Two routes toward compensation: An investigation into the mechanisms of compensation for high- and low-status groups

Laurent Cambon, Vincent Yzerbyt

Running head: Two routes toward compensation

Two routes toward compensation:

An investigation into the mechanisms of compensation for high- and low-status groups

Laurent Cambon

Université Côte d'Azur, France

Vincent Y. Yzerbyt

Université catholique de Louvain, Louvain-la-Neuve, Belgium

Please send correspondence to:

Laurent Cambon
Université Côte d'Azur
Laboratoire d'Anthropologie et de Psychologie Cognitive et Sociale
24 avenue des Diables Bleus
F-06357 Nice cedex 4 France
Email: laurent.cambon@unice.fr

or to

Vincent Yzerbyt
Université catholique de Louvain
Institute of Research in Psychological Sciences
Place du Cardinal Mercier 10
B-1348 Louvain-la-Neuve, Belgium
Email : vincent.yzerbyt@uclouvain.be

Abstract

In two experiments, we tested an account of compensation in intergroup perceptions as the manifestation of social creativity and magnanimity strategies for low- and high-status groups, respectively. For low-status groups, compensation would be a way to enhance self-esteem after a negative comparison. In contrast, for high-status groups, compensation would provide a means to appear non-discriminatory, in accordance with contemporary norms. Experiment 1 manipulated participants' need to protect their self-esteem by giving them or not giving them an opportunity to self-affirm. As predicted, this self-affirmation manipulation only affected low-status group members. Specifically, the ability to self-affirm protected their self-esteem from an upward comparison and precluded the emergence of compensation. Experiment 2 manipulated the pressures toward non-discrimination by activating a non-discrimination vs. an honesty norm. As expected, this norm manipulation only affected high-status group members. Compared to the non-discrimination norm, the honesty norm decreased the perceived pressures toward non-discrimination and paved the way to ingroup favoritism on both dimensions rather than compensation. We discuss the role and complementarity of these two routes in producing the compensation pattern.

Key words: compensation, competence, warmth, intergroup relations, ingroup bias.

Two routes toward compensation:

An investigation into the mechanism of compensation for high- and low-status groups

Compensation in intergroup relations corresponds to the fact that group members favor their ingroup on one of the two fundamental dimensions of social perception, either competence or warmth, while favoring the other group on the second dimension (Yzerbyt, Provost, & Corneille, 2005). Specifically, low-status (high-status) groups are perceived and characterize themselves as warmer (more competent) but less competent (warm) than high-status (low-status) groups (Bettencourt, Dorr, Charlton, & Hume, 2001; Cambon, Yzerbyt, & Yakimova, 2015; Judd, James-Hawkins, Kashima, & Yzerbyt, 2005; for a review, see Yzerbyt, 2016). Drawing on Social Identity Theory, compensation theorists interpreted this judgment pattern as the manifestation of social creativity and magnanimity strategies for low- and high-status groups, respectively (Kervyn, Judd & Yzerbyt, 2010; Yzerbyt, Kervyn & Judd, 2008). To date, however, the empirical efforts testing the motivational mechanisms underpinning these strategies have been sparse and mostly correlational (but see Yzerbyt & Cambon, 2017). Rather than relying on mediational approaches, the present studies provide experimental evidence of the existence of these two mechanisms underlying compensation.

Social creativity and magnanimity

Compensation in intergroup relations emerges in the context of harmonious relations between groups characterized by a subjectively legitimate status difference (Cambon et al., 2015; Yzerbyt et al., 2005). In such a situation, both groups seem to be satisfied with their standing. After all, the two parties involved find a way to achieve a decent level of positivity on one dimension while admitting to be outperformed by the outgroup on the other dimension. Yzerbyt and colleagues (2008) interpreted compensation as a social creativity strategy[1] (more precisely, a compensating bias, Tajfel, 1974; see also Lemaine, 1974) to re-establish positive social identity and self-esteem for low-status groups and as a magnanimity strategy to secure positive self-regard for high-status groups. In the first

case, although low-status or otherwise disadvantaged groups have to acknowledge ingroup inferiority on the status-related dimension, namely competence, they seek to achieve positive distinctiveness and to bolster their esteem on the other fundamental dimension, namely warmth (Niens & Cairns, 2003; Tajfel & Turner, 1979, van Knippenberg, 1978). For the high-status or advantaged groups, to the extent that their superior status is secured, they may want to appear non-discriminatory by giving up superiority on every dimension and instead showing outgroup bias on the status-irrelevant dimension (Bettencourt, Dorr, Charlton & Hume, 2001; Leach, Snider, & Iyer, 2002; Vanbeselaere, Boen, Van Avermaet, & Buelens, 2006).

Social creativity: The low-status groups' route toward compensation

In the first empirical demonstration of the compensation effect in intergroup relations, Yzerbyt et al. (2005) explicitly defined compensation as a social creativity strategy (Mummendey & Schreiber, 1983). Their study as well as subsequent research regularly revealed the presence of patterns of judgment suggesting that this strategy may indeed be at work. As a matter of fact, low-status groups often show the joint emergence of an outgroup bias on the status-related dimension (i.e., competence) and an ingroup bias on the other dimension (i.e., warmth) (Judd et al., 2005; Kervyn, Yzerbyt, Demoulin, & Judd, 2008). This is especially the case when group members see the status difference between the groups as stable and legitimate (Cambon et al., 2015; Cambon & Yzerbyt, 2016).

Although suggestive, the observed means and correlations remained silent as to the actual mechanisms underlying this creativity strategy. To explore this issue in an experimental context, Yzerbyt and Cambon (2017) tested whether low-status group members would continue to show compensation if they were denied the possibility to rate themselves more favorably than the outgroup on warmth. As predicted, participants who did not have an opportunity to shine on the warmth dimension failed to show compensation. This pattern confirms that the possibility for low-status group members to secure positive self-regard is a key factor in the emergence of compensation and

suggests that self-esteem is at the heart of compensation judgments expressed by low-status group members. In line with SIT, we argue here that following a comparison with members of high-status groups, members of low-status groups will have little choice but to feel outperformed on the status-related dimension, that is, competence. Because this inferiority threatens their self-esteem, members of low-status groups will be tempted to favor their group on the second fundamental dimension, that is, warmth, in order to restore a proper level of self-regard. This account casts self-esteem as both an instigator and a consequence of compensation.

Although no direct evidence established the role of self-esteem in compensation, some efforts provide additional indirect support in favor of this idea. A first line of evidence comes from studies using an experimental paradigm simulating a social creativity strategy and its impact on self-esteem (Becker, 2012; Derks, van Laar, & Ellemers, 2007, 2009; Spencer-Rodgers, Major, Forster, & Peng, 2016). In this research, participants are generally threatened (or not) on an important dimension before being confronted with information affirming (or not) their superiority on another dimension. Note that participants never directly describe their group and another group on judgmental dimensions. Rather, they encountered facts threatening them on one dimension and, then, information reassuring them on another dimension. Although these situations are not strictly comparable to situations of compensation, they are interesting because they parallel Yzerbyt and Cambon's (2017) pattern of findings: The acknowledgment on the ingroup's weakness on one dimension and the affirmation of its distinctiveness on another dimension. Using this kind of paradigm, Becker (2012, but see also Derks et al., 2007, 2009) showed that the social creativity manipulation led to enhanced personal self-esteem. In a slightly different scenario, Spencer-Rodgers et al. (2016) showed that a group affirmation task (writing on important group values), that came after a threat exposure, effectively buffered personal self-esteem. The results also suggested that the threat negatively affected participant's self-esteem and that the affirmation task bolstered it. Interesting as these results may be, they do not deal with the relation between self-esteem and a search for positive

distinctiveness by means of trait attributions on the fundamental dimensions, which is at the heart of our hypothesis.

A second line of evidence resides in studies on stereotyping, and most notably in Fein and Spencer's (1997, Expt. 3) work on the self-esteem maintenance role of stereotypes. These authors gave negative feedback to some of their participants and found that their self-esteem predicted the use of a stereotype to describe a target in a second 'unrelated' experiment. Moreover, the application of this stereotype improved participants' self-esteem. Importantly, other participants who received positive feedback showed no bias and their esteem did not change during the course of the experiment. Branscombe and Wann (1994) obtained similar results in an intergroup context. After having been threatened, their American participants derogated Russians in a personality description task. Moreover, (collective) self-esteem was both an antecedent and a consequence of derogation in that more derogation followed from low collective self-esteem and collective self-esteem improved after derogation.

Taken together, and in line with other work (Abrams & Hogg, 1988; Lemyre & Smith, 1985), these two lines of research suggest that self-esteem is both an instigator and an outcome of the search for positive distinctiveness through trait attribution between groups. In other words, people tend to compensate negative standing on one dimension with positive standing on a second dimension. Still, despite the evidence supporting the self-esteem hypothesis, one may also wonder why low-status group members should use the compensation strategy when other strategies are available. For instance, people could discount the dimension on which their group is outperformed by high-status outgroup (Crocker & Major, 1989; Schmader & Major, 1999). More radically, they could simply negate the reality of the superiority of the outgroup (Blanz, Mummendey, Mielke, & Klink, 1998; Cadinu & Cerchioni, 2001; Mackinnon, Smith, & Carter-Rogers, 2015).

