistency resolution results in a generalized, reduced tendency to use the heuristic thinking system. This processes weakens individuals' reliance on memory structures such as stereotypes, and facilitates positive impressions based on the application of individuating and non-stereotypic attributes. Accordingly, recent work has shown that after thinking about gender counter-stereotypic targets, participants demonstrated a generalized decrease in linguistic outgroup derogation and also decrease dehumanization towards a range of unrelated target groups, an effect that can be explained by reduced reliance on heuristic thinking (Prati, Menegatti, & Rubini, 2015; Prati, Vasiljevic, Crisp, & Rubini, 2015). Thinking about counter-stereotypes also improves cognitive outcomes such as creativity (Gocłowska & Crisp, 2013), lateral thinking and generalized intergroup tolerance (Vasiljevic & Crisp, 2013) as well as, attenuating preference for stereotypic leaders (Leicht, Randsley de Moura, & Crisp, 2014). In these studies, where a baseline condition has been included, no difference has been found between baseline and stereotypic conditions. This supports the idea that stereotypic thinking is the default mode of the social mind, and that counter-stereotype exposure shifts perceivers' processing out of this default in to new territory (Vasiljevic & Crisp, 2013).

Despite the growing amount of research on the (cognitive) outcomes of counter-stereotypic exposure, until now only one contribution has considered the extended benefits of this cognitive shift on *affective* outcomes. Ray, Mackie, Smith, and Terman (2012) showed that Greek Democrats and non-Greek Republicans (partial ingroup, partial outgroup) were evaluated equivalently by non-Greek Democrats, but the former were elicitors of ambivalent emotions, such as admiration and disgust, whereas the latter were affectively neutral, eliciting neither admiration nor disgust. These different emotional reactions are important because they predict specific action tendencies (e.g. see Cuddy, Fiske & Glick, 2007; Mackie, Devos & Smith, 2000; Pettigrew & Tropp, 2008). Non-Greek Republicans could be treated with ambivalence, whereas Greek Democrats might simply be

ignored. In this vein, intergroup emotions can be said to permit more refined predictions of intergroup outcomes of cross-categorized groups. However, an emotional reaction many not only follow from the cognitive process of inconsistency resolution, but also precede it. In some of the earliest research on counter-stereotypes, Kunda et al. (1990) argued that the emotion of surprise is experienced when individuals are encouraged to think about counter-stereotypic conjunctions (see also Hutter and Crisp, 2005). This then triggers an 'on-line' process of inconsistency resolution which helps to modify existing stereotypic representations. This suggests that counter-stereotypes should produce beneficial outcomes through the experience of surprise.

The Emotional Experience of Surprise

Studies have shown that surprise is one of the most basic and universal of emotions (e.g. Ekman, 1972; Fisk, 2002; Izard, 1977; Plutchik, 1980; Roseman, 1996) which usually results in the interruption of ongoing thoughts and activities and motivates people to pay attention to the unexpected (e.g., Kunda et al., 1990; Meyer, Reisenzein, & Schützwohl, 1997; Ortony & Partridge, 1987; Schützwohl & Reisenzein, 1999). Distinct from emotions such as joy or fear, surprise does not presuppose the appraisal of the eliciting information as positive (motive-congruent) or negative (motive-incongruent), and the feeling of surprise is *per se* hedonically neutral rather than pleasant or unpleasant (Reisenzein, 2009). Meyer et al. (1997) contended that the major evolutionary function of surprise is to monitor individuals' cognitive schemas by updating them in the face of unexpectedness. In line with this view, Maguire, Maguire & Keane (2011) showed that surprise manifests individuals' need for information-seeking. Accordingly, the feeling of surprise serves to inform the self about the occurrence of a schema-discrepancy and it provides an initial motivation for the analysis and resolution of the schema-discrepancy by eliciting curiosity about its nature and cause. In this vein, surprise can be conceived as the emotional link between perceived inconsistencies and awareness of expectancy violation. Broadly speaking, this is also in line with theoretical models that highlight the primary role of emotions in influencing cognition and behavior (e.g., Frijda, Manstead, & Bem, 2000; Zajonc, 1980) and evidence that mood states affect memories

and judgments (Clark & Fiske, 1982; Fiedler & Forgas, 1988; Forgas, 1991, 2000; Martin & Clore, 2001; Moscatelli, Albarello, Prati & Rubini, 2014). In particular, dealing with surprising (counter-stereotypic) category conjunctions elicits a cognitive shift to more careful, and accurate, information processing in order to resolve the stereotypic inconsistency (Hutter & Crisp, 2005). Following this line of research, we propose that the emotion of surprise, resulting from counter-stereotype exposure, can lead to an inconsistency resolution process that, in turn, reduces the application of stereotyped emotions.

The Current Research

In the light of the aforementioned literature we aimed to examine the impact of counter-stereotypic category conjunctions on intergroup emotions experienced towards outgroup members. To achieve these goals, three experiments were carried out[1] using different targets of discrimination: on the basis of gender (Eagly & Steffen, 1984) and on the basis of national group membership (Pratto, Sidanius, Stallworth, & Malle, 1994). Specifically, Experiment 1 tested whether after thinking about gender counter-stereotypic targets (*female mechanic, male nurse*), gender stereotyped emotions (pity towards women and envy towards men; see the SCM, Fiske et al., 2002) would be experienced to a lesser extent than after thinking about stereotypic targets (*female nurse, male mechanic*). Moreover, surprise was predicted to explain effects of counter-stereotypic targets on stereotyped emotions.

Experiment 2 aimed to replicate the findings of Experiment 1 using an alternative category conjunction, including national groups (ingroup - Italian *versus* outgroup - Romanian) and occupation categories (manager *versus* window-cleaner). Romanians form a highly discriminated and threatening group in many Western countries and they are perceived as a low competence and low warmth group (Albarello & Rubini, 2011; Bilewicz & Kogan, 2014). Thus, we expected that when Romanians are portrayed in counter stereotypic (*versus* stereotypic) terms participants would experience contempt, the stereotyped emotion elicited by low competence and low warmth group, to a lesser extent. We here also examined an alternative affective outcome in the form of

dehumanization. Growing evidence shows that outgroup members are dehumanized via the attribution to them of less uniquely human emotions and attributes than to ingroup members. In particular, Leyens et al. (2000; 2001; 2003) have repeatedly shown the *infrahumanization* effect, whereby participants attribute fewer secondary emotions (i.e. uniquely human emotions such as, pride, melancholy, shame, embarrassment, remorse, compassion) to outgroup members than to ingroupers (for a review, see Vaes, Leyens, Paladino, Pires, 2012; see also Haslam, Loughnan, Kashima, & Bain, 2008; Struch & Schwartz, 1989). However, while dehumanization represents one of the most robust forms of intergroup bias, little is known of the socio-cognitive processes that can inhibit this phenomenon. An exception is work by Albarello and Rubini (2012; 2015) that recently demonstrated reduced dehumanization of Black people following a multiple categorization intervention. Furthermore, a task in which participants were required to put together surprising conjunctions of social categories promoted the attribution of humanness to members of different outgroups (Prati, Vasiljevic, Crisp, & Rubini, 2015) and intergroup tolerance (Vasiljevic & Crisp, 2013). However, no studies so far have tackled counter-stereotypic category conjunctions as a humanization strategy. In Experiment 2, we examined whether counter-stereotype exposure could humanize the target outgroup considered. We also introduced a brand new methodological approach to measuring humanization. We aimed to investigate whether counter-stereotypic *versus* stereotypic category conjunctions of Romanians would enhance the generation of humanizing *versus* dehumanizing traits (i.e., uniquely human and human nature ones) in depicting the target (Haslam, 2006).

Experiment 3 aimed to go beyond the findings of Experiment 2 using an alternative measure of humanization. It tested whether counter-stereotypic category conjunctions of Romanians would enhance the generation of secondary or uniquely human emotions they are capable of (Leyens et al., 2000). Finally, Experiments 2 and 3 examined the mediating process accounting for the effects of counter-stereotypic conjunctions on the reduction of stereotyped emotions and the humanization of target groups. Specifically, based on research that has highlighted the primary role of emotions in

influencing cognition and behavior (e.g., Frijda et al., 2000; Zajonc, 1980) and on previous research of Hutter and Crisp (2005) who measured surprise as a direct outcome of counter-stereotype exposure, we tested whether the emotion of surprise would lead to the reduced application of stereotyped emotions and the increased humanization of outgroups.

Experiment 1

Experiment 1 tested the role of gender counter-stereotypic *versus* stereotypic category conjunctions on stereotyped emotions. Gender is one of the most chronically accessible categories and it is almost universally applied in impression formation (cf. Bargh & Thein, 1985; Higgins & King, 1981; Stangor, Lynch, Duan, & Glas, 1992). However, this leads to a wide and resilient use of gender stereotypes that are difficult to be eliminated. According to the SCM, females are perceived as warm and incompetent eliciting the emotion of pity, whereas males are perceived as competent and not warm, eliciting envy. The gender counter-stereotype exposure manipulation taken from Hutter & Crisp (2005) relies on unexpected information about female and male occupational roles (*female mechanic and male nurse*). It was thus expected that thinking about gender counter-stereotypic (*female mechanic / male nurse*) *versus* stereotypic category conjunctions (*female nurse / male mechanic*) would diminish the expression of stereotyped emotions; that is, pity towards females and envy towards males (the *stereotyped emotions reduction hypothesis*). Moreover, because mechanic and nurse occupations share the same status[2], no difference on stereotyped emotions between them was expected.

