RESEARCH ARTICLE

Warmth and Competence in Interpersonal Comparisons: The Quiz Master Paradigm through the Lens of Compensation

Julie Terache, Stéphanie Demoulin and Vincent Yzerbyt

Research shows that the two fundamental dimensions of warmth and competence often relate negatively in intergroup stereotypes. This 'compensation effect' emerges in group and person perception but has never been examined in situations of interpersonal comparisons involving the self as an individual. In three experiments, we adapted the Quiz Master paradigm to create a difference of competence between two contestants. Questioners and answerers (Expt. 1 to 3) as well as observers (Expt. 2 and 3) rated both contestants on warmth and competence. Results replicated the Quiz Master pattern in that questioners appeared more competent both for observers and answerers but not for questioners (Expt. 1 to 3). As predicted, observers manifested compensation by judging answerers warmer than questioners (Expt. 2 and 3). Whereas no compensation emerged for answerers, questioners generally perceived the answerer as warmer than themselves (Expt. 2 and 3). Overall, our results suggest that although observers rely on compensation, this was not the case for members of the interaction who showed reluctance to express any self-favoring judgments. These findings constitute a first contribution to the understanding of compensation outside of intergroup relations, and more broadly of interpersonal dynamics at work in social comparison situations.

Everyday life offers numerous occasions in which the comparison with other people proves unpleasant. The successes and achievements of others create a sense of self-failure, forcing us to look for alternative domains of accomplishment (Steele, 1988; Tesser, 2000, 2001). As it turns out, abundant research shows that perceivers appraise their social world in terms of two fundamental dimensions, namely warmth and competence (Fiske et al., 2002; for reviews, Abele at al., 2019; Abele & Wojciszke, 2014; Fiske, 2015; Yzerbyt, 2016). In all likelihood, people should thus rely on these two dimensions when they compare to others. Although research investigating naturalistic and spontaneous comparisons supports this conjecture (Locke, 2014; Locke & Nekich, 2000), to our knowledge, however, the literature on interpersonal comparison has not formalized situations of comparison in terms of these two dimensions. Rather, the available work focused on the comparison of performance (i.e., competence) and its consequences for performance and self-perception (for a meta-analysis, Gerber, Wheeler & Suls, 2017), leaving aside the warmth dimension. What happens on this front when people are confronted with a more competent other? Does this comparison in terms of competence have an impact on people's self-regard in terms of warmth? Drawing on existing work that looked at the relation between the two dimensions in intergroup comparisons (Yzerbyt, 2018), the present studies offer a first attempt at investigating these situations at the interpersonal level.

The Big Two of Social Perception
A vast body of research suggests that social perception revolves around two fundamental dimensions (for reviews, Abele & Wojciszke, 2014; Fiske, 2015; Yzerbyt, 2016). Going beyond the idea of a single evaluative dimension orchestrating social perception, Rosenberg et al. (1968) proposed that two (nearly) orthogonal dimensions provided a much better fit for people's impressions about others. These early efforts paved the way for the Stereotype Content Model (SCM; Fiske et al., 1999; Fiske et al., 2002) which focuses on the role of structural factors of interdependence and power resulting in perceptions of two dimensions of warmth and competence, respectively (Kervyn, Fiske, & Yzerbyt, 2015). Impression formation researchers similarly stressed the importance of a bidimensional space organized around communion and agency (Abele & Wojciszke, 2014; Wojciszke, 1994).

Whereas warmth relates to the perceived intent of a social target and materializes in such traits as sociable, sensitive, sincere or reliable, competence relates to its perceived ability and translates into traits like competent,

Université catholique de Louvain, BE
Corresponding author: Julie Terache (julie.terache@uclouvain.be)

efficient, ambitious, and self-assured (e.g., Fiske, Cuddy & Glick, 2007; Yzerbyt, Kervyn & Judd, 2008). This two-fold distinction spreads far beyond the spectrum of social psychology, emerging in such fields as political science (Abelson et al., 1982; Wojciszke & Klusek, 1996), cognitive and neuropsychology (Harris & Fiske, 2006; Willis & Todorov, 2006), or even human-animal psychology (Sevillano & Fiske, 2016). Recent work also suggests that these two dimensions further split into two facets each (Abele et al., 2016). Specifically, warmth comprises facets of *friendliness* and *morality* (Abele et al., 2016; Leach, Ellemers, & Barreto, 2007) whereas competence covers both *ability* and *assertiveness* (Abele et al., 2016; Carrier et al., 2014).

Relations between the Two Dimensions

Work on the SCM (Fiske et al., 2002) reveals that, although some social groups come across as high or low on both dimensions of warmth and competence, people characterize many groups by means of mixed stereotypes (Fiske, Cuddy, & Glick, 2007; Fiske et al., 1999). This idea of mixed stereotypes initially emerged in research on ambivalent sexism (Glick & Fiske, 1996), with benevolent sexism portraying (cooperative) women as low in competence but high in warmth and hostile sexism depicting (competitive) women as high in competence but low in warmth. Such a negative relationship resonates well with work on the so-called 'compensation effect' as proposed in the Dimensional Compensation Model (DCM; Judd et al., 2005; Yzerbyt, Provost, & Corneille, 2005; for a review, Yzerbyt, 2018).

Squarely located within social identity approaches (Tajfel & Turner, 1979), the DCM proposes that many situations lead group and group members to manifest ambivalent stereotypes. Yzerbyt et al. (2005) initially examined the stereotypes characterizing the French and French-speaking Belgians. Both the members of these groups as well as third-party observers (Swiss) considered French as higher on competence but lower on warmth than Belgians. Compensation emerges for fictional groups created in laboratory contexts (e.g., Judd et al., 2005; Kervyn, Judd, & Yzerbyt, 2009a; Yzerbyt et al., 2008) as well as for existing national groups (Kervyn et al., 2008; Yzerbyt et al., 2005). It shows up both in the judgments by members of the groups (e.g., Cambon & Yzerbyt, 2018; Kervyn et al., 2008; Yzerbyt & Cambon, 2017; Yzerbyt et al., 2005) and by observers (Judd et al., 2005; Yzerbyt et al., 2005; Yzerbyt et al., 2008). Compensation occurs more readily when there is a legitimate and stable status difference between the groups (Cambon & Yzerbyt, 2016; Cambon, Yzerbyt, & Yakimova, 2015), suggesting that this pattern relates to the justification of social hierarchy. Finally, empirical work relying on both explicit and implicit measures (e.g., Carlsson & Björklund, 2010; Kervyn, Yzerbyt, & Judd, 2011) further stresses the robustness of the compensatory pattern of judgments (Yzerbyt, 2018; for a review, see Kervyn, Yzerbyt, & Judd, 2010).

Of note, compensation is specific to warmth and competence (Yzerbyt et al., 2008), as those two are considered to be the fundamental dimensions shaping our social world, and as most traits or domains end up relating to one dimension or the other (Fiske et al., 2007; Fiske et al., 2002). Additionally, this pattern is not the result of an opposition between two distinguished constructs. In fact, when measuring evaluations for a third dimension (i.e., healthiness), researchers found that compensation did not occur (Yzerbyt et al., 2008).

Compensation in Interpersonal Relations

Although the early work on compensation concerns intergroup relations, a few studies investigated how this effect may apply to interpersonal settings. The limited data available suggest that a comparison between two individuals elicits compensation in that comparing targets with one superior to the other on one dimension leads to a compensatory pattern on the other dimension (e.g., Judd et al., 2005, Expt. 3; Kervyn et al., 2016; Kervyn et al., 2009b, Expt. 3). Using a minimal group paradigm, Kervyn and colleagues (Kervyn et al., 2009b, Expt. 3) had external observers report their perception of two participants, one allegedly a member of a high competence group and the other of a low competence group. As expected, observers rated the high competence (warmth) participant less warm (competent) than the low competence (warmth) one.