We would argue that these alternative strategies are probably more difficult to implement when the relative status of the groups, and by way of consequence the competence dimension, is at

stake. Indeed, empirical evidence show that people have hard time depreciating personally and socially relevant dimensions (Major, Testa, & Bylsma, 1991; Tesser, Millar & Moore, 1988; Rosenberg, 1979) especially when it comes to cognitive competence (Harter, 1986). Moreover, Yzerbyt and Cambon (2017) found that competence comes across as more objective and is more difficult to disconfirm than warmth. In other words, the reality constraints are markedly stronger when it comes to status-related dimensions than to other dimensions, making it difficult for people to deny an attribution of (lack of) competence (see also, Tausch, Kenworthy, & Hewstone, 2007). Overall, thus, compensation comes across as a strategy of choice for members of low-status groups.

Social magnanimity: the high-status groups' route toward compensation

There are several reasons to argue that the social creativity found among low-status group members and in which self-esteem plays a pivotal role does not operate in the same way for members of high-status groups. First, the self-esteem of high-status group member is rarely suffering, as these groups generally possess sufficient prestige or resources. In fact, some negative feedback could even produce an increase in self-esteem among members of the high-status groups (see the "Failure-as-an-Asset" effect, Reinhard, Stahlberg, & Messner, 2009). Second, several lines of work suggest that while the competence dimension is rather objectively tied to the status of groups (Cuddy et al., 2008; Fiske, 2015; Yzerbyt & Cambon, 2017) fewer reality constraints prevail for warmth (Tausch et al., 2007). This allows high-status groups to challenge the reputation of superiority on warmth on the part of low-status groups if they want. The more 'subjective' interpretation of warmth should allow the latter not to feel threatened by a claim of superiority on warmth by the former.

What could then be the mechanism underlying the compensation pattern manifested by high-status group members? Following Yzerbyt and colleagues (Yzerbyt et al., 2005, 2008), we propose that some form of magnanimity may be at work. In essence, when members of high-status group are reassured with respect to their superiority on the dimension of comparison, i.e., the difference in competence between high- and low-status groups comes across as legitimate and intergroup relations

are not conflictual, they would be ready to concede superiority of the low-status group members on dimensions irrelevant to the status comparison. This pattern is close to what is known as the noblesse oblige effect in the power literature (Sachdev & Bourhis, 1985; Vanbeselaere, Boen, Van Avermaet, & Buelens, 2006).

The operating force behind such magnanimity is rather straightforward. Given the strong normative pressures that prevail in our societies regarding the (non-)expression of discrimination (Plant & Devine, 2001; Dambrun & Guimond, 2004), high-status group members may feel embarrassed to express an ingroup bias on all possible dimensions. Such a blanket judgment of superiority is likely to send out an image of intolerance and lack of understanding. To avoid this, members of high-status groups may choose to restrict their superiority to the one dimension that counts for them, indeed the one on which they consider their superiority undeniable. According to this rationale, compensation on the secondary dimension of warmth corresponds to a strategy used by member of high-status groups in order to appear non-discriminatory. In other words, the way high-status group members understand the norms of discrimination lays at the heart of their readiness to compensate.

We thus argue that the perception of their superiority in status, and by way of consequence on the dimension of competence, leads high-status group members to feel strong pressures toward non-discrimination. This, in turn, should affect positively their willingness to manifest outgroup bias on warmth. As it happens, the mediational role of pressures toward non-discrimination in compensation has been found in a number of studies (Cambon et al., 2015; Cambon & Yzerbyt, 2016; Owuamalam, Wong, & Rubin, 2016; Yzerbyt & Cambon, 2017; see also Jeffries, Hornsey, Sutton, Douglas, & Bain, 2012, for similar arguments). However, the available evidence remains correlational in nature. Here, we propose an experimental test of the role played by pressures toward non-discrimination in the emergence of compensation among high-status group members.

The present experiments

Using the minimal group paradigm, we conducted two experiments that relied on a scenario known to produce a compensation pattern of judgments, that is, asymmetrical status relations in the absence of conflict (Cambon et al., 2015; Cambon & Yzerbyt, 2016; Yzerbyt & Cambon, 2017; for a review, see Yzerbyt, 2016). Specifically, after participants received bogus feedback on the social status of their ingroup relative to an outgroup, they had the opportunity to describe their ingroup as well as the outgroup on a set of traits pertaining to warmth and competence. Experiment 1 tested our social creativity, i.e., self-esteem, hypothesis whereas Experiment 2 tested our magnanimity, i.e., norms of non-discrimination, hypothesis. Specifically, and building upon the 'testing-process-by-interaction-strategy' (Jacoby & Sassenberg, 2011; Sigall & Mills, 1998), we propose here an experimental test of the role played by these two different factors, i.e., self-esteem and pressures toward non-discrimination, in the emergence of compensation among low-status and high-status group members, respectively.

Experiment 1 aimed at manipulating participants' need to protect their self-esteem by turning to compensation. In line with the idea that compensation serves an affirmation function (for a related argument on social creativity, see Derks et al., 2007, 2009), we did or did not give participants the possibility to self-affirm before describing both groups. We predicted that participants given the opportunity to self-affirm would not need to compensate because the affirmation manipulation would have boosted their self-esteem. In contrast, participants given no possibility to self-affirm should need to protect their self-esteem by turning to compensation. We predicted that these effects would take place only among low-status group members for two reasons. First, because the feedback should not threaten high-status group members' self-esteem, we did not expect the latter to experience a need to restore it. Second, because high-status group members' self-esteem is likely to be high and because there is evidence that self-affirmation fails to affect high self-esteem participants (During & Jessop, 2015, Jaremka, Bunyan, Collins, & Sherman, 2011), we predicted that self-affirmation would not affect high-status group members.

Experiment 2 aimed to manipulate the pressures toward non-discrimination by activating either a non-discrimination or an honesty norm. We hypothesized that the activation of the non-discrimination norm would make high-status participants aware of the pressures toward non-discrimination. Consequently, they should manifest compensation. In contrast, activating an honesty norm should attenuate the awareness of these pressures and make high-status groups less reluctant to express ingroup bias on both fundamental dimensions. We predicted that these effects would only occur for high-status group members because their advantageous position in the experimental setting should make them more sensitive to the norms toward non-discrimination.

Experiment 1

We manipulated participants' need to protect their self-esteem via compensation. We did this by either giving them or not giving them the opportunity to self-affirm before presenting them with a negative comparison feedback and a group rating task. Importantly, we took great care to measure participants' self-esteem[2] on three occasions, namely before the experiment, after the status manipulation, i.e., after the reception of the comparative feedback, and after the group rating task. Because the status manipulation was designed to lead to compensation[3], high-status participants were expected to manifest compensation in all conditions and their level of self-esteem should not be altered at any time. In contrast, because the threat induced by the status manipulation targets their group, we predicted that low-status group members would be affected by the affirmation manipulation. Specifically, in the no affirmation condition, participants' self-esteem should be threatened by the status difference. Therefore, these participants should rely on compensation in order to restore their esteem. In sharp contrast, and as predicted by self-affirmation theory (Cohen & Sherman, 2014), the affirmation task should protect low-status group members' esteem. As a result, these participants should feel less of a need to rely on compensation when rating the groups.

The experiment tested these hypotheses at the level of participants' mean responses. We expected to observe compensation in all conditions except for the low-status group members who had

a chance to self-affirm. In this condition, we predicted outgroup favoritism on competence but compared to the low-status, no-affirmation condition, less ingroup favoritism on warmth. Importantly, as far as self-esteem is concerned, participants in the low-status, no-affirmation, condition, should show a decrease in their self-esteem right after the reception of the feedback followed by an increase in their self-esteem after compensation. No such dip and rebound in self-esteem should emerge in the low-status, affirmation, condition or for the high-status groups. Accordingly, we predicted a three-way interaction between a quadratic contrast involving the successive three self-esteem measures and our two independent variables, i.e., affirmation and status.

Method

Participants and design

In light of earlier work on compensation (for a review, see Yzerbyt, 2016), power calculations led us to aim for 80 participants. In total, 100 psychology students participated for partial course credit and data collection did not depend on any analysis of results. Three cases had to be omitted because participants doubted the manipulation and twelve were dropped because they failed to complete the self-esteem scale one week before the experiment. The final sample consisted of $N=85$. Ages ranged between 18 and 39 years ($M=20.1$; $SD=4.22$; females=71%). Participants were randomly assigned to a 2 (manipulation of self-affirmation: self-affirmation vs. no self-affirmation) x 2 (status: high vs. low) between-participants design.

Procedure

One week before the experiment, a male experimenter asked the students of a class (N=284) to complete a state and a trait self-esteem scale along with other filler questions as part of a study on students' health. After completion, a female experimenter asked them to register for an unrelated experiment on small group interactions scheduled one week later. On the day of the experiment, participants came in groups of either four or six persons. Upon their arrival, the female experimenter introduced the experiment as a study on group discussion and the role of competence in small group

interactions. First, participants underwent the self-affirmation manipulation by means of a questionnaire about values that the experimenter presented as a way to provide a conversation topic for the upcoming group discussion. Participants then completed a bogus test in order to assess their competence. A pretest ensured that the test came across as a plausible competence test and was ambiguous enough to prevent participants from making a realistic estimate of their score.