In line with previous findings (Hastie et al., 1990; Hutter and Crisp, 2005) a reduction of stereotypic attributes used to depict female or male targets was expected in the counter-stereotypic *versus* stereotypic category conjunction condition due to the experience of surprise. We therefore hypothesized that surprise would mediate the effects of counter-stereotype exposure on the reduced application of stereotyped emotions; that is, pity and envy but also independently on the reduction of stereotypic attributes used to depict female or male targets[3].

Method

Participants and Design

One hundred and twenty four students at a British University (86 females, 38 males; M_{age} = 17.87) were randomly allocated to either a counter-stereotypic (*female mechanic; male nurse*) or to a stereotypic category condition (*female nurse; male mechanic*).

Procedure

Following Hutter and Crisp's (2005) procedure, participants in the first page of the questionnaire were presented with one of the four category conjunctions which were *female mechanic* or *male nurse* or *female nurse* or *male mechanic* and asked to think about the conjunction for one minute. They were then required to rate the extent to which they experienced surprise about the target and subsequently to generate as many attributes of the target as they could think of. After that, they completed the other dependent variables measures. The target was consistently reported at the beginning of each page in capital letters in order to remind participants.

Dependent variables

Before the attribute generation task, participants rated on a 7-point scale (from 1 = *extremely unlikely* to 7 = *extremely likely*) the extent to which they experienced surprise (comprising surprise and astonishment = .92) about the target. After that, each participant was given precisely 180 seconds to describe his or her spontaneous thoughts as to what characteristics the type of person he or she was presented with might possess (see Table 1). After this they were asked to rate to what extent they felt pity (comprising pity and sympathy = .68), envy (comprising envy and jealousy, = .85), contempt (comprising contempt and disgust = .51) and admiration (comprising admiration and pride = .72) towards the target on a 7-point scale (from 1 = *extremely unlikely* to 7 = *extremely likely*). Emotion words, except for surprise, came from prior Stereotype Content Model studies (see Cuddy et al., 2007; Fiske et al., 2002).

Results and Discussion

Coding of non-stereotypic and stereotypic attributes. To calculate the number of 'non-stereotypic' and 'stereotypic' attributes for the combined categories, the procedure outlined by Hastie et al. (1990; see also Hutter & Crisp, 2005) was used. Accordingly, a *non-stereotypic attribute* was defined as a concept used only when describing the category conjunction and therefore, it should be absent when describing either of the categories considered (e.g., *overweight* for female mechanic; *approachable* for male nurse). In contrast, a *stereotypic attribute* was one used by participants describing either of the categories considered (e.g., *easy-going* for female mechanic; *helpful* for male nurse). Some participants used an attribute or its synonym more than once (e.g., *muscly* and masculine to describe a male mechanic), and such a repetition or synonym was not included in the count. Two independent coders who were blind to the hypotheses performed this redundancy check using a thesaurus to ensure words within each participant's response list that were repeated would count only once. Female stereotypic attributes were considered as non-stereotypic if they were used to describe male category conjunctions and vice versa. The inter-rater agreement on the classification of attributes across coders was acceptable for both non-stereotypic attributes, $r(120) = .88$, $p < .005$, and stereotypic attributes, $r(120) = .81$, $p < .005$. So, we took the mean of the two coders' ratings as indices of generated non-stereotypic ($\alpha = .91$) and stereotypic attributes ($\alpha = .84$).

Non-stereotypic and stereotypic attributes. A 2 (category conjunction: counter-stereotypic, stereotypic) × 2 (target gender: female, male) × 2 (participant gender: female, male) × 2 (attribute: non-stereotypic, stereotypic) ANOVA mixed design with repeated measures on the second factor revealed a main effect of attribute, $F(1, 120) = 147.94$, $p = .001$, $\eta^2 = .55$ qualified by a category conjunction × attribute interaction, $F(1, 120) = 38.01$, $p = .002$, $\eta^2 = .24$. Participants generated fewer stereotypic attributes to describe counter-stereotypic ($M = 0.18$, $SD = 0.12$) compared to stereotypic targets ($M = 1.46$, $SD = 0.10$), $F(1, 120) = 63.07$, $p = .003$, $\eta^2 = .34$. They also generated more non-stereotypic (i.e., courageous for the female target, compassionate for the

male target) attributes to describe counter-stereotypic (M = 4.60, SD = 0.35) compared to stereotypic targets (M = 2.91, SD = 0.31), $F(1, 120) = 12.94$, $p = .005$, $\eta^2 = .09$. Furthermore, there were fewer stereotypic attributes (M = 0.53, SD = 0.10) used to describe female than male targets (M = 1.12, SD = 0.12). No other effect was significant, category conjunction, $F(1, 120) = 0.31$, $p = .58$, $\eta^2 = .00$, target gender, $F(1, 120) = 0.05$, $p = .82$, $\eta^2 = .00$, participant gender, $F(1, 120) = 0.73$, $p = .39$, $\eta^2 = .00$. There were also no significant interactions between category conjunction and target gender, $F(1, 120) = 0.07$, $p = .78$, $\eta^2 = .00$, category conjunction × participant gender, $F(1, 120) = 0.10$, $p = .75$, $\eta^2 = .00$, target gender × participant gender, $F(1, 120) = 0.02$, $p = .90$, $\eta^2 = .00$, category conjunction × target gender × participant gender, $F(1, 120) = 1.17$, $p = .28$, $\eta^2 = .01$.

Stereotyped emotions. Table 2 presents the means, standard deviations for stereotyped emotions based on category conjunctions, target gender and participant gender. To test our hypotheses, 2 (category conjunction: counter-stereotypic, stereotypic) × 2 (target gender: female, male) × 2 (participant gender: female, male) between subjects ANOVAs were conducted on each emotion separately; that is, pity and envy.

Results for *pity*, the emotion related to the female stereotype, revealed a significant effect of target gender, $F(1, 120) = 33.38$, $p = .002$, $\eta^2 = .21$. Overall, participants expressed pity to a higher extent towards female compared to male targets (see Table 2). This main effect was qualified by an interaction between category conjunction and target gender, $F(1, 120) = 4.16$, $p = .03$, $\eta^2 = .04$. In line with predictions, participants reported less pity when they thought about female mechanic (M = 3.78, SD = 1.02) compared to female nurse (M = 4.54, SD = 1.23), $F(1, 57) = 6.39$, $p = .04$, $\eta^2 = .10$. In contrast, there was no significant difference in pity felt towards male nurse (M = 3.00, SD = 1.35) and male mechanic (M = 2.89, SD = 0.92), $F(1, 63) = 0.12$, $p = .73$. There were no significant effects of category conjunction, $F(1, 120) = 0.96$, $p = .33$, $\eta^2 = .00$ nor participant gender, $F(1, 120) = 0.09$, $p = .76$, $\eta^2 = .00$. There were no significant interactions between category conjunction and target gender, $F(1, 120) = 0.77$, $p = .38$, $\eta^2 = .00$, category conjunction × participant gender, $F(1,$

120) = 0.16, p = .68, η^2 = .00, target gender × participant gender, F(1, 120) = 0.17, p = .67, η^2 = .00, or category conjunction × target gender × participant gender, F(1, 120) = 0.33, p = .56, η^2 = .01.

Results for *envy*, the emotion related to the male stereotype, revealed main effects of category conjunction, F(1, 120) = 48.29, p = .003, η^2 = .07 and target gender, F(1, 120) = 0.42, p = .51, η^2 = .01, which were qualified by a category conjunction × target gender interaction, F(1, 120) = 32.71, p = .002, η^2 = .22. Participants who thought about male nurse (M = 3.41, SD = 1.21) reported less envy compared to those who thought about male mechanic (M = 4.00, SD = 0.89), F(1, 63) = 4.63, p = .03, η^2 = .08. In contrast, the female mechanic (M = 4.09, SD = 1.18) elicited higher ratings of envy compared to female nurse (M = 2.71, SD = 0.79), F(1, 57) = 59.04, p = .001, η^2 = .51. No other effect was significant, all F < 1.14, p_s > .07.There was no significant effect of participant gender, F(1, 120) = 1.14, p = .28, η^2 = .01 and there were no significant interactions between category conjunction and target gender, F(1, 120) = 3.58, p = .07, η^2 = .03, category conjunction × participant gender, F(1, 120) = 0.78, p = .37, η^2 = .01, target gender × participant gender, F(1, 120) = 0.20, p = .65, η^2 = .00, or category conjunction × target gender × participant gender, F(1, 120) = 0.42, p = .51, η^2 = .00.

Surprise. Counter-stereotypic targets elicited greater surprise than stereotypic targets, F(1, 120) = 111.18, p < .005, η^2 = .48 (see Table 2). No significant effect of target gender, F(1, 120) = 2.65, p = .10, η^2 = .02, nor participant gender, F(1, 120) = 0.03, p = . 86, η^2 = .00 were obtained. There was no significant interaction between category conjunction and target gender, F(1, 120) = 0.33, p = .56, η^2 = .00, nor category conjunction × participant gender, F(1, 120) = 0.82, p = .36, η^2 = .01, nor target gender × participant gender, F(1, 120) = 0.50, p = .48, η^2 = .00, nor category conjunction × target gender × participant gender, F(1, 120) = 0.52, p = .46, η^2 = .00.

Mediational analysis. Mediational analyses were performed to test the role of surprise on counter-stereotype exposure outcomes. Bootstrapping analyses (5000 re-samples) were conducted using the methods described by Hayes and Preacher (2013). As shown in Table 3, surprise was positively correlated with category conjunctions, so that surprise was higher following counter-

stereotype exposure. It also correlated negatively with stereotypic attributes, pity and envy and positively with non-stereotypic attributes. However, no multicollinearity relation was found among variables as a pre-condition to mediation testing (Blalock, 1963).