A key question concerns what happens for those individuals directly involved in the interaction. Collange, Fiske, and Sanitioso (2014) showed that when participants experienced threat about their competence, they displayed compensation and perceived the comparison partner as colder. Although these data illustrate how a specific self-judgment may influence the perception of a partner in a compensatory way, they remain silent regarding self-perception. To our knowledge, the above-mentioned study by Kervyn and colleagues (Kervyn et al., 2009b, Expt. 3) is the only published set of data that looked at compensation for both self- and other-perceptions in interpersonal interactions. Indeed, these authors not only collected impressions from observers but also from the members of the comparison proper. Results for the latter showed that compensation emerged when participants judged the member of the other group but only when competence was the manipulated dimension. No significant effect emerged for self-judgments, suggesting that it may be more difficult to alter perceptions of the self than of another person.

Despite its merits, Kervyn et al.'s (2009b) study fails to provide information about the direct comparison between judgments of the self and of another person. Two additional issues make it difficult to generalize these findings. First, the interaction that took place in the study unfolded in an intergroup context, as the two individuals thought they were members of groups that differed in terms of competence or warmth. In other words, the judgments concerned members of two groups and not necessarily specific individuals. Second, these judgments hardly derived from naturalistic behaviors but from the information provided about the groups via a prearranged questionnaire.

Although not directly about two-person interactions, recent work on compensation in impression management also speaks to the present concerns. Several studies show

that people strategically use warmth and competence in a compensatory way to manage self-presentation (Holoien & Fiske, 2013; Lindholm & Yzerbyt, 2018). For instance, Swencionis and Fiske (2016) found that high-status (lower-status) participants who imagined interacting with lower-status (higher-status) participants emphasized their own warmth (competence). The authors conjectured that this pattern aims at disconfirming stereotypes about their status. The spontaneous and strategic use of compensation for self-presentation shows that people are aware of the role the two dimensions hold in social perception, and intuitively know how to manage the tradeoffs between them to handle social situations.

In sum, no research to date has directly examined the emergence of compensation in comparisons involving the self in interpersonal (and not intergroup) interactions. To examine this issue, we decided to bring in participants to the laboratory and have them interact in such a way that one of the two partners was believed to be more competent than the other was. Because previous findings showed that the strength of compensation is stronger when the manipulation concerned competence rather than warmth, (Judd et al., 2005; Kervyn et al., 2009b), we chose to manipulate comparisons on the domain of competence only.

Overview of the Hypotheses and the Experiments

Because we were interested in the instigation of lower competence, we decided to turn to the classical Quiz Master paradigm (Ross, Amabile, & Steinmetz, 1977). Briefly, this paradigm has participants endorse the role of questioner, answerer or observer in the context of a quiz game. This situation creates a differential of competence whereby observers and answerers (but not questioners) judge questioners as more competent than answerers. Building on previous work on intergroup and interpersonal perception (for a review, Yzerbyt, 2018), we hypothesized that compensation would emerge in the eyes of observers (Judd et al., 2005; Kervyn et al., 2009b) such that they would compensate by perceiving the answerer as being warmer than the questioner.

Regarding the two members of the interaction proper, we can formulate our predictions for the evaluations of the answerers and the questioners at two different levels. One can compare how participants rated answerers and questioners, i.e., the interpersonal level, or one can compare how participants rated themselves versus the other, i.e., the intrapersonal.

First considering the interpersonal level, i.e., the comparison between the evaluations of both targets (the answerer and the questioner), the Quiz paradigm led us to formulate different hypotheses for the participants in the role of answerers (i.e., participants disadvantaged on competence) and those in the role of questioners (i.e., participants advantaged on competence). Turning first to answerers, we expected to replicate earlier findings in that that they would perceive the questioner as being more competent than themselves. Provided this is the case, the literature suggests two possible patterns as to their response on warmth.

On the one hand, one could argue that what applies for intergroup comparisons should also apply for interpersonal comparisons. Because warmth and competence are at the heart of self-perception and self-esteem (Abele & Wojciszke, 2007; Tafarodi & Swann, 2001; Wojciszke et al., 2011), individuals challenged on one dimension are likely to react by investing the other dimension. Hence, a possible reaction for answerers would be to admit their disadvantage on the threatened dimension and to compensate by expressing an advantage on the other dimension, i.e., warmth. Although unrelated to interpersonal relations, Helm et al.'s results (2017) support the idea that people threatened on one dimension respond by increasing self-perception on the other dimension, suggesting that self-perception is indeed governed by compensatory dynamics.

On the other hand, several studies support the idea that individuals find it difficult to rate themselves above others (Dunning et al., 1989; Klein, 2001; Muller & Butera, 2004), especially when judging a specific other rather than an abstract target (Alicke et al., 1995; Perloff & Fetzer, 1986). Alicke and colleagues (1995, Expt. 6) indeed found that participants would rather declare the comparison target as superior to themselves, even without feedback vouching for that target, than to express self-favoring evaluations. This reluctance to brag only disappears when people can rely on objective information and are therefore not at risk to appear self-indulging (Dunning et al., 1989; Klein, 2001). Additionally, work conducted on the 'hubris hypothesis' showed that people who claim self-superiority are disliked by their peers (Hoorens et al., 2012; Van Damme, Hoorens, & Sedikides, 2016), and this whether they express superiority in warmth-related roles (i.e., being a better friend than others) or competence-related ones (i.e., being a better student than others) (Hoorens et al., 2012). Overall, these findings indicate that not only individuals have trouble expressing self-superiority, but they actually run the risk of being negatively evaluated when they do so. Such data suggest that, even though they may be facing a negative comparison on competence (as ascertained by observers' ratings), answerers might be reluctant to express self-favoring differences on warmth, hence would not show compensation.

In short, building on previous work, we formulated two competing hypotheses as to answerers' response to their disadvantage on competence, such that they could either show compensation or refrain from expressing a self-advantage on warmth and therefore show no compensation.

Turning to questioners, the specific paradigm we used led us to expect no differences on the way they attributed competence to themselves and the other. According to the fundamental attribution error research (Ross et al., 1977), this is a consequence of questioner's awareness of their situational advantage. Logically, this means that we should observe no difference on warmth.

In line with Kervyn et al.'s (2009b) analytic strategy, it is also possible to examine ratings at the intrapersonal level, meaning that we examined ratings of self and other separately. Because people need to hold on to a stable (and positive) self-image (Steele, 1988), members of the

interaction should express differences on competence or warmth as a function of their role (i.e., questioner or answerer) more easily when judging others than when judging the self. Kervyn et al.'s findings (2009b, Expt. 3) support this reasoning in that compensation emerged in evaluations of the other participant but not of the self. In other words, we predicted that answerers would rate the questioner as more competent than the questioners would rate the answerer. Conversely, questioners would rate the answerer more positively on warmth than the answerers would rate the questioner. We expected no difference between answerers and questioners in their judgements of self-competence and of self-warmth.

In three experiments, participants came to the lab to take part in a short quiz game before completing scales assessing warmth and competence impressions. Because our primary interest concerned interpersonal compensation in comparisons involving the self, Experiment 1 comprised the two members of the interaction, i.e., answerer and questioner. In Experiment 2, we enriched this design by including an external observer of the interaction. Finally, Experiment 3 aimed to explore the warmth and competence impressions in more depth by assessing all four facets the dimensions (Abele et al., 2016). Study materials, data, and code are available via the website of the Open Science Foundation (https://osf.io/73xyv/?view_only=9072bc57963b449bbced9ec61304cf21).

Experiment 1

To address the issue of interpersonal compensation in comparisons involving the self, Experiment 1 had participants confronted with a two-person interaction modeled after the classical quiz paradigm used to study the fundamental attribution error (Ross et al., 1977). In their initial experiment, Ross and colleagues (1977, Expt. 1) assigned participants to the roles of answerer or questioner. In the experimental condition, questioners composed the questions whereas in the control condition they posed questions created by a third party beforehand. After the quiz interaction, participants rated both contestants on items assessing general knowledge. Results showed that, compared to the answerers in the control condition, those in the experimental condition perceived the questioner as more cultivated, whereas the questioners did not report any difference. These results replicated in a second experiment in which external observers evaluated accomplices re-enacting the two conditions from the first experiment.