After the experimenter had allegedly scored the test, she provided public feedback about participants' competence so that participants could know each other's scores. One half of the participants in the session received low scores whereas the other half received high scores and they were categorized into one of two groups according to their score. The two groups were then led to two distinct areas in the room and asked to fill in different questionnaires. Specifically, they completed (a) a state self-esteem questionnaire, (b) ratings of the ingroup and the outgroup on competence and warmth traits, (c) a second version of the state self-esteem questionnaire allegedly because the first version was the wrong one, (d) a question checking the manipulations of status, (e) a question about the perceived legitimacy of the competence test, and, finally, (f) a question assessing the objectivity of competence and warmth traits. All measures used in this experiment are reported below. Participants were then debriefed, thanked, and dismissed.

Materials

We manipulated self-affirmation by presenting participants with a series of 10 values selected to be independent of competence (i.e., an exciting life, liberty, curiosity, respect for tradition, etc.) and asking them to rank them from the most to the least preferred value. Participants then had to write a short statement about either their most preferred or their least preferred one, for the self-affirmation and the no self-affirmation conditions, respectively (see, Napper, Harris & Epton, 2009). At the end of the experiment, we checked for the effectiveness of this manipulation by asking four independent judges to evaluate if the text written by participants was (1) or was not (0) related to the selected value.

Measures

Self-esteem measures. The trait self-esteem scale, taken one week before the experiment during the recruitment phase, was Tafarodi and Swann's (1995) self-liking and self-competence scale. The state self-esteem scale consisted of Heatherton and Polivy's (1991) performance (i.e., "I feel confident about my abilities") and social (i.e., "I feel displeased with myself") items[4]. We used these scales because they each break down self-esteem in components close to warmth (i.e., self-liking and social) and competence (i.e., self-competence and performance). We introduced this distinction for exploratory purpose with the idea that, as the threat to self-esteem for low-status participants concerned their competence, the competence related component of self-esteem might be more affected than the warmth related component. Participants completed the latter scale on three occasions: (a) during the recruitment phase alongside the trait self-esteem scale, (b) during the experiment just after the competence feedback, and (c) after the ingroup and outgroup description phase. The necessity of a second completion of the questionnaire during the experimental phase was justified by a problem with the first version in which some items (those related to appearance) were absent. All questionnaires used 5-point scales ranging from 1 (=*strongly disagree*) to 5 (=*strongly agree*).

Dependent variable. Participants rated the ingroup and the outgroup on 12 traits, 6 of which pertained to warmth (nice, pleasant, sensitive, sincere, sociable, warm) and 6 to competence (ambitious, competent, efficient, intelligent, hardworking, self-assured) on scales ranging from 1 (= *not at all*) to 9 (= *totally*).

Other measures. One question checked the effectiveness of our status manipulation by asking participants the extent to which they thought that the ingroup had received a lower (=1) versus higher (=9) competence score than the outgroup. To assess the legitimacy of the status attribution (and as an index of a denial strategy, see MacKinnon et al. 2015), we asked participants to rate the face validity of the competence test on a 9-point scale ranging from 1 (= has no validity) to 9 (= has high validity).

Finally, we checked for the perceived objectivity of competence over warmth attributions by asking whether they thought the attribution of competence was more subjective versus more objective than the attribution of warmth on a scale ranging from 1 (=competence attributions are more subjective than warmth attributions) to 9 (=competence attributions are more objective than warmth attributions).

Results

We computed a global score for each dimension and each group by averaging the ratings on the six competence traits and on the six warmth traits for both the ingroup and the outgroup. Cronbach's alphas ranged from .81 to .85. The Cronbach's alphas of the performance and social components of each of the three measures of Heatherton and Polivy's self-esteem scale ranged from .65 to .91.

Preliminary analyses

We first checked whether participants proved responsive to our affirmation manipulation. The four judges agreed on all cases and assigned a 1, i.e., the written text is representative of the selected value, to all participants thus confirming the success of the manipulation.

We then submitted participants' answers to the status question to a 2 (self-affirmation: self-affirmation vs. no self-affirmation) x 2 (status: high vs. low) between-subjects design[5]. The status main effect, $F(1,81)=330.57$, $p<.001$, $\eta_p^2=.80$, confirmed the success of our manipulation. Participants perceived themselves as having higher status in the high ($M=7.33$, $SD=0.98$) than in the low ($M=2.93$, $SD=1.20$) status conditions.

We submitted participants' perception of the legitimacy of the competence test to the same ANOVA as above. Not surprisingly, the analysis revealed a status main effect, $F(1,81)=7.86$, $p<.01$, $\eta_p^2=.08$, showing that low-status participants perceived the test as less legitimate ($M=6.22$, $SD=2.18$) than high-status participants ($M=7.28$, $SD=1.17$). This effect was qualified by a significant

interaction, $F(1,81)=6.04$, $p<.01$, $\eta_p^2=.07$, indicating that this status effect emerged in the affirmation condition, $F(1,81)=14.03$, $p<.001$, but not in the no-affirmation one, $F<1$.

Finally, we checked whether participants saw competence as more objective than warmth. In line with earlier findings, the score was significantly higher than 5, the midpoint of the scale, both in general ($M=7.22$, $SD=1.23$, $t(84)>16.69$, $p<.0001$) and in every condition (all $Ms>6.87$, all ts>5.91), confirming that competence was perceived as more objective than warmth. Moreover, we observed no impact of our independent variables or their interaction (all $ps>.25$).

Compensation

We submitted the trait ratings to a 2 (affirmation: affirmation vs. no-affirmation) x 2 (status: high vs. low) x 2 (target group: ingroup vs. outgroup) x 2 (dimension: competence vs. warmth) mixed-model ANOVA with the first two factors varying between participants and the last two within them (see Table 1 for means and standard deviations). Confirming our hypothesis, and on top of several significant lower-order effects, the four-way interaction proved significant, $F(1,81)=21.25$, $p<.001$, $\eta_p^2=.21$. Compensation emerged in every condition except for low-status group participants who self-affirmed. In this condition, participants exhibited outgroup favoritism on competence while failing to display ingroup favoritism on warmth.

To examine more directly the patterns of judgments as a function of affirmation, status, and target group, we re-ran the analysis on the scores of ingroup favoritism by subtracting the ratings of the outgroup from those of the ingroup separately for competence and warmth. From the analysis of the competence ingroup favoritism scores, the main effect of status proved significant, $F(1,81)=221.98$, $p=.0001$, $\eta_p^2=.73$, with high-status groups showing ingroup favoritism ($M=1.96$, $SD=0.89$), $t(38)=13.74$, $p<.0001$, and low-status groups outgroup favoritism ($M=-1.48$, $SD=1.21$), $t(45)=-8.32$, $p<.0001$. The affirmation by status interaction was moderately significant, $F(1,81)=2.84$, $p=.10$, $\eta_p^2=.03$. High-status groups showed the same level of ingroup favoritism whether in the affirmation ($M=1.88$, $SD=0.95$) or in the no-affirmation ($M=2.05$, $SD=0.84$) condition, $F(1,37)<1$, *ns*.

Interestingly, low-status groups tended to show less outgroup favoritism in the affirmation (M=-1.18, SD=1.16) than in the no-affirmation (M=-1.78, SD=1.21) condition, $F(1,44)$=2.98, p<.10, η_p^2=.06.

Turning to warmth, the significant main effects of status, $F(1,81)$=182.06, p<.0001, η_p^2=.69, and affirmation, $F(1,81)$=29.75, p<.0001, η_p^2=.27, were qualified by the predicted significant interaction, $F(1,81)$=38.01, p<.0001, η_p^2=.32. As we expected, high-status groups showed similar levels of outgroup favoritism whether in the affirmation (M=-1.34, SD=1.11) or in the no-affirmation (M=-1.50, SD=0.84) condition, $F(1,37)$<1, ns. More importantly, low-status groups showed markedly less ingroup favoritism in the affirmation (M=0.29, SD=0.68) than in the no-affirmation (M=2.88, SD=1.32) condition, $F(1,44)$=69.79, p<.0001, η_p^2=.61. In sum, and in line with our hypothesis, the manipulation of affirmation only affected the judgments expressed by low-status groups and essentially on the warmth dimension.

Although our experimental approach lead us to cast our predictions at the level of the means, we also wanted to explore patterns of compensation at the correlational level. Building upon earlier compensation work (Judd et al., 2005, Yzerbyt et al., 2005), we conjectured that compensation would manifest itself with a negative correlation between ingroup favoritism on competence and ingroup favoritism on warmth. We thus hoped to find this pattern of compensation for high-status group members as well as for low-status group members in the no-affirmation condition. In the affirmation, low-status, condition, we expected no correlation between the two kinds of favoritism. As a matter of fact, and corroborating the analyses on the mean ratings in a somewhat different way, a significant negative correlation confirmed the presence of compensation in all conditions (all rs>-.45) except for low-status groups who had the possibility to self-affirm (r=-.18, ns).

Self-esteem

We submitted our three successive measures of self-esteem[6] to a 2 (manipulation of self-affirmation: affirmation vs. no-affirmation) x 2 (status: high vs. low) x 3 (occasion: time1 vs. time2 vs. time3) mixed-model ANOVA with the first two factors varying between participants and the last

within them (see Table 2). As it happens, every single effect came out significant. Most importantly, however, the three-way interaction was highly significant, $F(1,162)=19.18$, $p<.001$, $\eta_p^2=.18$. To probe this interaction in a way that addressed our specific prediction, we created a linear contrast, comparing the first and the last self-esteem measure (with codes -1 for time1 and +1 for time3), and a quadratic contrast (with codes +1/2 for time1, -1 for time2, and +1/2 for time3) and submitted these to a 2 (manipulation of self-affirmation: affirmation vs. no-affirmation) x 2 (status: high vs. low) factorial ANOVA. As for the linear contrast, there was only a significant affirmation main effect, $F(1,81)=6.54$, $p<.02$, $\eta_p^2=.07$, suggesting that the self-esteem of affirmation participants ($M=0.32$, $SD=0.50$) increased slightly more over the course of the experiment than what was the case for no-affirmation participants ($M=0.02$, $SD=0.44$).