First, a mediational analysis tested whether surprise mediated category conjunction effects on the generation of stereotypic attributes related to the targets considered. Category conjunction predicted stereotypic attributes ($\beta = -.42$, $p < .05$) and surprise ($\beta = .68$, $p < .005$). Surprise significantly predicted the dependent variable while controlling for category conjunction ($\beta = -.49$, $p < .005$). The relationship between category conjunction and stereotypic attributes became non-significant ($\beta = -.08$, $p = .42$) when surprise was entered in the same regression. The point estimate for the indirect effect equated to -1.46 ($SE = 0.37$), with the lack of the presence of a zero within the 95% bias corrected confidence intervals (LL: -2.26; UL: -0.77), confirming the mediating role of surprise.

Secondly, a mediational analysis was performed on non-stereotypic attributes. Category conjunction predicted non-stereotypic attributes ($\beta = 3.25$, $p = .001$) and surprise ($\beta = 1.00$, $p = .005$). Surprise significantly predicted the dependent variable while controlling for category conjunction ($\beta = 1.02$, $p = .005$). The relationship between category conjunction and non-stereotypic attributes became non-significant ($\beta = -.14$, $p = .39$) when surprise was entered in the same regression. The point estimate for the indirect effect equated to -1.46 ($SE = 0.38$), (LL: 2.59; UL: 4.09), confirming the mediating role of surprise.

Third, mediational analyses on each gender-related emotion (i.e., pity for female targets and envy for male targets) tested whether surprise mediated the stereotyped emotions of category conjunctions. In particular, it was tested whether surprise mediated the effects of female category conjunctions on pity. Category conjunction predicted pity ($\beta = -.31$, $p = .01$) and surprise ($\beta = .54$, $p < .005$). Surprise significantly predicted pity, while controlling for category conjunction ($\beta = -.63$, $p < .005$). Category conjunction did not predict pity ($\beta = -.05$, $p = .78$), when the mediator was entered in the same regression. The point estimate for the indirect effect equated to -1.21 ($SE =$

0.30), 95% BC CI (-1.94, -0.73) confirming the mediating role of surprise in female category conjunction effects on pity [4].

It was also tested whether surprise mediated the effects of male category conjunctions on the expression of envy. Category conjunction predicted envy (β = -.26, p = .03) and surprise (β = .67, p < .005). Surprise significantly predicted envy, while controlling for category conjunction (β = -.53, p < .005). When the mediator was entered in the same regression with the predictor, category conjunction did not predict envy anymore (β = -.09, p = .51). The indirect effect equated to -1.01 (*SE* = 0.23), 95% BC CI (-1.33, -0.39) confirmed that male category conjunction effects on envy was mediated by surprise. As hypothesized, surprise mediated the impact of gender counter-stereotype exposure on the reduced application of stereotyped emotions. This was the case for both pity and envy separately.

In sum, Experiment 1 revealed that counter-stereotype exposure can produce benefits, not only on cognitive, but also affective outcomes. This first experiment supported the prediction that exposure to gender counter-stereotypes attenuates the stereotyped emotions of the categories considered; that is pity towards females and envy toward males. Furthermore, surprise explained counter-stereotype exposure effects both in terms of cognitive outcomes and the reduced application of stereotyped emotions. Experiment 2 aimed to replicate and extend the generalizability of these findings by considering alternative category conjunctions.

Experiment 2

In Experiment 2, we used a completely different outgroup target: Romanians. Since the entrance of Romania into the European Community many Romanians have moved to other European countries. Some of them have dramatically improved their conditions whereas many others are still striving for better life conditions (Albarello & Rubini, 2011; Bilewicz & Kogan, 2014). Prejudice toward Romanians has became so pervasive that the Romanian Government has financed a press campaign to promote acceptance of Romanians by Italians, highlighting similarities between Romanians and Italians, while differentiating them from other outgroups

(https://socialmediares.wordpress.com/2010/04/07/romania-piacere-di-conoscerti/). In this experiment we focused on this group and adapted methodology used by Prati, Menegatti and Rubini (2015). Specifically, socially relevant ingroup and outgroup national groups (Italian *vs.* Romanian) were combined with occupation categories (car window-cleaner *vs.* manager) to create stereotypic and counter-stereotypic conjunctions. It is worth noting here that car window cleaning on street corners (i.e., by homeless people) is a stereotypical immigrant "job" in the Italian context. As such *Romanian car window-cleaner* and *Italian manager* were the counter-stereotypic conjunctions, whereas *Romanian manager* and *Italian car window-cleaner* were the stereotypic conjunctions[5] .

We expected that thinking about Romanian counter-stereotypic *versus* stereotypic category conjunctions would reduce *contempt*, the stereotyped emotion usually related to low competence and low warmth groups, which is how Romanians are generally perceived (Fiske et al., 2002). Second, we expected that counter-stereotypic *versus* stereotypic category conjunctions of Romanians would reduce dehumanization tendencies via the generation of uniquely human traits in the impression formation phase of the experiment. Moreover, we tested whether surprise, and in turn perceived expectancy violation, would sequentially mediate the effects of counter-stereotypic conjunctions on contempt, and the increased generation of human traits for the outgroup.

Method

Participants and Design

One hundred and nineteen students at a northern Italian University (61 females, 56 males, 2 declined to give gender; age: M = 20.20, SD = .90) were randomly assigned to either a counter-stereotypic (*Romanian manager; Italian window-cleaner*) or to a stereotypic category (*Romanian window-cleaner; Italian manager*) condition.

Procedure

The attribute generation task was identical to that used in Experiments 1 with the exception of the categories used and the implementation of a brand new measure of humanization. Specifically, we adapted the attribute generation task used in Experiment 1 (originally from Hutter

& Crisp, 2005) so that we could code for human traits, rather than non-stereotypic and stereotypic attributes, freely generated at the impression formation phase. In this way we were able to distinguish between dehumanizing *versus* humanizing attributes. As in Experiment 1, before the attribute generation task, participants completed surprise and perceived expectancy violation measures. After this, they completed the remaining dependent variables.

Dependent Variables

The measure of surprise (α = .79) was the same used in Experiment 1. Participants were also asked to indicate to what extent they perceived the target as a) atypical and b) unexpected on scales from 1 (*Not at all*) to 7 (*Very much*). On the basis of these two measures an index of perceived expectancy violation was obtained (α = .65) (Bettencourt, Dill, Greathouse, Charlton, & Mulholland, 1997).

The degree to which participants experience a series of emotions towards the target was measured on 7-point Likert scales ranging from 1 (*not at all*) to 7 (*very much*). On this basis indices of contempt (comprising contempt and disgust = .83), envy (comprising envy and jealousy = .89), pity (comprising pity and sympathy = .68) and admiration (comprising admiration and pride = .67) were computed.[6]

At the end of the experiment, participants completed two items on the extent to which the impression formation task on the target was perceived as a) complex and as b) difficult on a scale ranging from 1 (*Not at all*) to 7 (*Very much*). These measures formed a manipulation check index (α = .79) .

Results and Discussion

Manipulation check. Drawing from previous studies (Hutter & Crisp, 2005) we computed a 2 (category conjunction: counter-stereotypic, stereotypic) × 2 (national group: ingroup - Italian, outgroup - Romanian) between subjects ANOVA that revealed regardless of the national group, $F < 2.46$, $p = .80$, counter-stereotypic category conjunctions ($M = 5.83$, $SD = 0.15$) were perceived as

more complex and difficult to think about compared to stereotypic conjunctions (M = 3.58, SD = 0.15), $F(1, 115) = 101.46$, $p = .005$, $\eta^2 = .47$.

Coding of dehumanizing versus humanizing attributes. Two coders blind to the hypotheses of the experiment were instructed to code the attributes generated by participants in terms of those that were humanizing (that is uniquely human and human nature traits) *versus* dehumanizing (that is not uniquely human and not human nature traits) (see Table 4). Coders were told that the former traits are distinctive of human beings and not shared with other animals nor robots. In contrast, the latter traits are shared with other animals or robots. They were then provided with examples of the two categories drawing from literature on dehumanization (Haslam et al., 2008; Loughnan & Haslam, 2007; Loughnan, Haslam & Kashima, 2009). In order to form clear concept on these two trait categories, before the coding, they were invited to generate examples of attributes distinguishing them between the two categories. They then performed the coding on the attributes generated by participants. Specifically, for each participant they coded the number of humanizing and dehumanizing traits (see Table 4). The inter-coder agreement was acceptable for both categories of traits, r (111) = .83, $p < .005$; r (111) = .75, $p < .005$. The proportions of the two categories of traits were arcsine transformed to normalize the data and remove intercell dependency (for similar ways of handling proportions, see Menegatti & Rubini, 2012, Rubini & Kruglanski, 1997; Semin, Rubini, & Fiedler, 1995).

To test our hypotheses, 2 (category conjunction: counter-stereotypic, stereotypic) × 2 (national group: ingroup - Italian, outgroup – Romanian) between subjects ANOVAs were conducted on dehumanizing *versus* humanizing traits, stereotyped emotions and perceived expectancy violation.[7] Table 5 presents the scale means, standard deviations for main dependent variables based on category conjunctions.