Although this paradigm originally aimed at demonstrating the impact of the fundamental error of attribution (Ross et al., 1977), it offers an ideal setting to examine the consequences of interpersonal comparison on competence. Because we had access to a limited sample of participants, we only mimicked the experimental condition and asked questioners to select questions among a list. We randomly assigned pairs of participants to the role of answerer or questioner. After taking part in the quiz game, both participants reported their perception of self and other on competence and warmth.

Method
Participants

Ross et al.'s (1977) relied on 24 pairs of participants in the experimental condition. To accommodate for potential loss of information, we planned for 60 second-year psychology students (M_{Age} = 20.32, SD_{Age} = 2.27, 7 men) from a large European university to come to the laboratory in exchange for course credit.[1] Participants registered to the experiment in pairs along with another student whom they did not know. They learned that the study concerned interpersonal communication, and that they would have to participate in a quiz game before answering a short questionnaire.

Procedure and materials

Upon participants' arrival at the laboratory, an ostensibly random draw assigned them to the role of questioner or answerer. The experimenter then presented the task as a short quiz game. Questioners had to select 10 questions (out of 30) depending on their preferences and to avoid making comments or giving clues during the interaction.[2] Answerers had to answer questions with no time limit but learned that they could move on to the next question if they did not know the answer. After the quiz, the experimenter directed participants to individual rooms and asked them to complete explicit measures to rate themselves and their partner on warmth and competence. The questionnaire ended with a series of questions about demographics. Finally, participants were debriefed, thanked, and dismissed. We report all items that participants completed.

General knowledge questions

All 30 questions pertained to general knowledge. A pretest confirmed that they were of medium or high difficulty.[3]

Dependent variable

Participants first rated the other and the self on 6 competence traits ('has a good culture', 'has a good knowledge of culture',[4] 'competent', 'brilliant', 'intelligent', 'gifted', the first two based on Ross et al.) using scales ranging from 1 (Not at all) to 9 (Totally) before rating the self and the other on 10 traits measuring warmth ('nice', 'unlikeable', 'unpleasant', 'agreeable', 'warm', 'sociable', 'distant', 'disagreeable', 'likeable', 'cold').

Results

We averaged the ratings for both dimensions and created global scores of competence and warmth for answerer and questioner targets (Cronbach's alphas from 0.76 to 0.88) (see **Figure 1**). We conducted a 2 (role: answerer vs. questioner) × 2 (target: answerer vs. questioner) × 2 (dimension: competence vs. warmth) repeated measures ANOVA. The analysis revealed a highly significant dimension effect, $F(1,29) = 84.37$, $p < 0.001$, $\eta_p^2 = 0.74$, such that warmth ratings were higher ($M = 7.02$, $SD = 0.49$) than competence ratings ($M = 5.66$, $SD = 0.66$). There was no significant role effect, $F(1,29) = 3.17$, $p = 0.09$, $\eta_p^2 = 0.10$, nor target effect,
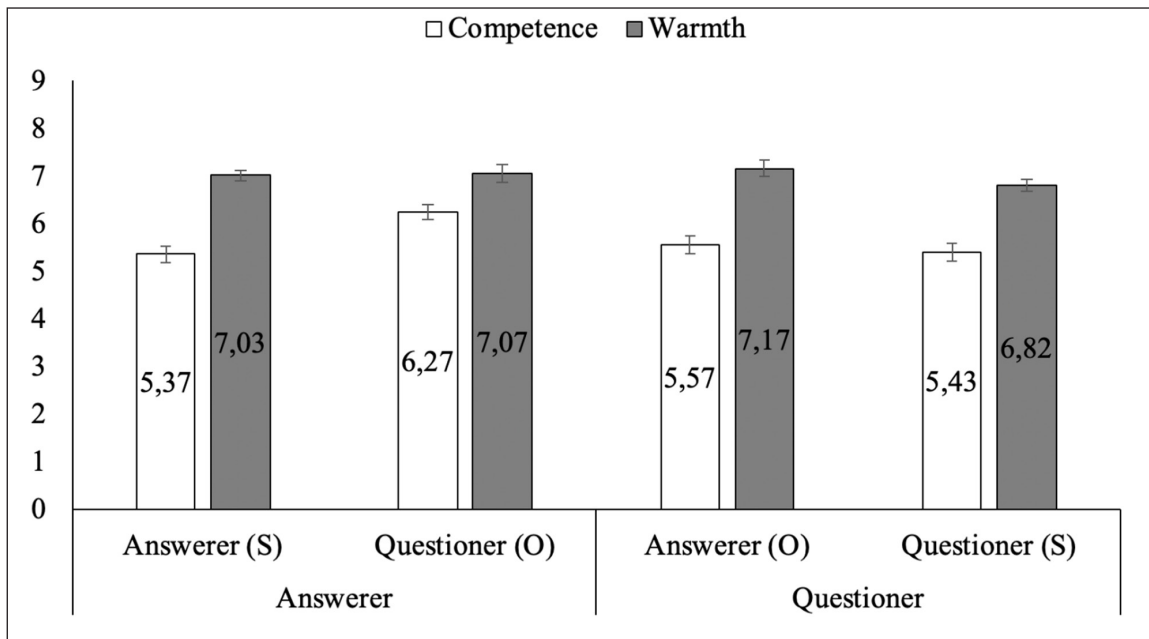
Figure 1: Means of competence and warmth ratings and Standard Errors of the means (as error bars) as a function of role and target (Experiment 1).
Note: Here (S) refers to the judgements of the self, whereas (O) to the judgments of other.

$F(1,29) = 2.09$, $p = 0.16$, $\eta_p^2 = 0.07$. Importantly, the three-way interaction between role, target, and dimension was significant, $F(1,29) = 5.46$, $p = 0.03$, $\eta_p^2 = 0.16$.

To probe this interaction, we first analyzed ratings at the interpersonal level and conducted separate 2 (role: answerer vs. questioner) × 2 (target: answerer vs. questioner) repeated measures ANOVAs on competence and warmth. The interaction for competence was significant, $F(1,29) = 20.40$, $p < 0.001$, $\eta_p^2 = 0.41$. As predicted and replicating the Quiz Master pattern (Ross et al., 1977), answerers saw the questioner as more competent ($M = 6.27$, $SD = 0.84$) than themselves ($M = 5.37$, $SD = 0.90$), $F(1,29) = 38.96$, $p < 0.001$, $\eta_p^2 = 0.57$, whereas questioners' ratings of the answerer ($M = 5.57$, $SD = 0.98$) and of themselves ($M = 5.43$, $SD = 1.03$) did not differ significantly, $F(1,29) = 0.70$, $p = 0.41$, $\eta_p^2 = 0.02$. As for warmth, the role by target interaction was not significant, $F(1,29) = 2.71$, $p = 0.11$, $\eta_p^2 = 0.09$.

Alternatively, we probed the three-way interaction with separate 2 (role: answerer vs. questioner) × 2 (dimension: competence vs. warmth) repeated measures ANOVA for other and self ratings separately. Turning to the other ratings first, this interaction was significant, $F(1,29) = 4.81$, $p = 0.04$, $\eta_p^2 = 0.14$. Competence ratings of the other were higher for answerers ($M = 6.27$, $SD = 0.84$) than for questioners ($M = 5.57$, $SD = 0.98$), $F(1,29) = 9.63$, $p < 0.01$, $\eta_p^2 = 0.25$, confirming that answerers considered their partner to be more competent than questioners did. Warmth ratings of the other did not differ between answerers ($M = 7.07$, $SD = 0.97$) and questioners ($M = 7.17$, $SD = 0.93$), $F(1,29) = 0.22$, $p = 0.64$, $\eta_p^2 = 0.01$. For the ratings of the self, the same 2 (role: answerer vs. questioner) × 2 (dimension: competence vs. warmth) repeated

measures ANOVA revealed no significant interaction, $F(1,29) = 1.29$, $p = 0.27$, $\eta_p^2 = 0.04$.