Turning to our quadratic contrast, both the affirmation and the status main effects proved significant. Specifically, participants manifested a stronger decrease in self-esteem at time2 compared to the other measurement occasions when they were not given a chance to affirm ($M=0.62$, $SD=0.84$) than when they had an opportunity to do so ($M=0.02$, $SD=0.37$), $F(1,81)=27.85$, $p<.0001$, $\eta_p^2=.26$. Also, low-status participants showed a stronger decrease ($M=0.62$, $SD=0.82$) than high-status participants ($M=-0.05$, $SD=0.27$), $F(1,81)=41.80$, $p<.0001$, $\eta_p^2=.34$. More importantly, and in line with our predictions, the two-way interaction qualified these main effects, $F(1,81)=32.80$, $p<.0001$, $\eta_p^2=.29$ (see table 2). Follow-up analyses confirmed that our affirmation manipulation had no impact among high-status participants, $t(81)<1$, ns. Quite a difference picture emerged among low-status participant. In this case, affirmation had a major effect such that the self-esteem of no-affirmation participants plummeted at time 2 ($M=1.19$, $SD=0.74$) compared to affirmation participants ($M=-0.07$, $SD=0.15$), $t(81)=8.12$, $p<.0001$.

Again, for exploratory purposes, we also checked the presence of significant correlations between self-esteem and group favoritism in the no-affirmation, low-status group, conditions. Specifically, we expected (a) a negative correlation between self-esteem measured right after the

feedback and ingroup favoritism on warmth and (b) a positive correlation between self-esteem measured following compensation and ingroup favoritism on warmth. According to our rationale, no such pattern should obtain in the other three conditions. None of the correlations reached significance for the high-status groups (rs<.36) or for the low-status groups who had the opportunity to self-affirm. For low-status groups who did not have the opportunity to self-affirm, the correlations proved significant but only for the competence component of self-esteem, $r(23)=-.43$, $p<.04$, and $r(23)=.44$, $p<.04$, for the first and the second change score, respectively.

Discussion

The rationale guiding Experiment 1 was that, among low-status groups, compensation finds its roots in a desire to secure positive distinctiveness for one's group. In the context of an unfavorable comparison on the competence dimension, warmth offers some comfort as to one's worth. When such positive distinctiveness is being obtained via alternative means, there should be less of a need to claim superiority on warmth. In other words, social creativity strategies should appear less necessary and compensation should not emerge. To test this hypothesis, Experiment 1 did or did not provide an opportunity to self-affirm to participants whose self-esteem had been measured one week before the experiment. Next, participants were informed that they belonged to an incompetent, i.e., low status, group or to a competent, i.e., high-status, group. We measured their self-esteem before asking them to evaluate both groups and to fill in the self-esteem scale one last time.

We predicted that the members of low-status groups who had not been given a chance to self-affirm would first show a decrease in their self-esteem compared to the initial self-esteem measure, particularly on the competence component of self-esteem. We also expected that these same participants would then show a compensatory pattern of judgments and, in turn, an increase on the final self-esteem measure. In sharp contrast, we predicted that the low-status groups given a chance to self-affirm would not show such diminished self-esteem after the competence feedback. More crucially, we expected to see no compensation among these participants and, of course, no need to

change their level of self-esteem on the final measure. Moreover, and in line with the hypothesis that self-esteem is a key factor underlying compensation only for low-status groups, we did not expect any impact of our self-affirmation manipulation on high-status groups.

Our data provided strong support to all these predictions. Moreover, complementing the findings on the mean ratings, our exploratory correlational analyses provide additional insight regarding the phenomenology of our participants and indeed corroborates our social creativity hypothesis. The more low-status group no-affirmation participants were hurt in their self-esteem, the more they manifested ingroup favoritism on the warmth dimension by denying the relevant traits to the high-status outgroup. In addition, the more they demonstrated ingroup favoritism on warmth, the more their self-esteem improved. It should be noted that the correlations with ingroup favoritism on warmth only appeared on the performance component of self-esteem even though the ANOVA revealed no differences between the two components.

If the present data go a long way to show that the preservation of self-esteem is at the heart of the compensation pattern manifested by members of low-status groups, they also suggest that self-esteem hardly plays a role in the emergence of compensation for members of high-status groups. In line with other research (During & Jessop, 2015), our affirmation manipulation also failed to have an impact effect on high-status participants' self-esteem. Moreover, high-status group members had no reason to feel threatened by the comparison with an outgroup that was deemed less competent. In spite of this, a clear compensation pattern emerged even for the high-status groups. Thus, although the above efforts shed light with respect to the dynamics of compensation among low-status group members, the puzzle remains as to what exactly drives compensation among members of high-status groups. Rather than a concern for self-esteem and the resulting social creativity strategy observed among low-status group members, we would like to suggest that a key factor contributing to the emergence of compensatory judgments among high-status group members is their concern for the

norms of non-discrimination, triggering what can be seen as a magnanimity strategy. We designed Experiment 2 to test this hypothesis.

Experiment 2

In Experiment 2, we relied on the same paradigm as in Experiment 1 but instead of manipulating the possibility to self-affirm, we manipulated the activation of a norm of non-discrimination versus a norm of honesty. Because the non-discrimination norm has been found to be the default norm in most social situations (see Dambrun & Guimond, 2004; Plant & Devine, 2001), we hypothesized that this is the norm that triggers compensation for high-status groups. In other words, high-status participants in Experiment 1 likely manifested compensation because of the fear to appear discriminant. If this is the case, making the norm explicit in Experiment 2 should not alter in any dramatic way the pattern of judgments found among high-status participants and should thus reveal compensation. In sharp contrast, we hypothesized that the presence of an honesty norm would disrupt the normal course of things and should prevent compensation from emerging in high-status participants' judgments.

We thus predicted that our manipulation would affect only high-status groups such that compensation should emerge in the non-discrimination norm condition whereas ingroup bias on the two dimensions should materialize in the honesty norm condition. We expected no effect of the norm manipulation for low-status groups because, in light of their predicament, their members should not refrain from showing ingroup bias on warmth.

Method

Participants

The sample consisted of 88 psychology students recruited for partial course credit. Four cases were omitted because they either doubted the manipulation or incorrectly completed the questionnaire. The final sample consisted of N=84. Ages ranged between 17 and 41 years (M=20.7; SD=4.76; females=66%). Data collection did not depend on any analysis of results.

Procedure and design

The procedure was the same as in Experiment 1 except for two important modifications. First, we replaced the self-affirmation manipulation with a norm manipulation. Second, the dependent measures were partially different. The norm manipulation took place at the beginning of the experiment with the presentation of the alleged results of a survey on political correctness introduced as a topic of conversation for the upcoming group discussion. The final questionnaire contained (a) manipulation check questions, (b) questions assessing the ingroup and the outgroup on competence and warmth, and (c) questions measuring the perception of pressures toward non-discrimination. All measures used in this experiment are reported below. Finally, participants were debriefed, thanked, and dismissed.

Participants were randomly assigned to a 2 (manipulation of the norm: honesty vs. non-discrimination) x 2 (status: high vs. low) between-subjects design.

Materials

The norm manipulation was adapted from Gabarrot, Falomir-Pichastor, and Mugny (2009) and consisted in the presentation of a fake survey on political correctness versus honesty. Participants learned about the results of a study carried out with a representative sample of psychology students. The results were displayed in graphical form using percentages of responses (i.e., 'Yes', 'No', and 'I don't know') to five questions. Specifically, the questions assessed whether people thought that preventing themselves from expressing negative opinions about any particular group in society was good and fair or bad and hypocritical (e.g., 'Do you think that today's political correctness is excessive and contrary to freedom of speech'; 'Do you think that hiding a negative opinion about a group is hypocritical'; 'Do you think that hiding a negative opinion about a group can have more negative consequences for this group than honestly expressing one's opinion'). We manipulated the alleged answers to the questions so that the majority was "no" in the non-discrimination condition and "yes" in the honesty condition[7] (depending on the question, the majority was conveyed through

percentages varying between 71.16% and 88.36% of the respondents whereas the "minority" option was chosen between 7.54% and 21.30%,, and the "don't know" responses varied between 1.97% and 7.54%).

Measures

Dependent variables. Participants rated the ingroup and the outgroup on the same 12 warmth and competence traits used in Experiment 1 on scales ranging from 1 (= *not at all*) to 9 (= *totally*). We computed a global score for each dimension and each group by averaging the ratings on the six competence traits and on the six warmth traits for the ingroup and for the outgroup. Cronbach's alphas ranged from .79 to .84.

Another series of questions dealt with the perception of pressures toward non-discrimination (e.g., 'Is it acceptable to express a negative attitude toward the other group?' (reverse-coded) on 9-point scales with 1 (9) corresponding to a weak (strong) perception of pressures (α=.92).

Manipulation check measures. We used the same questions as in Experiment 1 to examine the effectiveness of the status manipulation, the legitimacy of the status attribution, and the perceived objectivity of warmth and competence attributions. One question checked whether participants understood the results of the survey on "political correctness" by asking whether the results of the survey showed that most psychology students thought "it was better to express (hide) one's negative attitudes toward others because it was important to be frank with (not to hurt) others'.