Dehumanizing *versus* humanizing attributes. A 2 (category conjunction: counter-stereotypic, stereotypic) × 2 (national group: Italian, Romanian) × 2 (types of attributes: dehumanizing, humanizing) mixed design ANOVA with repeated measures on the second factor

revealed no main effect of category conjunction, $F(1, 111) = 0.01$, $p = .91$, $\eta^2 = .00$, group memberships, $F(1, 111) = 1.17$, $p = .28$, $\eta^2 = .01$, types of attributes, $F(1, 111) = 0.67$, $p = .41$, $\eta^2 = .00$ and an interaction between category conjunction, group memberships and attributes, $F(1, 111) = 6.30$, $p = .04$, $\eta^2 = .04$. Considering the Romanian targets, a attributes × category conjunction interaction, $F(1, 54) = 7.74$, $p = .02$, $\eta^2 = .12$ showed that humanizing attributes were generated to a greater extent by participants who thought about Romanian manger compared to those who thought about Romanian window-cleaner, $F(1, 35) = 4.75$, $p < .03$, $\eta^2 = .08$ (see Table 5). There was no significant difference in the generation of dehumanizing attributes between Romanian targets, $F(1, 55) = 2.33$, $p = .13$, $\eta^2 = .04$.

Considering the Italians target, there was an effect of attributes, $F(1, 58) = 40.70$, $p = .001$, $\eta^2 = .41$ and no effect of category conjunction, $F(1, 58) = 0.04$, $p = .84$. Results showed that participants attributed humanizing attributes ($M = 3.02$, $SD = 0.26$) to a greater extent than dehumanizing attributes ($M = 1.26$, $SD = 0.17$) to Italians.

Stereotyped emotion. Results for *contempt* revealed main effects of category conjunction, $F(1, 115) = 8.45$, $p = .004$, $\eta^2 = .07$, national group, $F(1, 115) = 9.04$, $p = .003$, $\eta^2 = .07$ and the predicted category conjunction × national group interaction, $F(1, 115) = 7.53$, $p = .007$, $\eta^2 = .06$. As expected, participants reported less contempt towards Romanian manager compared to Romanian window-cleaner, $F(1, 57) = 15.45$, $p = .001$, $\eta^2 = .21$. There was no significant difference between Italian targets, $F(1, 58) = 0.01$, $p = .90$, $\eta^2 = .00$.

Surprise. Results for surprise showed that Romanian and Italian counter-stereotypic conjunctions ($M = 4.83$, $SD = 0.13$) elicited a higher level of surprise than stereotypic conjunctions ($M = 3.63$, $SD = 0.13$), $F(1, 113) = 41.87$, $p = .002$, $\eta^2 = .27$, regardless of the national group, $F(1, 113) = 0.56$, $p = .45$. Surprise was correlated with the emotion of contempt and the humanizing traits (see Table 6).

Perceived expectancy violation. Counter-stereotypic targets (M = 4.07, SD = 0.21) were perceived as more atypical and unexpected than stereotypic targets (M = 3.04, SD = 0.21), $F(1, 115) = 9.49$, $p = .003$, $\eta^2 = .08$, regardless of the national group, $F(1, 115) = 0.15$, $p < .69$.[4]

Mediational analysis. We tested whether the effects of Romanian category conjunctions were sequentially mediated by variations in surprise and expectancy violation. Two sequential mediations were tested first on contempt and second on humanizing traits. To this end, we ran bootstrapping analyses (5000 re-samples) for testing direct and indirect effects in multiple step mediation models as prescribed by Hayes, Preacher, and Myers (2011; see also Taylor, MacKinnon, & Tein, 2008). The MEDTHREE macro for SPSS was employed. Since preliminary intercorrelation analyses revealed that the proposed mediators significantly correlated with the emotion of contempt and the generation of humanizing traits (Table 2) without multi-collinearity relations, these were entered as dependent variables in the mediational analyses. Figure 1 shows the sequential model with surprise and perceived expectancy violation as mediators and ratings of contempt. Romanian category conjunction significantly predicted the first mediator (path 1), $B_1 = 0.75$, $SE_1 = 0.31$ $p <$ .05. The first mediator predicted the second mediator, while controlling for the predictor (path 2), $B_2 = 0.32$, $SE_2 = 0.19$, $p = .02$. In turn, the second mediator predicted contempt, while controlling for category conjunction and the first mediator (path 3), $B_3 = -0.28$, $SE_3 = 0.11$, $p < .05$.

The point of estimate for the sequential indirect effect equated to -0.58 (SE = 0.28), 95% BC CI (-1.26, -0.14). Since zero falls outside this interval, the sequential indirect effect of category conjunction on contempt via surprise and perceived expectancy violation was significant. When the two mediators were included in the regression analysis, the relationship between category conjunction and contempt became non-significant (before inclusion of mediators, path 4: $B_4 = -1.13$, $SE_4 = 0.44$, $p <.05$; after inclusion of mediators, path 4: $B_4 = -0.55$, $SE_4 = 0.43$, $p = .21$).

The same sequential model was estimated entering humanizing attributes as dependent variable (Figure 2). The second mediator significantly predicted humanizing attributes while controlling for category conjunction and the first mediator (path 3), $B_3 = 0.33$, $SE_3 = 0.12$, $p < .005$.

The proposed sequential indirect effect was significant, with the point estimate for the sequential indirect effect equating to 0.74 (SE = 0.34), with 95% BC CI (0.19, 1.50). Including the mediators in the regression analysis the relationship between category conjunction and humanizing attributes became non-significant (before inclusion of mediators path 4: B_4 = 1.01, SE_4 = 0.46, p = .03; after inclusion of mediators path 4: B_4 = 0.26, SE_4 = 0.44, p = .54).

Although alternative models seemed less plausible, we tested whether expectancy violation and surprise sequentially mediated the effects of category conjunction. The analyses showed no significant indirect effect on either contempt or humanizing attributes, 95% BC CIs (-0.30, 0.02), and (-0.27, 0.16), respectively, when expectancy violation and surprise were entered as sequential mediators. Thus, these additional analyses supported the proposed order of the mediators.

This experiment extended the efficacy of counter-stereotype exposure for reducing stereotyped emotions toward minority and stigmatized groups. As predicted, thinking about a Romanian manager reduced contempt, the stereotyped emotion elicited by groups perceived as low in warmth as well as in competence. Furthermore, the analysis of self-generated traits used by participants to describe the targets supported the prediction that counter-stereotypic targets increase the generation of humanizing traits. Results also supported the sequential mediating role of surprise and expectancy violation in explaining the effects of Romanian counter-stereotypic category conjunction on both the reduced experience of contempt and "humanization" of this outgroup target. In Experiment 3, we further explored the "generative" role of counter-stereotype exposure on the uniquely human emotional experiences attributed to the outgroup.

Experiment 3

Experiment 3 aimed to build on the findings of Experiment 2 by testing whether counter-stereotypic *versus* stereotypic category conjunctions of Romanians would enhance the generation of dehumanizing *versus* humanizing characteristics attributed to this outgroup using our novel measure of humanization, but this time based on secondary emotions. As infra-humanization research shows, people tend to perceive outgroup members as less human than ingroup members, attributing

secondary emotions (i.e., emotions that are only distinctive of human beings and not shared with non-human primates) to a less extent to outgroup compared to ingroup members (Demoulin, Leyens, Paladino & Dovidio, 2004; Leyens et al., 2000; 2001). We hypothesized that participants who were asked to think about counter-stereotypic *versus* stereotypic conjunctions of Romanians would generate a higher number of secondary emotions to refer to the outgroup target. Moreover, we tested whether we would replicate the mediational model observed in Experiment 2.

Method

Participants and Design

Seventy-nine students at a northern Italian University (34 females, 35 males; M_{age} = 20.49) were randomly allocated to either a counter-stereotypic (*Romanian manager; Italian window-cleaner*) or to a stereotypic target (*Romanian window-cleaner; Italian manager*).

Procedure

As in Experiment 2, participants completed first the measures of surprise (α = .94) and perceived expectancy violation (α = .77). After that, they completed the attribute generating task and the other dependent variables.

Dependent variables.

The attribute generation task was similar to that used in previous experiments, but participants were asked to think and write down only emotional attributes to describe the target. It was followed by the measures of stereotyped emotions (contempt = .79; envy = .79; pity = .70; admiration = .66)[8] and manipulation check (α = .83) as in Experiment 2.

Results and Discussion

Manipulation check. A 2 (category conjunction: counter-stereotypic, stereotypic) × 2 (national group: Italian, Romanian) between subjects ANOVA revealed that counter-stereotypic conjunctions (*M* = 4.56, *SD* = 0.31) were perceived as more complex and difficult to think about than stereotypic conjunctions (*M* = 3. 43, *SD* = 0.31), $F(1, 73) = 6.62$, $p = .01$, $\eta^2 = .02$, regardless of national group, $F(1, 73) = 0.42$, $p = .46$, $\eta^2 = .00$.

Coding of dehumanizing versus humanizing emotional attributes. Two coders blind to the hypotheses of the experiment were instructed to code the emotional attributes generated by participants in terms of humanizing or secondary (i.e., embarrassment, remorse, melancholy, shame, compassion, pride) and dehumanizing or primary (i.e., pleasure, happiness, desire, fear, pain, and rage). They were provided with a list of primary and secondary emotions which was composed drawing from the literature on infrahumanization (Demoulin et al., 2004; Leyens et al., 2000; Rodríguez-Torres et al., 2005). For example, a Romanian manager was described as proud, suspicious and frustrated, whereas a Romanian window-cleaner was described as hopeless, sad and scared (see Appendix). They coded for each participant the number of humanizing and dehumanizing human emotional attributes generated. The inter-coder agreement was acceptable for both humanizing and dehumanizing emotional attributes, r (72) = .78, $p < .005$; r (72) = .81, $p < .005$. As in Experiment 2, the proportions of humanizing and dehumanizing emotional attributes were arcsine transformed to normalise the data and remove inter-cell dependency.