Discussion

The present experiment not only aimed at replicating the differential competence judgments of answerers of the classic Quiz paradigm but also, and more importantly, at examining the emergence compensation on the warmth judgment. Regarding the interpersonal level, that is the differences between the judgments of both answerer and questioner targets, our data showed that the predicted pattern emerged on competence. Specifically, answerers attributed more competence to the questioner than to themselves whereas questioners made no such difference. The differences in the judgements of warmth were not statistically significant, in line with the prediction that answerers would be reluctant to express differences between themselves and their partner on warmth. Turning to the intrapersonal level, the analysis of the ratings about the other revealed the predicted difference on competence. In line with Kervyn et al.'s (2009b) earlier findings, no differences emerged in the judgments of the self.

Because Experiment 1 only looked at evaluations of participants involved in the interaction, it remains difficult to probe whether answerers were reluctant to express self-favoring differences on warmth or if no compensation was at work. Any indication that answerers indeed appeared warmer than questioners in this setting would help tease apart these two interpretations. Clearly, one would need to know how uninvolved third parties appraise the members of the dyads. To address this issue, we conducted a second experiment in which we not only increased the number of

participants to secure a higher level of power but also, and more crucially, added an observer in the situation.

A first goal of Experiment 2 was to ascertain the fact that, replicating earlier findings (Judd et al., 2005; Kervyn et al., 2009b), a compensation pattern would indeed emerge for observers. Our second goal was to examine the concomitant pattern of judgments for answerers and questioners as we did in Experiment 1 albeit with more power. If compensation does occur for observers and but does not for answerers, this would certainly lend credence to our interpretation that answerers indeed come across as warmer compared to questioner but possibly refrain from saying so when evaluating themselves.

Experiment 2

In Experiment 2, participants came to the lab in triads and were randomly assigned to the roles of observer, answerer or questioner. All three participants rated their perception of the answerer and the questioner on warmth and competence. Our predictions for members of the interaction were the same as in Experiment 1, with the addition of observers whom we predicted would show compensation.

Method

Participants

We used the PANGEA web app (jakewestfall.shinyapps. io/pangea/) to compute the number of triads needed to achieve a power of 80% to detect a medium effect ($d = 0.45$) three-way interaction with a 0.05 alpha two-tailed criterion. The minimum required number of triads suggested by PANGEA was 40, meaning 120 participants. To accommodate for potential data loss, 153 second-year psychology students ($M_{Age} = 20.44$, $SD_{Age} = 3.26$, 19 men) from a large European university came to the laboratory in triads in exchange for course credit. The instructions made clear that they had to participate with students whom they did not know beforehand. We used the same cover story as in Experiment 1.

Procedure and materials

An ostensibly random draw assigned participants to one of three roles, i.e., questioner, answerer or observer. The instructions for answerers and questioners were the same as in Experiment 1. Observers had to monitor the way the two contestants behaved during the quiz interaction and to write down their impressions about them. After the quiz, the experimenter directed participants to individual rooms and asked them to rate the questioner and the answerer on warmth and competence. Before the debriefing, participants completed a series of questions pertaining to demographics. We report all items that participants completed.

Dependent variable

Participants first rated both the answerer and the questioner on the same 6 competence traits from Experiment 1, and on 4 warmth traits ('warm', 'sociable', 'agreeable', 'sympathetic') by way of a computerized questionnaire that used sliders ranging from 0 to 100, thereby following Ross et al.'s (1977) procedure. The traits appeared ran-

domly within each dimension, and the two sliders representing answerer and questioner targets always appeared simultaneously on the screen.

Results

We created two scores of competence and warmth for both targets (Cronbach's alphas from 0.81 to 0.91). Because observers evaluated targets other than themselves whereas answerers and questioners each rated themselves along with their counterpart, we conducted separate analysis for observers on the one hand and the two other roles on the other.

Observers

We submitted observers' competence and warmth ratings to a 2 (target: answerer vs. questioner) × 2 (dimension: competence vs. warmth) repeated measures ANOVA. There was a dimension main effect, $F(1,50) = 32.94$, $p < 0.001$, $\eta_p^2 = 0.40$, such that warmth ratings were higher ($M = 71.78$, $SD = 12.78$) than competence ratings ($M = 60.41$, $SD = 11.71$), but no target main effect, $F(1,50) = 0.01$, $p = 0.93$, $\eta_p^2 < 0.01$. More importantly, the interaction was significant, $F(1,50) = 13.45$, $p < 0.001$, $\eta_p^2 = 0.21$, revealing the presence of a compensation pattern. Observers rated the questioners more competent ($M = 62.75$, $SD = 11.52$) than the answerers ($M = 58.07$, $SD = 14.51$), $F(1,50) = 8.03$, $p < 0.01$, $\eta_p^2 = 0.14$, and perceived the answerers warmer ($M = 73.98$, $SD = 14.85$) than the questioners ($M = 69.57$, $SD = 15.53$), $F(1,50) = 3.72$, $p = 0.06$, $\eta_p^2 = 0.07$, although this latter difference was only marginal (see **Figure 2**).

Answerers and questioners

As far as the answerers and questioners were concerned, we conducted on a 2 (role: answerer vs. questioner) × 2 (target: answerer vs. questioner) × 2 (dimension: competence vs. warmth) repeated measures ANOVA. This analysis revealed a significant dimension main effect, $F(1,50) =$
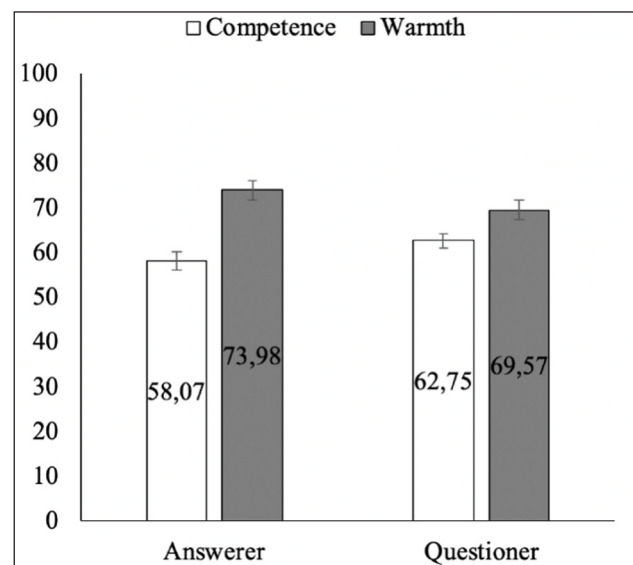
Figure 2: Means of competence and warmth ratings and Standard Errors of the means (as error bars) as a function of target for Observer participants (Experiment 2).

86.63, $p < 0.001$, $\eta_p^2 = 0.63$, such that ratings of warmth were higher ($M = 71.67$, $SD = 11.78$) than ratings of competence ($M = 54.92$, $SD = 9.93$). The role main effect was also significant, $F(1,50) = 4.82$, $p = 0.03$, $\eta_p^2 = 0.09$, meaning that questioners gave higher ratings ($M = 65.6$, $SD = 11.64$) than answerers did ($M = 60.99$, $SD = 11.42$). The target main effect was not significant, $F(1,50) = 2.82$, $p = 0.10$, $\eta_p^2 = 0.05$.