Results

Preliminary analyses

We first checked whether participants understood the norm manipulation. All but two participants in the non-discrimination norm (one in a high-status group and the other in a low-status group) correctly understood the message of the survey. We conducted all analyses with and without these two participants and no differences emerged. Below, we present the analyses that include them.

Next, we submitted participants' answers to the status question to a 2 (manipulation of norm: honesty vs. non-discrimination) x 2 (status: high vs. low) between-subjects design. Only the status effect was significant and revealed that participants perceived themselves as having higher status in the high ($M=6.80$, $SD=1.22$) than in the low status ($M=3.76$, $SD=1.10$) conditions, $F(1,80)=149.17$, $p<.001$, $\eta_p^2=.65$, confirming the success of our manipulation.

We submitted participants' perception of the legitimacy of the competence test to the same ANOVA. This analysis revealed a main effect of status, $F(1,80)=10.68$, $p<.002$, $\eta_p^2=.11$, showing that low-status participants perceived the test as less legitimate ($M=5.71$, $SD=1.69$) than high-status participants ($M=6.85$, $SD=1.52$). The effect of norm manipulation was also significant, $F(1,80)=5.78$, $p<.02$, $\eta_p^2=.06$, revealing that participants perceived the test as more legitimate in the non-discrimination ($M=6.74$, $SD=1.54$) than in the honesty norm condition ($M=5.90$, $SD=1.74$).

Finally, we checked whether competence came across as more objective than warmth. In all conditions, the score was significantly superior to 5, the midpoint of the scale (all ts>3.32), sanctioning that competence was perceived as more objective than warmth.

Compensation

We submitted the trait ratings to a 2 (manipulation of norm: honesty vs. non-discrimination) x 2 (status: high vs. low) x 2 (target group: ingroup vs. outgroup) x 2 (dimension: competence vs. warmth) mixed-model ANOVA with the first two factors varying between participants and the last two within them. Confirming our hypothesis, and on top of several significant lower-order effects, the four-way interaction proved significant, $F(1,80)=15.10$, $p<.0002$, $\eta_p^2=.16$. As Table 3 shows, compensation emerged in every condition except for high-status groups confronted with an honesty norm in which case participants exhibited ingroup favoritism on both dimensions.

To further probe the patterns of ingroup and outgroup favoritism as a function of norm, status, and target group, we conducted a new analysis using a score of ingroup favoritism on warmth and another one on competence, paralleling what we did in Experiment 1. Turning to competence first,

the main effect of status proved significant, $F(1,80)=116.50$, $p=.0001$, $\eta_p^2=.59$, with high-status

groups showing ingroup favoritism ($M=1.72$, $SD=1.10$), $t(45)=10.68$, $p<.0001$, and low-status groups

outgroup favoritism ($M=-0.83$, $SD=1.05$), $t(37)=-4.90$, $p<.001$. No other effect was significant.

As for the warmth dimension, the significant main effects of status, $F(1,80)=35.63$, $p<.0001$,

$\eta_p^2=.31$, and norm, $F(1,80)=25.97$, $p<.0001$, $\eta_p^2=.25$, were qualified by the predicted significant

interaction, $F(1,80)=29.02$, $p<.0001$, $\eta_p^2=.27$. Specifically, low-status groups showed similar levels of

ingroup favoritism whether in the honesty ($M=1.31$, $SD=1.11$) or in the non-discrimination ($M=1.38$,

$SD=1.38$) condition, $F(1,36)<1$, ns. More importantly, a significant norm effect emerged for the high-

status groups, $F(1,44)=80.83$, $p<.0001$, $\eta_p^2=.65$. Whereas members of high-status groups showed

outgroup favoritism in the non-discrimination condition ($M=-1.36$, $SD=0.73$), $t(23)=-9.13$, $p<.0001$,

they manifested ingroup favoritism in the honesty condition ($M=1.17$, $SD=1.15$), $t(21)=4.77$,

$p<.0001$. Clearly thus, and in line with our hypothesis, the norm manipulation only influenced the

strategy used by high-status groups and this again took place on the warmth dimension.

As in Experiment 1, we computed a series of correlations between ingroup favoritism on

competence and on warmth, predicting a negative correlation between ingroup favoritism on

competence and ingroup favoritism on warmth in all conditions but for the high-status groups in the

honesty condition. Although all correlations were in the predicted direction, only one out of the three

expected negative correlations came out clearly significant ($r_{high\text{-}status/non\text{-}discrimination}=-.48$, $p=.02$) ; $r_{low\text{-}}$

$_{status/non\text{-}discrimination}=-.38$, $p=.11$; $r_{low\text{-}status/honesty\text{-}discrimination}=-.38$, $p<.11$). In the honesty condition, the

correlation for high-status group was, as expected, clearly non-significant ($r=.12$, $p=.59$).

Pressures toward non-discrimination

We also expected that participants in the high-status condition and non-discrimination norm

condition would perceive more pressures toward non-discrimination than any other groups. A 2

(manipulation of norm: honesty vs. non-discrimination) x 2 (status: high vs. low) ANOVA on the

relevant score revealed a main effect of the norm manipulation, $F(1,80)=57.82$, $p<.0001$, $\eta_p^2=.42$, and

of status, $F(1,80)=64.65$, $p<.0001$, $\eta_p^2=.45$. More importantly, the interaction between norm and status proved significant, $F(1,80)=74.89$, $p<.0001$, $\eta_p^2=.48$. Probing this interaction revealed that, as expected, high-status groups perceived more pressures in the non-discrimination condition ($M=7.30$, $SD=1.20$) than in the honesty condition ($M=3.02$, $SD=0.79$), $F(1,80)=199.50$, $p<.0001$, $\eta_p^2=.82$, no such difference existed for low-status groups ($F<1$). Moreover, the pressures perceived by high-status group in the non-discrimination condition were significantly above 5, the midpoint of the scale, $t(23)=-9.40$, $p<.0001$, whereas the score fell below it in all the other conditions, all ts>5.43 ($M_{\text{low-status/non-discrimination}}=2.91$, $SD=1.30$, and $M_{\text{low-status/honesty}}=3.18$, $SD=1.46$).

Discussion

In Experiment 2, we reasoned that normative concerns are driving the emergence of compensation, at least among high-status group members. As in Experiment 1, we randomly assigned our participants to a low-status or a high-status group. We then manipulated normative beliefs by informing participants that a majority of their fellow students believed that it is important either to avoid harboring prejudice or manifesting discrimination, in order to promote tolerance and respect, or to express one's true beliefs and behave in accordance with one's convictions, so as to uphold honesty and to combat hypocrisy. Participants were then invited to evaluate both the ingroup and the outgroup on competence and warmth traits. We also collected information about the relative perceived objectivity of the two dimensions along with their views regarding the prevailing discriminatory norms.

The present findings fully confirm our hypotheses. As predicted, members of high-status, but not of low-status, group proved sensitive to our norm manipulation. Specifically, low-status group members manifested compensation whether they learned that the majority of their fellow students opposed the expression of discriminatory judgments or supported honesty and condemned hypocrisy. The same pattern emerged when high-status members believed that the prevailing views favored non-discrimination norms. In sharp contrast, when the (alleged) dominant view happened to promote

honesty and condemn hypocrisy when it comes to intergroup judgments, high-status group members derogated low-status group members not only on competence but also on warmth.

These data are in full agreement with our analysis that, at least for high-status group members, normative considerations play a crucial. When non-discrimination norms prevail, high-status group members display compensation. In all likelihood, such a reaction allows these participants to assert their superiority on competence while coming across as magnanimous and compliant with the dominant norms by stressing the outgroup's higher standing on warmth. Interestingly, this pattern replicates the compensation findings found in our earlier work, suggesting that a tacit rule used by high-status participants in previous studies was that norms of non-discrimination prevail (Yzerbyt & Cambon, 2017). When the norms of 'honesty' were salient, participants seemed liberated to claim superiority on both dimensions. The malleability of the warmth dimension apparently lends itself to the manifestation of ingroup favoritism across the board. The correlational evidence, though somewhat weaker than what we found for Experiment 1, proved globally encouraging in that the only correlation that was clearly non-significant emerged in the appropriate condition, namely the high-status groups in the honesty condition.

It is interesting to note how low-status group members appear to be largely insensitive to the nature of the dominant norms. Whether confronted with honesty or non-discrimination norms, they claim superiority on warmth. In our opinion, this is hardly surprising because these participants are in fact the underdogs in the present context. The fact that competence is perceived as more objective than warmth and that low-status group members are thus forced to admit their relative inferiority on this important dimension may give them the feeling that they cannot easily be accused of holding discriminatory views. In other words, because they concede superiority of the outgroup on the more undisputable of the two fundamental dimensions, i.e., competence, they probably feel entitled to some degree of partisanship on the more subjective dimension of warmth without experiencing any

feelings of inadequacy or guilt. In fact, these participants did not convey a sense that there were many pressures toward non-discrimination in the setting.