Dehumanizing *versus* humanizing emotional attributes. A 2 (category conjunction: counter-stereotypic, stereotypic) × 2 (national group: Italian, Romanian) × 2 (emotional attributes: dehumanizing, humanizing) mixed design ANOVA with repeated measures on the second factor was computed. There were no main effects of category conjunction, $F(1, 76) = 0.75$, $p = .35$, $\eta^2 = .00$, national group, $F(1, 76) = 0.90$, $p = .55$, $\eta^2 = .00$, and emotional attributes, $F(1, 76) = 0.90$, $p = .67$, $\eta^2 = .00$, but the analysis produced an emotional attributes × category conjunction × national group interaction, $F(1, 76) = 4.49$, $p < .05$, $\eta^2 = .07$. Further analyses on Romanian targets revealed a main effect of category conjunction, $F(1, 35) = 4.20$, $p < .05$, $\eta^2 = .11$ qualified by an emotional attributes × category conjunction interaction, $F(1, 35) = 5.90$, $p < .05$, $\eta^2 = .14$. As expected, humanizing emotional attributes were generated to a higher extent when the target was Romanian manager ($M = 2.06$, $SD = 0.96$) compared to Romanian window-cleaner ($M = 1.09$, $SD = 1.01$), $F(1, 35) = 8.55$, $p < .05$, $\eta^2 = .19$. There was no significant difference between conditions in the generation of dehumanizing emotional attributes (counter-stereotypic, $M = 1.80$, $SD = 0.67$;

stereotypic, M= 1.90, SD = 0.92), $F(1, 35) = 0.15$, $p = .69$, $\eta^2 = .00$. Furthermore, no significant effect was found for Italian targets, nor category conjunction, $F(1, 35) = 0.66$, $p = .42$, $\eta^2 = .00$ nor emotional attributes, $F(1, 35) = 0.07$, $p = .78$, $\eta^2 = .00$.

As in Experiment 2, 2 (category conjunction: counter-stereotype, stereotype) x 2 (national group: Italian, Romanian) ANOVAs were computed on contempt, surprise and perceived expectancy violation.

Stereotyped emotion. Results for *contempt*, the stereotyped emotion usually elicited by exposure to low warmth and low competence groups, revealed main effects of category conjunction, $F(1, 76) = 7.29$, $p = .03$, $\eta^2 = .09$ and national group, $F(1, 76) = 16.65$, $p = .001$, $\eta^2 = .18$ qualified by a category conjunction × national group interaction, $F(1, 76) = 6.21$, $p = .02$, $\eta^2 = .07$. In line with Experiment 2, participants reported less contempt towards Romanian manager ($M = 2.50$, $SD = 1.33$) compared to Romanian window-cleaner ($M = 4.07$, $SD = 1.50$), $F(1, 35) = 7.59$, $p = .02$, $\eta^2 = .18$. There was no significant difference between Italian targets (window-cleaner, $M = 2.02$, $SD = 0.75$; manager, $M = 2.08$, $SD = 1.04$), $F(1, 41) = 0.05$, $p = .82$, $\eta^2 = .00$.

Surprise. Regardless of national group, $F(1, 76) = 0.47$, $p = .49$, $\eta^2 = .00$, counter-stereotypic targets ($M = 5.18$, $SD = 0.26$) elicited a much higher level of surprise than stereotypic targets ($M = 3.62$, $SD = 0.27$), $F(1, 77) = 16.78$, $p < .005$, $\eta^2 = .18$.

Perceived expectancy violation. Counter-stereotypic targets ($M = 4.71$, $SD = 0.23$) were perceived as more atypical and unexpected than stereotypic targets ($M = 3.82$, $SD = 0.23$), $F(1, 75) = 7.44$, $p < .05$, $\eta^2 = .09$ regardless of national group, $F(1, 75) = 0.54$, $p = .46$, $\eta^2 = .01$.

Mediational analysis. Inter-correlation analyses showed that surprise correlated with contempt, expectancy violation and humanizing emotional attributes (see Table 3), but there were no multicollinearity relations among variables. As in Experiment 2, we tested the sequential mediating role of surprise and perceived expectancy violation in accounting for the effects of category conjunctions of Romanians first on contempt (Figure 3) and second on humanizing emotional attributes (Figure 4).

In the first sequential mediation, category conjunction predicted surprise (path 1), $B_1 = 1.41$, $SE_1 = 0.35$, $p < .005$. Surprise predicted perceived expectancy violation, while controlling for category conjunction (path 2), $B_2 = 0.31$, $SE_2 = 0.09$, $p < .005$. Finally, perceived expectancy violation predicted contempt while controlling for category conjunction and surprise (path 3), $B_3 = -0.29$, $SE_3 = 0.12$, $p = .02$. The point of estimate for the sequential indirect effect of category conjunction on contempt equated to -0.48 (SE = 0.21), with 95% BC CI (-0.97, -0.11). Including the two mediators in the regression analysis, the relationship between category conjunction and contempt became non-significant (before inclusion of mediators, path 4: $B_4 = -1.17$, $SE_4 = 0.34$, $p = .005$; after inclusion of mediators, path 4: $B_4 = -0.59$, $SE_4 = 0.35$, $p = .06$.

In the second mediation analysis, perceived expectancy violation predicted humanizing emotional attributes, while controlling for category conjunction and surprise (path 3), $B_3 = 0.32$, $SE_3 = 0.10$, $p < .005$. The point of estimate for the sequential indirect effect was significant, with the point estimate equating to 0.42 (SE = 0.18), with 95% BC CI (0.09, 0.80). Including the two mediators in the regression analysis, the relationship between category conjunction and humanizing emotional attributes became non-significant (before inclusion of mediators, path 4: $B_4 = 0.70$, $SE_4 = 0.28$, $p = .02$; after inclusion of mediators, path 4: $B_4 = 0.25$, $SE_4 = 0.28$, $p = .38$).

As before, we tested the alternative sequential mediational models; that is, to test whether expectancy violation and surprise sequentially mediated the effects of category conjunction. The analyses showed no significant indirect effect on either contempt or humanizing attributes, 95% BC CIs (-0.10, 0.02) and (-0.03, 0.11), respectively, when expectancy violation and surprise were entered as sequential mediators. Thus, these additional analyses supported the proposed order of the mediators.

The data gathered in this experiment brought further support to results of Experiment 2 by showing that counter-stereotypic information on Romanians attenuated contempt. The data also revealed that the counter-stereotype of Romanians promoted the generation of humanizing emotional attributes. In a convergent manner with the previous experiment evidence showed that

the reduction of contempt but also the generation of humanizing emotions to the target is explained by the sequential mediation of surprise and expectancy violation.

General Discussion

In three experiments we demonstrated, for the first time, the beneficial emotional and humanizing outcomes of counter-stereotypic *versus* stereotypic category conjunctions and the mediation processes explaining these effects. Specifically, Experiment 1 showed that gender counter-stereotype exposure attenuated pity and envy, the stereotyped emotions usually experienced towards females and males respectively. In addition, the study confirmed that counter-stereotypic conjunctions lead to the generation of non-stereotypic attributes in comparison to stereotypic attributes, indicating a shift from a heuristic mode of thinking. Moreover, this was the first experiment to reveal a mediating role of surprise in explaining these positive effects of counter-stereotype exposure. Using a novel measure of humanization, Experiment 2 demonstrated that the counter-stereotypic exposure to Romanians heightened the application of human traits in impressions formed of them. We also observed a lesser tendency to attribute the stereotyped emotion of contempt towards them. Experiment 3 replicated and extended the results of Experiment 2 by showing that the counter-stereotype exposure heightened the generation of humanizing emotional attributes to Romanians. Across these experiments it was consistently found that surprise, elicited by the counter-stereotype of Romanians, leads to awareness of expectancy violation that in turn explained the reduction in stereotyped emotions attributed to them. It is noteworthy that the beneficial effects of counter-stereotype exposure were consistent across the different target groups used (gender in Experiment 1; national group membership in Experiments 2 and 3), suggesting this social cognitive intervention is generalizable with potentially wide appeal.

Implications for theory and research on intergroup emotions

This contribution goes beyond conventional models of cognition and emotion (Fiske et al., 2002) to show how multiple categorization, specifically counter-stereotypic conjunctions, can inhibit the attribution of stereotyped emotions towards the groups at hand. It should be also noted that across the three experiments, the stereotype dependent emotions revealed good alphas (also considering that the variables were composed of only two items, Cortina, 1993; Schmitt, 1996) and their ratings were around the mid-point of each scale showing a good reliability of the obtained indices (Smith, Seger & Mackie, 2007). Furthermore, according to the intergroup emotion theory (Mackie et al., 2000), intergroup emotions predict behaviours. In this vein, challenging intergroup emotions through counter-stereotypic category conjunctions should improve intergroup attitudes and behaviours. Future research should extend the evidence we have gathered by testing whether intergroup emotions can explain counter-stereotypic category conjunction effects, not only at the judgemental level, but also in terms of intergroup behaviour.