The predicted three-way interaction between role, target and dimension was not significant, $F(1,50) = 2.83$, $p = 0.10$, $\eta_p^2 = 0.05$. Because this interaction was significant in the previous experiment, and given that our *a priori* hypotheses focused on the interpersonal level and the interactions for each dimension, we further probed the three-way interaction by conducting separate 2 (role: answerer vs. questioner) × 2 (target: answerer vs. questioner) repeated measures ANOVAs on competence and warmth. The interaction for competence was significant, $F(1,50) = 16.53$, $p < 0.001$, $\eta_p^2 = 0.25$, replicating the Quiz Master pattern (Ross et al., 1977). Answerers attributed more competence to the questioner ($M = 59.7$, $SD = 13.76$) than to themselves ($M = 46.6$, $SD = 18.01$), $F(1,50) = 22.29$, $p < 0.001$, $\eta_p^2 = 0.31$, whereas questioners' ratings of the answerer ($M = 56.28$, $SD = 15.67$) and of themselves ($M = 57.11$, $SD = 11.39$) were not statistically different, $F(1,50) = 0.30$, $p = 0.59$, $\eta_p^2 < 0.01$. Regarding warmth, the interaction between role and target was close to significance, $F(1,50) = 3.94$, $p = 0.05$, $\eta_p^2 = 0.07$. Answerers' ratings of the questioner ($M = 69.21$, $SD = 19.18$) and of themselves ($M = 68.45$, $SD = 19.33$) did not differ significantly, $F(1,50) = 0.06$, $p = 0.80$, $\eta_p^2 < 0.01$. In contrast, questioners judged the answerer warmer ($M = 77.32$, $SD = 15.41$) than themselves ($M = 71.71$, $SD = 16.46$), $F(1,50) = 11.98$, $p < 0.001$, $\eta_p^2 = 0.19$ (see **Figure 3**).

Again, we also we probed the three-way interaction with separate 2 (role: answerer vs. questioner) × 2 (dimension: competence vs. warmth) repeated measures ANOVA for ratings of the other and of the self. This interaction was significant for the other ratings, $F(1,50) = 9.53$, $p < 0.01$, $\eta_p^2 = 0.16$. Competence ratings of the other did not differ between answerers ($M = 59.7$, $SD = 13.76$) and questioners ($M = 56.28$, $SD = 15.67$), $F(1,50) = 1.5$, $p = 0.23$, $\eta_p^2 = 0.03$. Regarding warmth, ratings of the other were higher for questioners ($M = 77.32$, $SD = 15.41$) than answerers ($M = 69.21$, $SD = 19.18$), $F(1,50) = 6.56$, $p = 0.01$, $\eta_p^2 = 0.12$, showing that questioners considered their partner to be warmer than answerers did. Finally, the same ANOVA on the ratings of the self confirmed the presence of a significant interaction, $F(1,50) = 5.29$, $p = 0.03$, $\eta_p^2 = 0.10$. Regarding competence, ratings of the self were lower for answerers ($M = 46.6$, $SD = 18.01$) than questioners ($M = 57.11$, $SD = 11.39$), $F(1,50) = 15.79$, $p < 0.001$, $\eta_p^2 = 0.24$, indicating that answerers considered themselves to be less competent than questioners did. For warmth, ratings of the self did not differ between answerers ($M = 68.45$, $SD = 19.33$) and questioners ($M = 71.71$, $SD = 16.45$), $F(1,50) = 0.88$, $p = 0.35$, $\eta_p^2 = 0.02$.

Discussion

Experiment 2 not only aimed at replicating results from Experiment 1 but also at ascertaining the emergence of compensation in the eyes of external observers. Consistent with our predictions, compensation emerged for observers, although the difference on warmth only came close to significance. Regarding the contestants, interpersonal differences showed that the predicted pattern emerged on competence. Specifically, answerers attributed more competence to the questioner than to themselves whereas
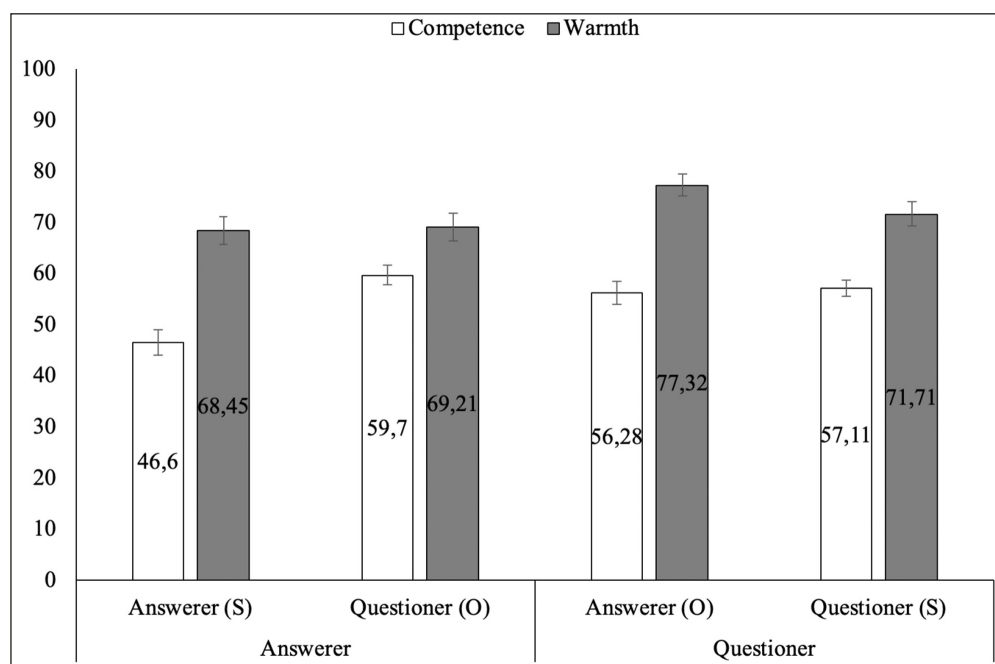
Figure 3: Means of competence and warmth ratings and Standard Errors of the means (as error bars) as a function of target for Answerer and Questioner participants (Experiment 2).
Note: Here (S) refers to the judgements of the self, whereas (O) to the judgments of other.

questioners made no such difference. As for warmth, we found no significant difference between judgments of both targets for answerers. In light of what we found for observers' data, this suggests that answerers hesitated to declare their superiority on warmth even though they admitted a disadvantage on competence. Regarding questioners, they saw their partner to be warmer than they judged themselves. Turning to intrapersonal ratings, and contrary to Experiment 1, the judgments about the self revealed differences on competence whereas judgments about the other revealed differences on warmth.

Interestingly enough, both members of the interaction distinguished one contestant as being the more competent one, and the other as being the warmer one. However, these differences never resulted from participants placing themselves above the other. It seems that both participants were reluctant to admit their advantage on their more favorable dimension, even though observers reported such a difference. This pattern suggests that, when it comes to judging ourselves compared to another person, admitting our own superiority comes across as undesirable at best and risky at worst.

To see whether these findings would replicate, we conducted a third experiment again using triads. A second goal of Experiment 3 was to explore if these patterns would extend to a more comprehensive set of traits. Indeed, recent research suggests that both competence and warmth each encompass two facets (Abele et al., 2016; Carrier et al., 2014). Specifically, competence comprises *ability* and *assertiveness* facets and warmth comprises *friendliness* and *morality* facets. While *ability* refers to the skills that individuals possess to act on their intention and *assertiveness* to the motivation to carry out these intentions (Carrier et al., 2014), *friendliness* relates to maintaining cooperative and affectionate relations with others and *morality* to maintaining relations in adherence with shared moral values (Abele et al., 2016). In light of this, we wanted to check whether the obtained pattern on competence (i.e., ability in the previous experiments) would be the same on the assertiveness facet. Given that assertiveness refers more directly to social status than to ability (Carrier et al., 2014), we were interested to see whether questioners who did not voice any superiority on ability would acknowledge their situational upper hand on assertiveness given that they are those in charge of asking questions and checking answers. Similarly, it would be interesting to see if the pattern obtained on warmth (i.e., friendliness in the previous experiments) would also emerge when looking at the facet of morality. Work by Abele et al. (2016) indicated that friendliness more strongly relates to the Big Five's 'Extraversion' and 'Agreeableness' personality factors than morality. This distinction suggests that our previous pattern should replicate more readily on friendliness than on morality. Having said this, Experiment 3 remained mostly exploratory because very little work examined the four facets and no studies concerned compensation.