General discussion

Building upon Yzerbyt et al.'s (2005) initial research, a sizeable number of studies in intergroup research found that people compensate the privileged position of their group on one of the two fundamental dimensions by rating the outgroup more favorably than the ingroup on the other fundamental dimension (Judd et al., 2005; Kervyn et al., 2009; Yzerbyt et al., 2008). In recent years, researcher identified a series of factors that facilitate or prevent the emergence of compensation (for reviews, see Kervyn et al., 2010; Yzerbyt, 2016). For instance, in line with Yzerbyt et al.'s (2005) intuitions, compensation only materialize when the two groups involved in the comparison setting are not engaged in a conflict and requires group members to consider the existing status differential between the groups as legitimate (Cambon et al., 2015, Cambon & Yzerbyt, 2016).

While the empirical evidence in line with compensation has kept accumulating, it is fair to say that comparatively little attention has been devoted to the underlying mechanisms leading to compensation. Although researchers proposed that the search for positive distinctiveness along with social creativity and noblesse oblige strategies are key mechanisms underlying compensation, the available evidence was only correlational. To examine the role of positive distinctiveness more directly, Yzerbyt and Cambon (2017) had participants think about members of a more or a less prestigious and respected outgroup in a context devoid of any conflict, a situation known to generate compensatory judgments. Interestingly, they provided some of their participants with traits pertaining to only one of the two dimensions in order to rate both groups before presenting the traits characteristic of the other dimension. These authors predicted that high-status and low-status groups would manifest compensatory judgments only if they were initially able to compare favorably with the outgroup on their preferred dimension, i.e., competence and warmth, respectively. As predicted, high-status group members invited to rate initially both groups only on warmth expressed ingroup

favoritism whereas strong compensation emerged when both dimensions were presented simultaneously or when competence traits came first. For low-status group members, compensation was much weaker when they initially rated both groups solely on competence. Interestingly, no outright ingroup favoritism emerged in the latter case, contrary to what took place for the high-status group members. According to Yzerbyt & Cambon (2017), this is the case presumably because the subjective nature of the warmth dimension makes it easy to forge partisan views whereas the reality constraints on competence are much stronger.

Interestingly enough, the fact that the differences in status and their associated judgments of competence are less malleable than perceptions of warmth likely places the high-status and low-status groups in somewhat different positions with respect to compensation. Whereas low-status groups have little choice than to secure positive distinctiveness on warmth, high-status groups could well claim superiority on both dimensions. This distinction directly resonates with the social creativity and magnanimity accounts that Yzerbyt and colleagues (2005, 2008) proposed for the emergence of compensation among low-status and high-status groups respectively.

For low-status group members, the comparison with the outgroup offers a limited number of possibilities to secure positive distinctiveness. Given that the competence dimension is constrained, and unless some alternative option presents itself, their self-esteem very much hangs on the possibility to shine on warmth and to derogate the outgroup on that same dimension, i.e., the social creativity strategy. Conversely, any possibility to comfort one's self-views should weaken their need to claim superiority on warmth. In contrast, given that high-status group members can take comfort in their superiority on competence, any other boost to their self-esteem should not otherwise affect their judgment pattern. In fact, high-status group members possibly face another predicament, one that concerns the norms regarding discrimination. The dominant, and often implicit, non-discrimination norms may well encourage high-status groups to concede inferiority on warmth, the more negotiable

dimension. To be sure¸ altering the prevailing norms may suffice for the temptation to derogate the outgroup on all counts to kick in again.

Experiments 1 tested the self-esteem hypothesis by assigning participants to a high- or a low-status group and giving half of the participants in each status a chance to self-affirm before learning about the position of their group. As predicted, low-status group members who did not self-affirm manifested a drop in self-esteem upon learning about their inferior position, compensated as a result, which restored their self-esteem. In sharp contrast, self-affirming allowed low-status group members to maintain their self-esteem intact. Importantly, this took away the need to derogate the high-status group. In line with our analysis, high-status group members remained impervious to our affirmation manipulation. Not only did their self-esteem remain high in all conditions but they also manifested compensation, which suggests that self-esteem per se is not a key factor in their case.

Experiment 2 focused on the concerns driving members of high-status groups and investigated the role of normative beliefs in the emergence of compensation. We reasoned that the expression of compensatory judgments essentially rests on the prevalence of non-discrimination norms. We thus hoped that having participants believe that a majority of people think it is appropriate to express one's true beliefs would have high-status group members expose their conviction that they are better than the outgroup on both dimensions. In contrast, low-status group members should prove insensitive to the nature of the norms, given their objective inferiority on competence. The data supported our conjecture at the nomothetic level, and, to some extent, at the idiothetic level.

As in earlier work, our participants in both experiments confirmed that competence is more objective a dimension than warmth. In line with this state of affairs, the effects observed in the two experiments essentially concern warmth. In Experiment 1, the one judgment that runs counter to the compensation pattern is when self-affirming low-status group members rate the warmth of the outgroup, judging them no colder than themselves. In Experiment 2, the one judgment that is disrupting compensation is when high-status group members in the honesty norm condition rate the

outgroup less warm than themselves. In sum, playing around with the evaluation of warmth offers the means to serve self-esteem for low-status group members and to comply with the norms for high-status group members.

A limitation of our research, which concerns Experiment 1, is that we theorized one of the mechanisms at the heart of compensation at the intergroup level while operationalizing it at the individual level. Indeed, we measured esteem and manipulated affirmation at the individual level whereas a collective version might have been used (Derks et al., 2009; Sherman, Kinias, Major, Kim, & Prenovost, 2007). To a certain extent, the nature of the minimal group paradigm dictated this choice because participants' identification to their group is known to be generally low in this context. Moreover, research shows that self-affirmation mainly affects individual esteem (Derks et al., 2009). Finally, and perhaps ironically, self-esteem has been reported to be a better antecedent and consequence of social creativity (Becker, 2012) but also, more generally, of positive distinctiveness, than collective esteem (Lemyre & Smith, 1986, Oakes & Turner, 1980; for a review, see Brown, 2000). To be sure, we acknowledge the fact that future efforts should test our hypotheses with real groups and should assess individual as well as collective forms of esteem. At the same time, previous work on compensation with real groups (Cambon & Yzerbyt, 2016) leads us to think that the present results are not restricted to the minimal group paradigm and could definitely emerge in situations involving actual groups and higher levels of identification.

Finally and in line with recent methodological claims (Bullock, Green, & Ha, 2010; Fiedler, Schott, &Meiser, 2011; Spencer, Zanna, & Fong, 2005; Stone-Romero & Rosopa, 2008; for the most recent one, see Kline, 2015), the current efforts looked at the processes underlying compensation by going beyond a measurement-of-mediation strategy (Yzerbyt & Cambon, 2017) and turning instead to a testing-process-by-interaction strategy (Jacoby & Sassenberg, 2011, Sigall & Mills, 1998; for a recent review, see Judd, Yzerbyt & Muller, 2014). Specifically, because we manipulated both the independent and the mediating variable, the present results allow for stronger inferences with respect

to the causal chain of events leading to compensation. A most encouraging outcome of our research program is that both approaches, i.e., measurement-of-mediation and testing-process-by-interaction, lead to the same conclusion and point to self-esteem (for low-status groups) and non-discrimination norms (for high-status groups) as core mechanisms of compensation. Building upon these findings, future research should continue to study the contours of compensation in intergroup relations.

References

Abrams, D., & Hogg, M. A. (1988). Comments on the motivational status of self-esteem in social identity and intergroup discrimination. *European Journal of Social Psychology, 18,* 317-334. DOI: 10.1002/ejsp.2420180403

Becker, J., C. (2012). The system-stabilizing role of identity management strategies: Social creativity can undermine collective action for social change. *Journal of Personality and Social Psychology, 103*, 647-662. DOI: 10.1037/a0029240.

Bettencourt, B.A., Dorr, N., Charlton, K., & Hume, D.L. (2001). Status differences and ingroup bias: A meta-analytic examination of the effects of status stability, status legitimacy, and group permeability. *Psychological Bulletin, 127,* 520-542. DOI: 10.1037//0033-2909.127.4.520

Blanz, M., Mummendey, A., Mielke, R., & Klink, A. (1998). Responding to negative social identity: A taxonomy of identity management strategies. *European Journal of Social Psychology, 28*, 697–729. DOI: 10.1002/(SICI)1099-0992(199809/10)28:5<697::AID- EJSP889>3.0.CO;2-#

Branscombe, N.R., & Wann, D.L. (1994). Collective self-esteem consequences of outgroup derogation when a valued social identity is on trial. *European Journal of Social Psychology, 24*, 641-657. DOI: 10.1002/ejsp.2420240603

Bullock, J. G., Green, D. P., & Ha, S. E. (2010). Yes, but what's the mechanism? (Don't expect an easy answer). *Journal of Personality and Social Psychology, 98*, 550–558. DOI:10.1037=a0018933

Brown, R. (2000). Social identity theory: Past achievements, current problems and future challenges. *European Journal of Social Psychology, 30*, 745–778. DOI: 10.1002/1099-0992(200011/12)30:6<745::AID-EJSP24>3.0.CO;2-O

Cadinu, M. R., & Cerchioni, M. (2001). Compensatory biases after ingroup threat: 'Yeah, but we have a good personality'. *European Journal of Social Psychology, 31*, 353–367. DOI:10.1002/ejsp.46

Cambon, L., & Yzerbyt, V.Y. (2016). Compensation is for real: Evidence from existing groups in the context of actual relations. *Group Processes and Intergroup Relations.* DOI: r

Cambon, L., Yzerbyt, V.Y., and Yakimova, S. (2015). Compensation in intergroup relations: An investigation of its structural and strategic foundations. *British Journal of Social Psychology, 54*, 140-158. DOI: 10.1111/bjso.12067