Implications for theory and research on dehumanization

This evidence extends the role of counter-stereotype exposure as a prejudice-reduction strategy (Hutter & Crisp, 2005) by showing that this intervention can lead to humanization of outgroup targets. We would also highlight how the combination of these two areas of research has produced a methodological advance: a new attribute generation measure of humanization. Whereas previous studies on dehumanization have adopted a fixed response format, in these studies we used an attribute generation task to code for dehumanizing *versus* humanizing characteristics. Moreover, this measure revealed that the generative role of counter-stereotypes was consistent across different dimensions of humanness, that is, humanizing traits (i.e., uniquely human and human nature traits, Haslam 2006) in Experiment 2 and humanizing emotions (i.e., secondary emotions, Leyens et al., 2000) in Experiment 3. Taken together these findings show that dehumanizing and humanizing characteristics are components of self and other definition that individuals freely use to form impressions following counter-stereotype exposure. This is a remarkable finding considering the extensive evidence for the pervasiveness of prejudicial thinking.

Counter-stereotypic conjunctions may provide a important tool for tackling one of the most heinous forms of prejudice as dehumanization.

The interplay of surprise and expectancy violation

In Experiment 1 surprise was shown to drive the effects of counter-stereotype (vs. stereotype) exposure on pity and envy as well as the increased use of non-stereotypic attributes used to describe the targets. In Experiments 2 and 3 we observed, and replicated, a sequential mediational of surprise and perceived expectancy violation for explaining the effects of counter-stereotypes on contempt and the generation of (de-)humanizing traits or (de-)humanizing emotions. The test of alternative models with perceived expectancy violation as the first mediator in the sequential mediation was not significant in both studies and on all dependent variables whether stereotyped emotions or (de-)humanizing traits.

Taken together these results speak to the important role of surprise as an emotion that can trigger accuracy in the impression formation process, which can help to attenuate pre-existing biased emotions and representations of the target at hand. Moreover, although we relied on cross-sectional data the evidence we collected highlights the role of surprise on the cognitive process of expectancy violation in motivating individuals to solve cognitive discrepancies elicited by counter-stereotypes. Indeed, as the alternative sequential models were not significant, this converges with previous studies demonstrating the motivating and eliciting role of affective processes on social cognition (Frijda, et al., 2000; Zajonc, 1980). To go deeper into the understanding of the interaction between emotional and cognitive factors underlying counter-stereotypic conjunctions effects, future research could test the relationship between the cognitive outcome of generating attributes and the stereotyped emotions elicited by the categories at stake. Moreover, even if the status of the occupation categories used did not influence category conjunctions effects (as shown in footnote 5), future research should replicate these findings using occupation categories that do not differ in terms of status (e.g., teacher and nurse).

Conclusions

In this research we showed that being surprised by unexpected category conjunctions prompts individuals to think more thoroughly about the ‘essence’ of the target at hand, including attributing uniquely human characteristics and less biased emotions towards them. Taken together these effects suggest counter-stereotype exposure has benefits for intergroup relationships and interactions. In our world, exposure to social and cultural diversity is increasing due to the progressive breaking down of traditional, geographical and social boundaries that have characterized the previous century. This research shows that in this social climate, not only it is possible to meet counter-stereotypic targets, such as a female mechanic or a Romanian manager, but doing so may play an important role in enhancing social mobility and in driving social change.

References

Albarello, F., & Rubini, M. (2011). Outgroup Projection: Il caso degli stereotipi negativi attribuiti a Rom e Rumeni. [Outgroup Projection: Negative Stereotypes of Roms and Romanians]. *Psicologia Sociale, 6,* 355–366.

Albarello, F., & Rubini, M. (2012). Reducing dehumanization outcomes towards Blacks: The role of multiple categorization and of human identity. *European Journal of Social Psychology, 42,* 801–938.

Albarello, F., & Rubini, M. (2015). The role of reduced humanity in producing linguistic discrimination. *Personality and Social Psychology Bulletin,* 41, 224–236.

Bargh, J. A., & Thein, R. D. (1985). Individual construct accessibility, person memory, and the recall-judgment link: The case of information overload. *Journal of Personality and Social Psychology, 49,* 1129–1146.

Bettencourt, B., Dill, K. E., Greathouse, S. A., Charlton, K., & Mulholland, A. (1997). Evaluations of ingroup and outgroup members: The role of category-based expectancy violation. *Journal of Experimental Social Psychology, 33,* 244–275.

Bilewicz, M., & Kogan, A. (2014). Embodying imagined contact: Facial feedback moderates the intergroup consequences of mental simulation. *British Journal of Social Psychology, 53,* 387–395.

Blalock, H. M. (1963). Correlated independent variables: The problem of multicollinearity. *Social Forces, 42,* 233–237.

Clark, M. S. & Fiske, S. T., (1982). *Affect and Cognition: The 17th Annual Carnegie Mellon Symposium on Cognition.* Hillsdale, NJ: Erlbaum.

Cortina, J. M. (1993). What is coefficient alpha? An examination of theory and applications. *Journal of Applied Psychology, 78,* 98–104.

Crisp, R. J., & Hewstone, M. (2007). Multiple social categorization. *Advances in Experimental Social Psychology, 39,* 163–254.

Crisp R. J., & Turner, R. (2011). Cognitive adaptation to the experience of social and cultural diversity. *Psychological Bulletin, 137,* 242-266.

Cuddy, A. J. C., Fiske, S. T., & Glick, P. (2007). The BIAS Map: Behaviors from intergroup affect and stereotypes. *Journal of Personality and Social Psychology, 92,* 631–648.

Cuddy, A. J. C., Norton, M. I., & Fiske, S. T. (2005). This old stereotype: The pervasiveness and persistence of the elderly stereotype. *Journal of Social Issues, 61,* 265–283.

Demoulin, S., Leyens, J. P., Paladino, M. P., Rodriguez-Torres, R., Rodriguez-Perez, A., & Dovidio, J. F. (2004). Dimensions of "uniquely" and "non-uniquely" human emotions. *Cognition and Emotion, 18,* 71-96.

Eagly, A. H., & Steffen, V. J. (1984). Gender stereotypes stem from the distribution of women and men into social roles. *Journal of Personality and Social Psychology, 46,* 735–754.

Ekman, P. (1972). Universals and cultural differences in facial expressions on emotion. In J. Cole (Ed.), *Nebraska symposium on motivation,* (pp. 207-283). Lincoln: University of Nebraska Press.

Fiedler, K. & Forgas, J. P., eds. (1988). *Affect, Cognition, Social Behavior*. Toronto: Hogrefe.

Fisk, J. E. (2002). Judgments under uncertainty: Representativeness or potential surprise? *British Journal of Psychology, 93,* 431–449.

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.

Forgas, J. P. (1991). *Emotion and Social Judgments.* Oxford: Pergamon Press.

Forgas, J. P. (2000). *Feeling and Thinking.* Cambridge: Cambridge University Press.

Frijda, N. H., Manstead, A. S., & Bem, S. (2000). *Emotions and beliefs: How feelings influence thoughts*. Cambridge University Press.

Gocłowska, M. A., & Crisp, R. J. (2013). On counter-stereotypes and creative cognition: When interventions for reducing prejudice can boost divergent thinking. *Thinking Skills and Creativity, 8,* 72-79.

Haslam, N. (2006). Dehumanization: An integrative review. *Personality and Social Psychology Review, 10,* 252-264.

Haslam, N., Loughnan, S., Kashima, Y., & Bain, P. (2008). Attributing and denying humanness to others. *European Review of Social Psychology, 19,* 55-85.

Hastie, R., Schroeder, C., & Weber, R. (1990). Creating complex social conjunction categories from simple categories. *Bulletin of the Psychonomic Society, 28,* 242–247.

Hayes, A. F., & Preacher, K. J. (2013). Conditional process modeling: Using structural equation modeling to examine contingent causal processes. *Structural equation modeling: A second course, 2.*

Hayes, A. F., Preacher, K. J., & Myers, T. A. (2011). Mediation and the estimation of indirect effects in political communication research. *Sourcebook for political communication research: Methods, measures, and analytical techniques,* 434-465.

Higgins, E. T., & King, G. (1981). Accessibility of social constructs: Information-processing consequences of individual and contextual variability. In N. Cantor & J. F. Kihlstrom (Eds.), *Personality, Cognition, and Social Interaction* (pp. 69-121). Hillsdale, N J: Erlbaum.

Hutter, R. R. C., & Crisp, R. J. (2005). The composition of category conjunctions. *Personality and Social Psychology Bulletin, 31,* 647-657.

Hutter, R. R. C., & Crisp, R. J. (2006) Implications of cognitive busyness for the perception of category conjunctions. *Journal of Social Psychology, 146,* 253-256.

Hutter, R. R. C., Crisp R. J., Humphreys G. W., Waters G. M., & Moffitt G. (2009). The Dynamics of Category Conjunctions. *Group Processes Intergroup Relations, 12,* 673-686.

Izard, C. E. (1977). *Human emotions*. New York: Plenum.

Kunda, Z., Miller, D. T., & Claire, T. (1990). Combining social concepts: The role of causal reasoning. *Cognitive Science, 14,* 551–577.

Leicht, C., de Moura, G. R., & Crisp, R. J. (2014). Contesting gender stereotypes stimulates generalized fairness in the selection of leaders. *The Leadership Quarterly, 25,* 1025-1039.

Leyens, J. P., Paladino, M. P., Rodriguez, R. T., Vaes, J., Demoulin, S., Rodriguez, A. P., et al. (2000). The emotional side of prejudice: The role of secondary emotions. *Personality and Social Psychology Review, 4,* 186–197.

Leyens, J. P., Rodriguez-Perez, A., Rodriguez-Torres, R., Gaunt, R., Paladino, M. P., Vaes, J., & Demoulin, S. (2001). Psychological essentialism and the differential attribution of uniquely human emotions to ingroups and outgroups. *European Journal of Social Psychology*, *31*, 395–411.