Experiment 3

Experiment 3 used the same procedure and design as Experiment 2. Participants rated their perception of the answerer and the questioner on warmth, i.e., friendliness and morality, and competence, i.e., ability and assertiveness. Our predictions were similar to Experiment 2 for answerers and observers, but given the results of Experiment 2 for questioners, we expected to replicate this pattern with our new set of traits.

Method

Participants

A total of 123 psychology students (M_{Age} = 20.95, SD_{Age} = 4.51, 13 men) from a large European university came to the laboratory in triads in exchange for course credit. As in the previous experiments, they participated with students they did not previously know.

Procedure and materials

The experimenter assigned participants to the roles of questioner, answerer or observer again using an ostensibly random draw. The instructions were the same as in Experiment 2. After the quiz, participants individually rated the questioner and the answerer on warmth and competence using the four facets. Before the debriefing, participants provided their demographics. We report all items that participants completed.

Dependent variable

We adapted the traits in order to tap all four facets of the two dimensions. Participants first rated the answerer and the questioner on competence, i.e., ability ('competent', 'efficient', 'intelligent') and assertiveness ('competitive', determined', 'self-confident'), and warmth, i.e., friendliness ('warm', 'sociable', 'considerate') and morality ('reliable', 'moral', 'sincere') (Abele et al., 2016; Carrier et al., 2014) using sliders ranging from 0 to 100. As in Experiment 2, we randomized the order of traits within dimension, and the two sliders representing both targets appeared simultaneously on the screen.

Results

We computed ability, assertiveness, friendliness, and morality scores for both targets (Cronbach's alphas from 0.44 to 0.84, for the analysis of general warmth and competence, see supplementary materials here: https://osf.io/73xyv/?view_only=9072bc57963b449bbced9ec61304cf21). Again, because ratings differed in nature for observers compared to answerers and questioners, we conducted separate analysis for observers.

Observers

The 2 (target: answerer vs. questioner) × 4 (facet: ability vs. assertiveness vs. friendliness vs. morality) interaction was significant, $F(3,40)$ = 11.35, p < 0.001, η_p^2 = 0.22. In light of our concern for the different facets, we decided to look at the facets for each dimension separately. Looking first at the facets for competence, there was a facet main effect, $F(1,40)$ = 43.80, p < 0.001, η_p^2 = 0.52, showing that overall ratings of ability were higher (M = 64.3, SD = 15.43) than ratings of assertiveness (M = 47.49, SD = 16.58). Moreover, the target main effect was significant, confirming that observers generally attributed more ability and more assertiveness to the questioner (M = 62.1,

$SD = 13.06$) than to the answerer ($M = 49.69$, $SD = 19.08$), $F(1,40) = 20.52$, $p < 0.001$, $\eta_p^2 = 0.34$. The interaction between target and facet was not significant, $F(1,40) = 0.87$, $p = 0.36$, $\eta_p^2 = 0.02$, confirming that the pattern was the same for both facets of competence.

As for the facets of warmth, the target main effect was not significant, $F(1,40) = 0.54$, $p = 0.47$, $\eta_p^2 = 0.01$, but there was a facet main effect, $F(1,40) = 9.88$, $p < 0.01$, $\eta_p^2 = 0.20$, showing that overall ratings of morality were higher ($M = 69.63$, $SD = 14.15$) than ratings of friendliness ($M = 62.91$, $SD = 16.58$). Moreover, the interaction between target and facet was significant, $F(1,40) = 9.00$, $p < 0.01$, $\eta_p^2 = 0.18$. Observers attributed more friendliness to answerers ($M = 65.02$, $SD = 15.97$) than questioners ($M = 60.8$, $SD = 19.5$), $F(1,40) = 4.33$, $p = 0.04$, $\eta_p^2 = 0.10$, whereas they was no statistical difference on morality between their ratings of the answerers ($M = 68.63$, $SD = 14.26$) and of the questioners ($M = 70.63$, $SD = 15.83$), $F(1,40) = 1.55$, $p = 0.22$, $\eta_p^2 = 0.04$ (see **Figure 4**).

In sum, the expected pattern on competence emerged in observers' ratings on both its facets of ability and assertiveness. The predicted pattern also emerged on warmth although it only showed on friendliness.

Answerers and questioners

We again conducted our analysis on the four facets of competence and warmth. While the 2 (role: answerer vs. questioner) × 2 (target: answerer vs. questioner) × 4 (facet: ability vs. assertiveness vs. friendliness vs. morality) interaction was significant, $F(3,40) = 8.21$, $p < 0.001$, $\eta_p^2 = 0.17$, we again looked at the facets for each dimension separately.

We first conducted our analysis on competence, using a 2 (role: answerer vs. questioner) × 2 (target: answerer vs. questioner) × 2 (facet: ability vs. assertiveness). Because the three-way interaction was significant, $F(1,40) = 10.71$, $p < .01$, $\eta_p^2 = .21$, we conducted separate 2 (target:

answerer vs. questioner) × 2 (facet: ability vs. assertiveness) repeated measures analyses for each role. Regarding answerers, there was a facet main effect, $F(1,40) = 14.90$, $p < 0.001$, $\eta_p^2 = 0.27$, such that ratings of ability were higher ($M = 61.23$, $SD = 11.78$) than ratings of assertiveness ($M = 53.73$, $SD = 13.77$). There was also a target main effect, $F(1,40) = 40.55$, $p < 0.001$, $\eta_p^2 = 0.50$, as ratings of questioners were higher ($M = 66.35$, $SD = 11.65$) than ratings of answerers ($M = 48.61$, $SD = 16.58$). There was also a two-way interaction, $F(1,40) = 9.25$, $p < 0.01$, $\eta_p^2 = 0.19$. Specifically, answerers attributed more assertiveness to the questioner ($M = 59.68$, $SD = 16.07$) than to themselves ($M = 47.78$, $SD = 19.27$), $F(1,40) = 11.65$, $p < 0.001$, $\eta_p^2 = 0.23$, and even more ability to the questioner ($M = 73.02$, $SD = 13.27$) than to themselves ($M = 49.44$, $SD = 17.91$), $F(1,40) = 51.76$, $p < 0.001$, $\eta_p^2 = 0.56$.

As for questioners, whereas the target effect had no impact, $F(1,40) = 2.26$, $p = 0.14$, $\eta_p^2 = 0.05$, the facet effect came out significant, $F(1,40) = 31.90$, $p < 0.001$, $\eta_p^2 = 0.44$, such that ratings of ability were higher ($M = 57.6$, $SD = 16.58$) than ratings of assertiveness ($M = 45.4$, $SD = 15.88$). The interaction between target and facet was not significant, $F(1,40) = 2.33$, $p = 0.14$, $\eta_p^2 = 0.06$. Combined with the absence of a target main effect, this suggests that questioners did not report any difference between themselves and the answer on neither facet of competence.