Cohen, G.L.; Sherman, D.K (2014). The psychology of change: self-affirmation and social psychological intervention. *Annual Review of Psychology, 65*, 333-371. DOI: 10.1146/annurev-psych-010213-115137

Crocker, J., & Major, B. (1989). Social stigma and self-esteem: The self-protective properties of stigma. *Psychological Review, 96,* 608–630. DOI:10.1037/0033-295X.96.4.608

Cuddy, A. J. C., Fiske, S.T., & Glick, P. (2008). Warmth and competence as universal dimensions of social perception: The stereotype content model and the BIAS Map. *Advances in Experimental Social Psychology, 40*, 61-149. DOI: 10.1016/S0065-2601(07)00002-0

Dambrun, M., & Guimond, S. (2004). Implicit and explicit measures of prejudice and stereotyping: Do they assess the same underlying knowledge structure. *European Journal of Social Psychology, 34*, 663-676. DOI: 10.1002/ejsp.223

Derks, B., Van Laar, C., & Ellemers, N. (2007). Social creativity strikes back: Improving low status group members' motivation and performance by valuing ingroup dimensions. *European Journal of Social Psychology, 37*, 470–493. DOI:10.1002/ejsp.375

Derks, B., Van Laar, C., & Ellemers, N. (2009). Working for the self or working for the group: How personal and social self-affirmation promote collective behavior among members of devalued groups. *Journal of Personality and Social Psychology, 96*, 183–202. DOI:10.1037/ a0013068

During, C., Jessop, D.C. (2015). The moderating impact of self-esteem on self-affirmation effects. *British Journal of Health Psychology, 20*, 274-289. DOI: 10.1111/bjhp.12097

Fiske, S.T. (2015). Intergroup biases: A focus on stereotype content. *Current Opinion in Behavioral Sciences, 3,* 45-50. DOI: 10.1016/j.cobeha.2015.01.010

Fein, S., & Spencer, S.J. (1997). Prejudice as self-image maintenance : Affirming the self through derogating others. *Journal of Personality and Social Psychology, 73*, 31-44. DOI: 10.1037/0022-3514.73.1.31

Fiedler, K., Schott, & Meiser, T. (2011). What mediation analysis can (not) do. *Journal of Experimental Social Psychology, 47*, 1231-1236. DOI: 10.1016/j.jesp.2011.05.007

Gabarrot, F., Falomir-Pichastor, J.M., & Mugny, G. (2009). Being similar versus being equal: Intergroup similarity moderates the influence of in-group norms on discrimination and prejudice. *British Journal of Social Psychology, 48*, 253-273. DOI: 10.1348/014466608X342943

Harter, J. (1986). Processes underlying the construction, maintenance, and enhancement of the self-concept in children. In J. Suls & A.G. Greenwald (Eds), *Psychological Perspectives on the self* (Vol. 3, pp 136-182). Hillsdale, NJ: Erlbaum.

Heatherton, T.F., & Polivy, J. (1991). Development and validation of a scale for measuring state self-esteem. *Journal of Personality and Social Psychology, 60*, 895-910. DOI: 10.1037/0022-3514.60.6.895

Jacoby, J., & Sassenberg, K. (2011), Interactions do not only tell us *when*, but can also tell us *how*: Testing process hypotheses by interaction. *European Journal of Social Psychology, 41*, 180–190. doi:10.1002/ejsp.762

Jaremka, L.M., Bunyan, D.P., Collins, N.L., Sherman, D.K. (2011). Reducing defensive distancing: Self-affirmation and risk regulation in response to relationship threats. *Journal of Experimental Social Psychology, 47*, 264-268. DOI: 10.1016/j.jesp.2010.08.015

Jeffries, C.H., Hornsey, M.J., Sutton, R.M., Douglas, K.M., & Bain, P.G. (2012). The David and Goliath principle. Cultural, ideological, and attitudinal underpinnings of the normative

protection of low-status groups from criticism. *Personality and Social Psychology Bulletin, 38*, 1053-1065. DOI: 10.1177/0146167212444454

Judd, C.M., James-Hawkins, L. Yzerbyt, V.Y., & Kashima, Y. (2005). Fundamental dimensions of social judgment: Understanding the relations between judgments of competence and warmth. *Journal of Personality and Social Psychology, 89*, 899-913. DOI: 10.1037/0022-3514.89.6.899

Judd, C. M., Yzerbyt, V. Y., & Muller, D. (2014). Mediation and moderation. In H. T. Reis & C. M. Judd (Eds.), *Handbook of research methods in social and personality psychology* (2nd edition, pp. 653-676). Cambridge, UK: Cambridge University Press.

Kenny, D.A., Mannetti, L., Pierro, A., Livi, S., & Kashy, D.A. (2002). The statistical analysis of data from small groups. *Journal of Personality and Social Psychology, 83,* 126-137. DOI: 10.1037/0022-3514.83.1.126.

Kervyn, N., Yzerbyt, V., & Judd, C. (2010). Compensation between warmth and competence: Antecedents and consequences of a negative relation between the two fundamental dimensions of social perception. *European Review of Social Psychology, 21,* 155-187. DOI: 10.1080/13546805.2010.517997

Kervyn, N., Yzerbyt, V., Demoulin, S., & Judd, C. (2008). Competence and warmth in context: The compensatory nature of stereotypic views of national groups. *European Journal of Social Psychology, 38,* 1175–1183. DOI:10.1002/ejsp.526

Kline, R.B. (2015). The mediation myth. *Basic and Applied Social Psychology, 37*, 202-213. DOI: 10.1080/01973533.2015.1049349

Leach, C.W., Snider, N., & Iyer, A. (2002). "Poisoning the conscience of the fortunate": The experience of relative advantage and support for social equality. In I. Walker & H. J. Smith (Eds.), *Relative deprivation: Specification, development, and integration* (pp. 136-163). New York: Cambrdige University Press. DOI: 10.1017/CBO9780511527753

Lemaine, G. (1974). Social differentiation and social originality. *European Journal of Social Psychology, 4*, 17–52. DOI:10.1002/ejsp.2420040103

Lemyre, L., & Smith, P.M. (1985). Intergroup discrimination and self-esteem in the minimal group paradigm. *Journal of Personality and Social Psychology, 49*, 660-670. DOI: 10.1037/0022-3514.49.3.660

Mackinnon, S.P., Smith, S.M., & Carter-Rogers, K. (2015). Multidimensional self-esteem and test derogation after negative feedback. *Canadian Journal of Behavioural Science, 47*, 123-126. DOI: 10.1037/a0038364

Major, B., Testa, M., & Bylsma, W. H. (1991). Responses to upward and downward social comparisons: The impact of esteem-relevance and perceived control. In J. Suls & T. A. Wills (Eds.), *Social comparison: Contemporary theory and research* (pp. 237–260). Hillsdale, NJ: Erlbaum.

Mummendey, A., & Schreiber, H. (1983). Better or just different? Positive social identity by discrimination against, or by differentiation from outgroups. *European Journal of Social Psychology, 13*, 389-397. DOI: 10.1002/ejsp.2420130406

Napper, L., Harris, P.R., & Epton, T. (2009). Developing and testing a self-affirmation manipulation. *Self and Identity, 8,* 45-62. DOI: 10.1080/15298860802079786

Niens, U., & Cairns, E. (2003). Explaining social change and identity management strategies: New directions for future research. *Theory and Psychology, 13*, 489-509. DOI: 10.1177/09593543030134003

Oakes, P.J., & Turner, J.C. (1980). Social categorization and intergroup behavior: Does minimal intergroup discrimination make social identity more positive. *European Journal of Social Psychology, 10*, 295-301.

Owuamalam, C.K., Wong, K.X., & Rubin, M. (2016). Chubby but cheerful? Investigating the

compensatory judgments of high, medium, and low status weight groups in Malaysia. *Cogent*

Psychology, 3, 1188441. DOI: 10.1080/23311908.2016.1188441.

Plant, E. A., & Devine, P. G. (2001). Responses to other - imposed pro - Black pressure: Acceptance

or backlash? *Journal of Experimental Social Psychology, 37*, 486-501. DOI:

10.1006/jesp.2001.1478

Rosenberg, M. (1979). *Conceiving the self*. New York: Basic Books.

Reinhard, M.-A., Stahlberg, D., & Messner, M. (2008). Failure as an asset for high-status persons –

relative group performance and attributed occupational success. *Journal of Experimental Social*

Psychology, 44, 501-518. DOI: 10.1016/j.jesp.2007.07.006

Sachdev, I., & Bourhis, R. Y. (1985). Social categorization and power differentials in group relations.