Loughnan, S., & Haslam, N. (2007). Animals and Androids Implicit Associations Between Social Categories and Nonhumans. *Psychological Science*, *18*, 116-121.

Loughnan, S., Haslam, N., & Kashima, Y. (2009). Understanding the relationship between attribute-based and metaphor-based dehumanization. *Group Processes & Intergroup Relations*, *12*, 747-762.

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.

Maguire, R., Maguire, P., & Keane, M. T. (2011). Making sense of surprise: an investigation of the factors influencing surprise judgments. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*, 176-186.

Martin, L. L. & Clore, G. L., eds. (2001). *Theories of Mood and Cognition.* Mahwah, NJ: Erlbaum.

Menegatti M., & Rubini M. (2012). From the individual to the group: The enhancement of linguistic bias, *European Journal of Social Psychology*, *42*, 36–40.

Meyer, W. U., Reisenzein, R., & Schützwohl, A. (1997). Towards a process analysis of emotions: The case of surprise. *Motivation and Emotion, 21*, 251–274.

Moscatelli S., Albarello, F., Prati., F., & Rubini, M. (2014). Badly off or better off than them? The impact of relative deprivation and relative gratification on implicit and overt intergroup discrimination. *Journal of Personality and Social Psychology*, *107,* 248-264.

Ortony, A., & Partridge, D. (1987). Surprisingness and expectation failure: What's the difference? In *Proceedings of the 10th International Joint Conference on Artificial Intelligence* (pp. 106–108). Los Altos, CA: Kaufmann.

Pettigrew, T. F., & Tropp, L. R. (2008). How does intergroup contact reduce prejudice? Meta-analytic tests of three mediators. *European Journal of Social Psychology*, *38*, 922-934.

Plutchik, R. (1980). *Emotion: A Psychoevolutionary Synthesis* (p. 440). New York: Harper & Row.

Prati, F., Menegatti, M., & Rubini, M. (2015). Reducing Linguistic Outgroup Derogation: A Matter of Multiple Categorization and Intergroup Contact. *Journal of Language and Social Psychology.* Doi: 0261927X14567777.

Prati, F., Moscatelli, S., Pratto, F., & Rubini, M. (2015). Predicting Support for Arabs' Autonomy from Social Dominance: the Role of Identity Complexity and Dehumanization. *Political Psychology.* Advance online publication.

Prati, F., Vasiljevic, M., Crisp, J.R., & Rubini, M. (2015). Some Extended Psychological Benefits of Challenging Social Stereotypes: Decreased Dehumanization and a Reduced Reliance on Heuristic Thinking. *Group Processes and Intergroup Relations*. Doi:1368430214567762.

Pratto, F., Sidanius, J., Stallworth, L. M., & Malle, B. F. (1994). Social dominance orientation: A personality variable predicting social and political attitudes. *Journal of Personality and Social Psychology*, *67*, 741-763.

Ray, D. G., Mackie, D. M., Smith, E. R., & Terman, A. W. (2012). Discrete emotions elucidate the effects of crossed-categorization on prejudice. *Journal of Experimental Social Psychology, 48,* 55-69.

Reisenzein, R. (2009). Emotional experience in the computational belief–desire theory of emotion. *Emotion Review*, *1*, 214-222.

Roccas S., & Brewer M.B. (2002). Social identity complexity. *Personality and Social Psychology Review, 6,* 88-106.

Rodríguez-Torres, R., Leyens, J. P., Pérez, A. R., Rodriguez, V. B., Quiles del Castillo, M. N., Demoulin, S., & Cortés, B. (2005). The lay distinction between primary and secondary emotions: A spontaneous categorization?. *International Journal of Psychology*, *40,* 100-107.

Roseman M. (1996). Managing complexity in TeamRooms, a Tcl-based Internet groupware application. In *Proceedings of the 1996 Tcl/Tk Workshop*, Usenix Press.

Rubini M., & Kruglanski, A. W. (1996). Brief encounters ending in estrangement: motivated language use and interpersonal rapport in the question-answer paradigm. *Journal of Personality and Social Psychology, 72,* 1047–1060.

Schmitt, N. (1996). Uses and Abuses of Coefficient Alpha. *Psychological Assessment,* 8, 350-353.

Schützwohl, A., & Reisenzein, R. (1999). Children's and adult's reactions to a schema discrepant event: A developmental analysis of surprise. *International Journal for Behavioral Development, 23,* 37-62.

Semin, G. R., Rubini, M., & Fiedler, K. (1995). The answer is in the question: The effect of verb causality on locus of explanation. *Personality and Social Psychology Bulletin*, *21*, 834-841.

Smith, E.R., Seger, C.R., & Mackie, D.M. (2007). Can Emotions Be Truly Group Level? Evidence Regarding Four Conceptual Criteria. *Journal of Personality and Social Psychology*, *93*, 431–446.

Stangor, C., Lynch, L., Duan, C., & Glas, B. (1992). Categorization of individuals on the basis of multiple social features. *Journal of Personality and Social Psychology*, *62*, 207-218.

Struch, N., & Schwartz, S. H. (1989). Intergroup aggression: its predictors and distinctness from in-group bias. *Journal of Personality and Social Psychology*, *56*, 364-373.

Taylor A.B., & MacKinnon, D.P., Tein, J-Y. (2008). Tests of the three-path mediated effect. *Organizational Research Methods*, *11*, 241–269.

Vaes, J., Leyens, J. P., Paladino, M.P., & Pires, M. M. (2012). We are human, they are not: Driving forces behind outgroup dehumanisation and the humanisation of the ingroup. *European Review of Social Psychology*, *23*, 64-106.

Vanbeselaere, N. (1987). The effect of dichotomous and crossed social categorizations upon intergroup discrimination. *European Journal of Social Psychology*, *17*, 143-156.

Vasiljevic M., Crisp, R.J. (2013). Tolerance by Surprise: Evidence for a Generalized Reduction in Prejudice and Increased Egalitarianism through Novel Category Combination. *PLoS ONE* 8(3): e57106.

Zajonc, R. B. (1980). Feeling and thinking: Preferences need no inferences. *American Psychologist*, *35*, 151-175.

Footnotes

[1] Across the three experiments, stereotypic categorization was used as baseline condition, basing on the accepted view that stereotypic thinking is the default mode of person perception (Fiske & Neuberg, 1990). Previous work on counter-stereotypes has established that this is the case. Vasiljevic and Crisp (2013) found no difference on cognitive flexibility between a stereotypic categorization and a no categorization condition compared to a counter-stereotypic categorization condition.

[2] A pilot test showed no significant status difference between nurse (M = 2.38, SD = 1.58) and mechanic (M = 2.79, SD = 1.09) occupations, $t(43) = 3.84$, $p = .001$.

[3] The reason we focused on the reduction of stereotypic attributes is because these are the measure of reduced heuristic thinking specified by Crisp & Turner (2011) as indicative of reduced reliance on the heuristic thinking system. Non-stereotypic or emergent attributes should be affected by this process, but are theoretically incidental to the resulting reduced application of stereotyped emotions. Furthermore, for completeness we measured all four stereotyped emotions illustrated in the SCM (Fiske et al., 2002). However, no difference on contempt and admiration between gender category conjunctions was expected, since these two emotions are not usually elicited by the category conjunction at stake. Findings on these additional emotions supported this expectation. Participants reported less *contempt* after thinking about counter-stereotypic (M = 3.71, SD = 1.01) compared to stereotypic (M = 4.33, SD = 0.95) targets, $F(1, 120) = 20.38$, $p = .004$, $\eta^2 = .14$ independently of target gender (M_{male}= 4.13, SD = 0.90, M_{female} = 3.90, SD = 1.01), $F(1, 120) = 0.00$, $p = .94$, $\eta^2 = .00$ and participant gender (M_{male} = 4.26, SD = 0.86, M_{female} = 3.95, SD = 0.98), $F(1, 120) = 2.82$, $p = 009$, $\eta^2 = .02$. There were no significant interactions between category conjunction × target gender, $F(1, 120) = 0.02$, $p = .86$, $\eta^2 = .00$, category conjunction × participant gender, $F(1, 120) = 0.00$, $p = .94$, $\eta^2 = .00$, target gender × participant gender, $F(1, 120) = 1.90$, p =

.17, η^2 = .02, category conjunction × target gender × participant gender, $F(1, 120) = 0.16$, $p = .68$, η^2 = .00.

No significant effects on *admiration* were found for category conjunction ($M_{\text{counter-stereotypic}}$ = 3.42, SD = 1.35, $M_{\text{stereotypic}}$ = 3.03, SD = 1.10), $F(1, 120) = 2.17$, $p = .14$, $\eta^2 = .02$, target gender (M_{male} = 2.97, SD = 1.21, M_{female} = 3.30, SD = 1.30), $F(1, 120) = 1.29$, $p = .25$, $\eta^2 = .01$, participant gender (M_{male} = 3.14, SD = 1.18, M_{female} = 3.13, SD = 1.29), $F(1, 120) = 0.01$, $p = .90$, $\eta^2 = .00$ (see Table 2). There were also no significant interactions between category conjunction × target gender, $F(1, 120) = 0.12$, $p = .72$, $\eta^2 = .00$, category conjunction × participant gender, $F(1, 120) = 2.79$, $p = .09$, $\eta^2 = .02$, target gender × participant gender, $F(1, 120) = 0.39$, $p = .53$, $\eta^2 = .00$, category conjunction × target gender × participant gender, $F(1, 120) = 0.74$, $p = .39$, $\eta^2 = .00$.