Turning to the facets of warmth, we conducted a 2 (role: answerer vs. questioner) × 2 (target: answerer vs. questioner) × 2 (facet: friendliness vs. morality) repeated ANOVA. The three-way interaction was significant, $F(1,40) = 4.01$, $p = 0.05$, $\eta_p^2 = 0.09$. In light of this, we conducted separate analyses for each role. Turning to answerers first, there was no target main effect, $F(1,40) = 2.46$, $p = 0.13$, $\eta_p^2 = 0.06$, but a facet main effect, $F(1,40) = 18.92$, $p < 0.001$, $\eta_p^2 = 0.32$, showing that ratings of morality were higher ($M = 73.67$, $SD = 13.77$) than ratings of friendliness
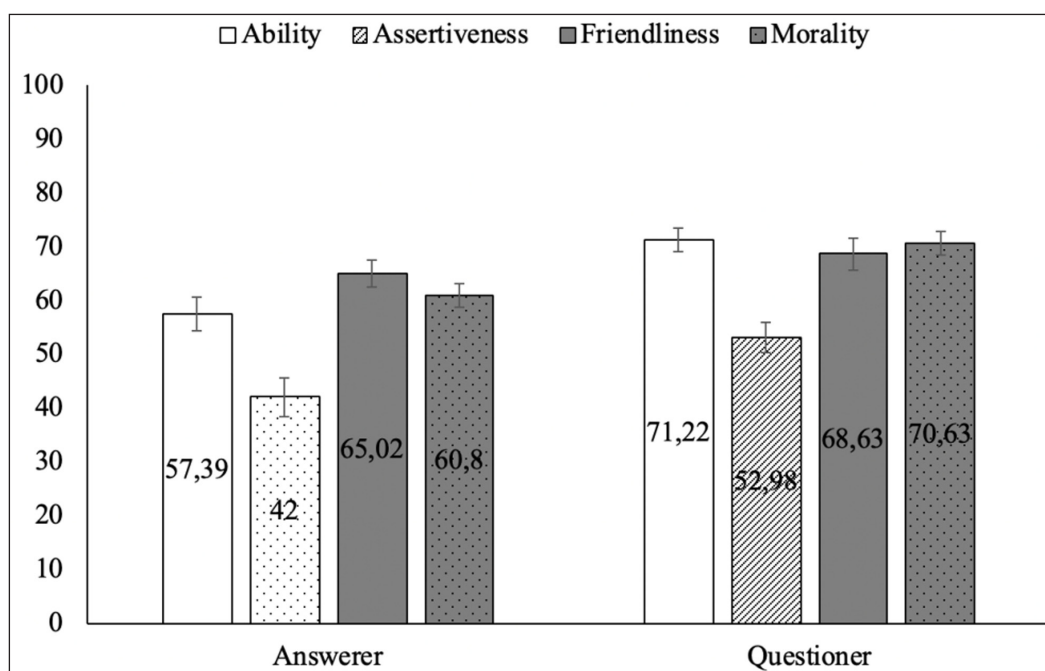
Figure 4: Means of ability, assertiveness, friendliness and morality ratings and Standard Errors of the means (as error bars) as a function of target for Observer participants (Experiment 3).

($M = 63.76$, $SD = 16.78$). The two-way interaction was significant, $F(1,40) = 24.61$, $p < 0.001$, $\eta_p^2 = 0.38$. Specifically, answerers did not make any difference on friendliness between themselves ($M = 65.83$, $SD = 16.04$) and their partner ($M = 61.68$, $SD = 21.12$), $F(1,40) = 2.51$, $p = 0.12$, $\eta_p^2 = 0.06$, thereby replicating the findings in Experiment 2. Unexpectedly, they attributed more morality to their partner ($M = 78.54$, $SD = 14.24$) than to themselves ($M = 68.8$, $SD = 15.72$), $F(1,40) = 27.85$, $p < 0.001$, $\eta_p^2 = 0.41$.

There was a target main effect for questioners, $F(1,40) = 7.15$, $p = 0.01$, $\eta_p^2 = 0.15$, such that ratings were higher for answerers ($M = 65.26$, $SD = 17.29$) than for questioners ($M = 62.16$, $SD = 17.22$). There was also a facet main effect, $F(1,40) = 12.93$, $p < 0.001$, $\eta_p^2 = 0.24$, such that ratings of morality were higher ($M = 69.39$, $SD = 17.93$) than ratings of friendliness ($M = 58.02$, $SD = 21.26$). The patterns between friendliness and morality were also different, $F(1,40) = 6.93$, $p = 0.01$, $\eta_p^2 = 0.15$. Specifically, questioners attributed more friendliness to their partner ($M = 61.15$, $SD = 22.27$) than to themselves ($M = 54.9$, $SD = 22.23$), $F(1,40) = 9.47$, $p < 0.01$, $\eta_p^2 = 0.19$, again replicating the pattern observed in Experiment 2, while there was no statistical difference on morality between the evaluation of their partner ($M = 69.37$, $SD = 18.44$) and of themselves ($M = 69.41$, $SD = 18.21$), $F(1,40) < 0.01$, $p = 0.97$, $\eta_p^2 < 0.01$ (see **Figure 5**).

We also conducted our analysis for ratings of self and other separately. Results showed that, as in Experiment 2, differences in judgments emerged in both self and other evaluations (for the complete analysis, see supplementary materials here: https://osf.io/73xyv/?view_only=9072bc 57963b449bbced9ec61304cf21).

Discussion
Experiment 3 aimed at replicating Experiment 2 and the Quiz Master paradigm (Ross et al., 1977) using a more comprehensive set of traits and exploring how previous patterns would extend to all dimension facets. Again,

observers saw questioners as being more competent, that is, more able and more assertive, than answerers. As predicted and in line with Experiment 2, a compensation pattern emerged in the eyes of observers on warmth, albeit only on the friendliness facet.

Turning to the two contestants, the predicted pattern emerged on competence when looking at interpersonal ratings. Specifically, answerers attributed more competence (ability and assertiveness) to the questioner than to themselves. In contrast, questioners' ratings did not differ significantly. As in Experiment 2, no significant difference on warmth (i.e., friendliness) emerged for answerers, once again supporting the hypothesis that answerers might be reluctant to express self-superiority on warmth. Morality ratings revealed an unexpected pattern in that answerers attributed more morality to their partner than to themselves. As in Experiment 2, questioners favored their partner on warmth although this was restricted to friendliness ratings with no difference on morality ratings. In our view, these findings suggest that future research on these issues would benefit from making finer distinctions between the facets of warmth and competence (Abele et al., 2008, 2016; Carrier et al., 2014, 2019).

Bayesian Mini Meta-analyses
The frequentist approach used in the three experiments does not allow us to draw any firm conclusion in support of the null hypotheses with respect to the judgments of questioners on competence and warmth on the hand and the judgements of answerers on warmth on the other. We therefore turned to Bayesian mini meta-analyses (Rouder & Morey, 2011) in order to determine the ratio (Bayes Factor$_{01}$ or BF_{01}) of the probability to observe our data given H0 (i.e., no differences between the judgements of the answerer and the questioner) to the probability to observe our data given H1 (differences between the judgements of the answerer and the questioner; Etz et al., 2018). To analyze these effects in the most coherent way across experi-
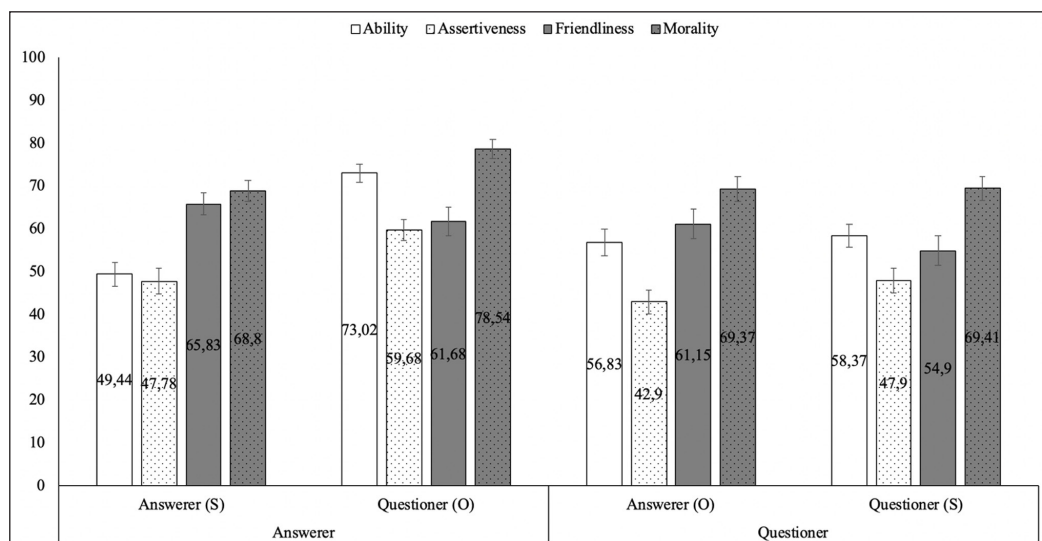
Figure 5: Means of ability, assertiveness, friendliness and morality ratings and Standard Errors of the means (as error bars) as a function of target for Answerer and Questioner participants (Experiment 3).
Note: Here (S) refers to the judgements of the self, whereas (O) to the judgments of other.