European Journal of Social Psychology, 15, 415-434. DOI: 10.1002/ejsp.2420150405

Spencer, S. J., Zanna, M. P., & Fong, G. T. (2005). Establishing a causal chain: Why experiments are

often more effective than mediational analyses in examining psychological processes. *Journal*

of Personality and Social Psychology, 89, 845–851. DOI:10.1037/0022-3514.89.6.845

Schmader, T., & Major, B. (1999). The impact of ingroup vs. outgroup performance on personal

values. *Journal of Experimental Social Psychology, 35*, 47-67. DOI: 10.1006/jesp.1998.1372

Sherman, D.K., Kinias, Z., Major, B., Kim, H.S, Prenovost, M. (2007). The group as a resource:

Reducing biased attributions for group success and failure via group affirmation. *Personality*

and Social Psychology Bulletin, 33, 1100-1112. DOI: 10.1177/0146167207303027

Sigall, H., & Mills, J. (1998). Measures of independent variables and mediators are useful in social

psychology experiments: But are they necessary? *Personality and Social Psychology Review*, *2*,

218-226. doi.org/10.1207/s15327957pspr0203_5

Spencer-Rodgers, J., Major, B., Forster, D.E., & Peng, K. (2016). The power of affirming group values: Group affirmation buffers the self-esteem of women exposed to blatant sexism. *Self and Identity, 15*, 413-431. DOI: 10.1080/15298868.2016.1145593

Stone-Romero, E. F., & Rosopa, P. J. (2011). Experimental tests of mediation models: Prospects, problems, and some solutions. *Organizational Research Methods, 14,* 631–646. DOI:10.1177=1094428110372673

Tafarodi, R.W., & Swann, W.B. (1995). Self-liking and self-competence as dimensions of global self-esteem: Initial validation of a measure. *Journal of Personality Assessment, 65*, 322-342.

Tajfel, H. (1974). Social identity and intergroup behavior. *Social Science Information/sur les sciences sociales, 13*, 65–93. DOI:10.1177/053901847401300204

Tajfel, H., & Turner, J.C. (1979). An integrative theory of intergroup conflict. In W. G. Austin & S. Worchel (Eds.), *The social psychology of intergroup relations* (pp. 33- 47). Monterey, CA: Brooks/Cole.

Tausch, N., Kenworthy, J.B., & Hewstone, M. (2007). The confirmability and disconfirmability of trait concepts revisited: Does content matter? *Journal of Personality and Social Psychology, 92,* 542-556. DOI: 10.1037/0022-3514.92.3.542

Tesser, A., Millar, M., & Moore, J. (1988). Some affective consequences of social comparison and reflection processes: The pain and pleasure of being close. *Journal of Personality and Social Psychology Journal, 54*, 49-61. DOI: 10.1037/0022-3514.54.1.49

Vanbeselaere, N., Boen, F., van Avermaet, E., & Buelens, H. (2006). The Janus face of power in intergroup contexts: A further exploration of the noblesse oblige effect. *The Journal of Social Psychology, 146*, 685-699. DOI: 10.3200/SOCP.146.6.685-699

van Knippenberg, A. (1978). Status differences, comparative relevance and intergroup differentiation. In H. Tajfel (Ed.), *Differentiation between social groups: studies in the social psychology of intergroup relations* (pp. 171-199). London, UK: Academic.

Yzerbyt, V. (2016). Intergroup stereotyping. *Current Opinion in Psychology, 11*, 90-95. DOI: 10.1016/j.copsyc.2016.06.009

Yzerbyt, V., Cambon, L. (2017). The dynamics of compensation: When ingroup favoritism paves the way for outgroup praise. *Personality and Social Psychology Bulletin, 43*, 587-600. DOI: 10.1177/0146167216689066

Yzerbyt, V., Kervyn, N., & Judd, C. (2008). Compensation versus halo: The unique relations between the fundamental dimensions of social judgment. *Personality and Social Psychology Bulletin, 34*, 1110-1123. DOI: 10.1177/0146167208318602

Yzerbyt, V., Provost, V., & Corneille, O. (2005). Not so competent but warm… Really ? Compensatory stereotypes in the French-speaking world. *Group Processes and Intergroup Relations, 8,* 291-308. DOI: 10.1177/1368430205053944

Table 1. Ratings as a function of affirmation, dimension, ingroup status, and target group

(Experiment 1)

Dimension	Affirmation		No-affirmation	
	Competence	Warmth	Competence	Warmth
Low-status groups				
Ingroup	5.80a (1.06)	7.38b (0.45)	5.17a (1.04)	7.39b (0.71)
Outgroup	6.98b (0.71)	7.09b (0.71)	6.96b (0.64)	4.51a (1.19)
High-status groups				
Ingroup	6.95b (0.81)	5.69a (1.13)	7.13b (1.08)	5.39a (0.86)
Outgroup	5.08a (0.74)	7.03b (0.85)	5.08a (1.09)	6.88b (0.21)

Note: Means with different subscripts indicate a significant difference between ingroup and outgroup.

All tests were conducted using $p<.005$. Standard deviations are in parentheses.

Table 2. Self-esteem at the three measurement occasions and contrast values as a function of affirmation and ingroup status (Experiment 1)

	No-affirmation		Affirmation	
	Low-status	High-status	Low-status	High-status
Time1 Self-esteem (T1)	3.49 (0.54)	3.72 (0.35)	3.43 (0.52)	3.66 (0.36)
Time2 Self-esteem (T2)	2.22 (0.71)	3.91 (0.31)	3.54 (0.61)	3.85 (0.48)
Time3 Self-esteem (T3)	3.32 (0.72)	3.96 (0.35)	3.75 (0.50)	3.99 (0.69)
Linear Contrast [T3 - T1]	-0.17 (0.50)	0.24* (0.21)	0.31* (0.58)	0.32* (0.62)
Quadratic Contrast [(T1 + T3)/2 - T2]	1.19* (0.74)	-0.07 (0.15)	0.05 (0.39)	-0.03 (0.35)

Note: Standard deviations are in parentheses. Contrast values with a * are significantly different from 0 at $p<.05$;

Table 3. Ratings as a function of norm activation, dimension, ingroup status, and target group

(Experiment 2)

Dimension	Honesty		Non-discrimination	
	Competence	Warmth	Competence	Warmth
Low-status groups				
Ingroup	5.92a (0.69)	7.12b (0.77	5.78a (1.09)	7.09b (1.02)
Outgroup	6.57b (0.77)	5.82a (0.93)	6.79b (1.10)	5.71a (1.22)
High-status groups				
Ingroup	7.30b (1.00)	6.52b (1.10)	7.15b (1.03)	5.27a (0.70)
Outgroup	5.57a (1.04)	5.36a (0.64)	5.43a (1.13)	6.63b (0.85)

Note: Means with different subscripts indicate a significant difference between ingroup and outgroup.

All tests were conducted using $p<.005$. Standard deviations are in parentheses.

Notes

[1] According to Social Identity Theory (Tajfel & Turner, 1979); social creativity refers to a set of identity management strategies that people use when they want to remain group members and when group boundaries are impermeable while the status relations are perceived to be legitimate. These strategies aim at coping with ingroup's disadvantage and at establishing positive distinctiveness. At least three social creativity strategies have been identified. In the first, positive distinctiveness can be established by engaging in a downward comparison. A second strategy is to re-evaluate the value or attribute of the comparison dimension. The last one, called compensating bias by Tajfel (1974), consists in changing the comparison dimension, that is the group compensates for negative distinctiveness on a dimension on which the outgroup exceeds the ingroup by comparing the groups on a dimension on which the ingroup is superior.

[2] One should note that we manipulated affirmation and measure self-esteem at the individual level and not at the collective level. The rationale for this choice is that in the minimal group paradigm, participants' identification to their group is generally low. As Derks et al. (2009) showed, a self-affirmation strategy is more effective for low identifiers than for high identifiers and affects their individual esteem more than their collective esteem.

[3] Cambon et al., (2015) showed that compensation appear in asymmetrical relations that are perceived to be legitimate and are devoid of competition or conflict. Several experiments showed that the minimal group paradigm fulfill these conditions provided that the status manipulation be perceived as legitimate.

[4] The correlations between the two components of trait and state self-esteem measured during the recruitment phase were positive (r=.83 and r=.88 for the self-competence/performance and self-liking/social components, respectively). We therefore decided to look only to the state self-esteem scale as comparison point and dropped from the analysis the results pertaining to trait self-esteem.

[5] Even if participants were randomly assigned to each group and our recruitment procedure ensure that none of the participants knew each other, it is possible that a common fate or mutual influence had affected group members, creating a lack of independence of observations in groups (Kenny, Mannetti, Pierro, Livi, & Kashy, 2002). In order to test non-independence, we estimated the intraclass coefficient (ICC) using ANOVA (Kenny et al., 2002). Group was treated as the independent variable in a one-way ANOVA. We computed four ANOVAs, one for each experimental condition (affirmation versus non-affirmation) and for each dependent variable (ingroup favoritism on warmth versus on competence). The ICC's were all negative and close to 0 (-.05 ≤ ICCs ≥ -.10) and none reached significance, meaning that we can reject the hypothesis that the scores within groups are equal while the group means differ. We repeated the same analysis in Experiment 2. The ICCs were 0.05, -0.14, -0.16, and -0.16, the last three being significant. For, these last three, the results mean that the scores within groups vary and the group means are equal. In sum, there was either no proof of dependence and even proof of independence.

[6] No differences emerged concerning the two dimensions of self-esteem (performance and social), thus we dropped them from the analysis.

[7] We pretested this manipulation asking two groups of participants to what extent they share the opinion expressed by the majority of students in the survey on a 9-point scales ranging from 1 (=I totally disagree with the opinion expressed by the majority) to 9 (=I totally agree with the opinion expressed by the majority). One group (N=20) was exposed to the non-discrimination version whereas the second group (N=20) was exposed to the honesty version. In order to compare the two groups, participants' scores in the discrimination condition were reverse-coded. As we expected, participants' perception differed as a function of the experimental condition ($M_{non\text{-}discrimination}$=7.00; $M_{honesty}$=2.55), $t(38)$= 10.60, $p<.001$.