[4] Given that gender counter-stereotypic conjunctions also reduced the negative emotion of contempt towards both males and females, for further information, we tested the mediating role of surprise in the effect of counter-stereotypic conjunctions on this emotion. Results showed that category conjunction predicted contempt ($\beta = 3.20$, $p = .001$) and also surprise predicted contempt ($\beta = 0.20$, $p = .001$). However, the relationship between category conjunction and contempt was still significant ($\beta = 0.51$, $p = .02$) when surprise ($\beta = 0.03$, $p = .54$) was entered in the same regression. The point estimate for the indirect effect equated to 0.9 (SE = 0.17), (LL: -.23; UL:.44), indicating no mediation. This confirmed our hypothesis that surprise explained only the reduction of stereotyped emotions related the categories considered (such as pity and envy) and not general negative emotions such as contempt.

[5] An additional study was conducted to rule out the possibility that differential status of the occupational categories we used (window cleaner vs. manager) could explain the effects observed our dependent measures. To do this we carried out a replication of the main experiment, but also measured perceived status as a covariate..Forty four students at a northern Italian University (27 females, 17 males,; age: M = 22.50, SD = .90) were randomly assigned to either a counter-stereotypic (*Romanian manager; Italian window-cleaner*) or to a stereotypic category (*Romanian*

window-cleaner; Italian manager) condition. Participants were assigned to either a counter-stereotypic (*Romanian manager; Italian window-cleaner*) or to a stereotypic category (*Romanian window-cleaner; Italian manager*) conjunction condition and then they were asked to describe the target presented. After this, they were asked to complete the measures of surprise and humanizing traits as in the main experiment. Here, however, they were also asked to indicate on a 7-point Likert scale the status attributed to the occupations categories considered. A 2 (category conjunction: counter-stereotypic, stereotypic) × 2 (group memberships: ingroup - Italian, outgroup - Romanian) ANCOVA, including status as a covariate, replicated the main experiment's main effect of category conjunction, $F(1, 40) = 19.36$, $p = .00$, $\eta^2 = .32$ and revealed no effect of manager status, $F(1, 40) = 0.72$, $p = .40$, $\eta^2 = .01$, no effect of window-cleaner status, $F(1, 40) = 0.62$, $p = .43$, $\eta^2 = .01$ and no interactions, $F(1, 40) = 6.30$, $p = .04$, $\eta^2 = .04$. Moreover, the mediating roles of surprise and status on the effect of multiple categorization on generation of humanizing traits was tested. Category conjunction predicted humanizing traits, $B = 0.95$, $SE = 0.37$, $p = .01$. Category conjunction was also related to surprise, $B = 1.37$, $SE = 0.40$, $p = .001$ and status, $B = -.72$, $SE = 0.31$, $p = .02$. When surprise and status were added to the regression of category conjunction on humanization, the predictor became not significant, $B = 0.26$, $SE = 0.42$, $p = .53$ and surprise was still significant, $B = 0.47$, $SE = 0.13$, $p = .001$, whereas status was not, $B = -0.06$, $SE = 0.17$, $p = .70$. The point of estimate for the sequential indirect effect of category conjunction on contempt equated to 0.003 (SE = 0.03), with 95% BC CI (-0.04, 0.10). In sum, the findings of this experiment suggest that status does not play a role in explaining the effects of counter-stereotype exposure revealed in the current research.

[6] For the sake of completeness, we measured all four stereotyped emotions based on SCM (Fiske et al., 2002).

Envy revealed significant effects of category conjunction, $F(1, 115) = 34.17$, $p = .000$, $\eta^2 = .22$ and group memberships, $F(1, 115) = 20.05$, $p = .000$, $\eta^2 = .14$ qualified by category conjunction × group memberships interaction, $F(1, 115) = 58.50$, $p = .000$, $\eta^2 = .33$. Italian window-cleaner ($M = 1.71$,

SD = 1.07) elicited less envy compared to Italian manager (*M* = 4.70, *SD* = 1.28), *F*(1, 58) = 93.62, *p* = .001, η^2 = .61 and no significant difference was found between Romanian targets ($M_{\text{window-cleaner}}$ = 2.01, *SD* = 1.54, M_{manager} = 2.41, *SD* = 0.77), *F*(1, 57) = 1.58, *p* = .21, η^2 = .02.

Results for *pity* demonstrated a main effect of category conjunction, *F*(1, 115) = 4.36, *p* = .03, η^2 = .04 qualified by a category conjunction × group memberships interaction, *F*(1, 115) = 8.13, *p* = .005, η^2 = .06. Romanian manager (*M* = 2.45, *SD* = 1.36) elicited less pity compared to Romanian window-cleaner (*M* = 3.70, *SD* = 1.66), *F*(1, 57) = 10.10, *p* = .005, η^2 = .15. There was no significant differences between Italian targets ($M_{\text{window-cleaner}}$ = 2.80, *SD* = 1.14, M_{manager} = 2.60, *SD* = 1.31), *F*(1, 58) = 0.36, *p* = .54, η^2 = .00.

Results for *admiration* revealed no main effect of category conjunction, *F*(1, 115) = 1.80, *p* = .18, η^2 = .01 nor of group memberships, *F*(1, 115) = 0.61, *p* = .43, η^2 = .00 but an interaction between category conjunction and group memberships, *F* (1, 115) = 7.12, *p* < .05, η^2 = .06. Admiration was elicited to a higher extent by Romanian manager (*M* = 2.75, *SD* = 1.12) compared to Romanian window-cleaner (*M* = 1.82, *SD* = 1.34), *F*(1, 57) = 8.17, *p* = .005, η^2 = .12 and no significant difference was found between Italian targets ($M_{\text{window-cleaner}}$ = 2.30, *SD* = 1.08, M_{manager} = 2.62, *SD* = 1.29), *F*(1, 58) = 1.07, *p* = .30, η^2 = .01.

[7] To rule out the impact of participant gender a series of 2 (category conjunction: counter-stereotypic, stereotypic) × 2 (group memberships: ingroup - Italian, outgroup – Romanian) × 2 (participant gender: female, male) between subjects ANOVAs on dehumanizing *versus* humanizing traits, stereotyped emotions and perceived expectancy violation were conducted. Results showed no significant effect of participant gender on dehumanizing *versus* humanizing traits, *F*(1, 113) = 0.90, *p* = .34, η^2 = .00, nor on contempt, *F*(1, 113) = 1.06, *p* = .30, η^2 = .01, pity, *F*(1, 113) = 0.00, *p* = .95, η^2 = .00, envy, *F*(1, 113) = 0.01, *p* = .90, η^2 = .00, admiration, *F*(1, 113) = 0.18, *p* = .66, η^2 = .00, surprise, *F*(1, 113) = 0.02, *p* = .88, η^2 = .00, nor on expectancy violation, *F*(1, 115) = 0.00, *p* = .94, η^2 = .00.

[8] Results for *envy* showed main effects of category conjunction, $F(1, 76) = 23.01$, $p = .001$, $\eta^2 = .23$ and group memberships, $F(1, 76) = 35.62$, $p = .001$, $\eta^2 = .31$ qualified by a category conjunction × group memberships interaction, $F(1, 76) = 50.57$, $p = .001$, $\eta^2 = .40$. Italian window cleaner ($M = 1.68$, $SD = 0.65$) elicited less envy compared to Italian manager ($M = 3.97$, $SD = 1.04$), $F(1, 41) = 77.27$, $p = .005$, $\eta^2 = .65$. Romanian manager ($M = 1.90$, $SD = 0.55$) and Romanian window-cleaner ($M = 1.45$, $SD = 0.55$) showed no significant difference in ratings of envy, $F(1, 35) = 2.47$, $p = .125$, $\eta^2 = .06$.

Results for *pity* showed no significant effects of category conjunction, $F(1, 76) = 0.56$, $p < .45$, $\eta^2 = .00$ and group membership, $F(1, 76) = 0.16$, $p = .69$, $\eta^2 = .00$ but a category conjunction × group memberships interaction, $F(1, 76) = 13.55$, $p < .005$, $\eta^2 = .15$. Romanian manager ($M = 2.20$, $SD = 0.43$) evoked less pity compared to Romanian window-cleaner ($M = 3.56$, $SD = 0.28$), $F(1, 35) = 9.26$, $p = .004$, $\eta^2 = .20$. Conversely, Italian window-cleaner ($M = 3.46$, $SD = 0.27$) evoked higher pity compared to Italian manager ($M = 2.55$, $SD = 0.32$), $F(1, 41) = 4.59$, $p = .03$, $\eta^2 = .10$.

Results for *admiration* revealed no significant effects of category conjunction, $F(1, 76) = 0.41$, $p = .25$, $\eta^2 = .00$ and group membership, $F(1, 76) = 2.82$, $p = .09$, $\eta^2 = .03$ and an interaction between category conjunction and group memberships, $F(1, 76) = 25.81$, $p = .03$, $\eta^2 = .25$. Romanian manager led participants to feel higher admiration towards this outgroup ($M = 2.80$, $SD = 0.26$) compared to Romanian window-cleaner ($M = 1.38$, $SD = 0.22$), $F(1, 35) = 16.69$, $p = .01$, $\eta^2 = .32$. Italian window-cleaner led participants to feel less admiration ($M = 1.96$, $SD = 0.22$) compared to Italian manager ($M = 3.05$, $SD = 0.26$), $F(1, 41) = 9.92$, $p = .02$, $\eta^2 = .19$.