ments, we took into account only the facets of ability and friendliness from Experiment 3 as they related the most with the traits used in the two previous experiments. We used the R 'bayesmeta' package (Röver, in press), and the prior recommended by Dienes (2011), which assumed a half normal distribution and a scale of 0.5 (i.e., the equivalent of a SD of 0.50 for the full normal distribution). This distribution assumes that the effect sizes are greater than zero, with a mode at the minimum (0), which is conservative if H1 is true. The SD of 0.5 means that most effects are under a d = 0.50 (a medium effect size according to Cohen [1988]). Note, however, that increasing the SD up to 1 does not have any noticeable effect on the Bayes factors (see script here: https://osf.io/73xyv/?view_only=9072bc57963b449bbced9ec61304cf21).

The results for the three experiments strongly support the null hypothesis for two predicted null effects, i.e., the absence of differences in answerers' judgements on warmth and in questioners' judgements on competence. In contrast, there was no support for H0 or H1 for the non-predicted effect that emerged in two of the three experiments, that is, a difference in questioners' judgements of warmth. Specifically, the data that we observed for answerers' judgements of warmth are more likely to occur under H0 ($BF_{01} = 47.71$), as are the data for questioners' judgements of competence ($BF_{01} = 34.24$). As for the data for questioners' judgements of warmth, the analysis revealed no support for either H0 or H1 ($BF_{01} = 1.46$). Given that a BF_{01} greater than 10 is considered as 'strong evidence' in favor of H0, whereas a BF_{01} smaller than 3 is considered to be 'anecdotal evidence' (Jeffreys, 1961), the data for answerers' judgements of warmth and questioners' judgements of competence showed strong support in favor of H0, whereas the data for questioners' judgements of warmth failed to lean toward either H0 or H1.

General Discussion

An impressive and still growing amount of research shows that the social world is governed by two fundamental dimensions, i.e., warmth and competence (for reviews, Abele & Wojciszke, 2018; Fiske, 2018; Yzerbyt, 2018). The relation characterizing these two dimensions is still a matter of debate but a substantial body of evidence supports the idea of a negative relation between warmth and competence when it comes to intergroup perception, a pattern known as compensation. But although numerous efforts reveal that compensation shows up in intergroup perception, interpersonal perception, person perception, and self-presentation, only one study to date examined the emergence of a compensatory pattern of judgment in interpersonal comparisons involving the self (Kervyn et al., 2009b).

In three experiments, we adapted the classical Quiz Master paradigm to create a situation of interpersonal interaction producing a differential on competence. The data consistently replicated the Quiz Master pattern in that both observers and answerers, but not questioners, perceived questioners as more competent. More importantly, and consistent with previous findings on compensation (Kervyn et al., 2009b; Yzerbyt, 2018), our results

showed that when evaluating the two targets, observers compensated questioners' advantage on competence by perceiving answerers as warmer (although this difference was marginal in Experiment 2).

Turning to the members of the interaction, ratings of the answerer and the questioner showed that participants were reluctant to express self-favoring differences. Indeed, answerers did not compensate their disadvantage on competence, i.e., ratings on warmth between themselves and their counterpart did not differ significantly. As for questioners, they also did not manifest their upper hand on competence, even though they perceived the answerer to be warmer than themselves (Experiments 2 and 3). In other words, both members of the interaction elevated their partner on the dimension on which they themselves were disadvantaged, rather than elevating themselves on the opposite dimension. Regarding the intrapersonal level of analysis and contrary to our predictions, these differences emerged in both ratings about the self and the other. Bayesian mini meta-analyses showed that the non-difference found in answerers' judgments of warmth and in questioners' judgments of competence were extremely more likely to occur under H0, thereby conferring strong support for this absence of effect, whereas the non-predicted difference expressed by questioners on warmth did not favor H1 more than H0.

Experiment 3 offered the opportunity to explore further the facets of the two dimensions. Regarding competence, the differential Quiz Master pattern surfaced in both facets in the eyes of observers and answerers, such that questioners came across higher on ability and assertiveness than answerers. Turning to warmth, compensation emerged on friendliness, such that observers judged the answerers friendlier than they judged questioners. A similar pattern emerged among questioners.

Our results represent an initial contribution in the investigation of compensation in interpersonal comparison involving the self as an individual. To the best of our knowledge, only Kervyn et al. (2009b) examined compensation in this type of situation. However, in their experiment, the self was a representative of its assigned group and not a unique individual. Our preliminary results support the idea that compensation might not occur similarly in intergroup versus interpersonal comparisons. Indeed, it seems to be more difficult for individuals to rely on compensation when they compare to another person as an individual than as a member of their ingroup.

Cambon and Yzerbyt (2018) showed that groups use compensation for different reasons depending on their status. Specifically, low-status groups rely on compensation as a means to protect self-esteem (whereas high-status groups see it as a response to norms of non-discrimination). Indeed, when low-status group members had a chance to self-affirm and protect their self-esteem before the comparison with the high-status group, compensation did not emerge. Because our situation created a threat on answerers' sense of competence, and because this dimension is closely linked to self-esteem (Abele & Wojciszke, 2014), one hypothesis was that Cambon and Yzerbyt's pattern could replicate in our situation and that

answerers would manage this threatening comparison by using compensation. Yet this hypothesis was not confirmed by our data. This suggests that although Cambon and Yzerbyt (2018) showed that compensation is a sensible strategy to justify social hierarchy; this same strategy might not apply to interpersonal forms of hierarchy.

As it turns out, we anticipated that an absence of difference in self-other evaluations of warmth for answerers could be the manifestation of a reluctance to appear self-favoring. Indeed, previous findings show that people are often unwilling to express a superiority over others (Dunning et al., 1989; Klein, 2001; Muller & Butera, 2004), especially when the other is a specific individual (Alicke et al., 1995; Perloff & Fetzer, 1986). Studies suggest that this reluctance disappears when participants' self-esteem is threatened (Muller & Butera, 2004), or when the comparison target is in competition with the self (Butera & Mugny, 1995).

Clearly, the inclusion of an observer role in our experimental situation and, more importantly, the emergence of compensation in the eyes of these uninvolved third parties certainly supports the argument that the members of the interaction proper refrained from admitting self-advantage, although external observers did perceive it. In order to understand whether the pattern obtained in our studies is the result of a pressure toward modesty or a genuine absence of compensation in interpersonal comparisons involving the self, future research should examine situations that presumably take away this reluctance to brag from participants. This would require manipulating more explicitly the degree of threat to self-image and/or participants' possibility to reinterpret and minimize their lower standing on competence.

Because the Quiz paradigm presumably has questioners be well aware of the nature of their situational advantage, we did not predict that they would attribute less competence to answerers. This also led us to predict that questioners would not see more warmth in answerers than in themselves. Surprisingly, however, a difference on warmth emerged in the data of Experiments 2 and 3. If this pattern were to be further replicated, it would depart from the rather optimistic interpretation put forth by Ross and colleagues (1977) that questioners simply discount their higher competence because of their situational advantage. It raises the intriguing possibility that questioners did not express self-favoring bias on competence for different reasons than those proposed by Ross and colleagues. Specifically, it may be the case that questioners thought they were more competent than answerers were but were actually reluctant to stress this superiority too bluntly. This rationale allows accounting for the pattern of judgments found on warmth as a form of *noblesse-oblige* effect (Vanbeselaere et al., 2006; Yzerbyt & Cambon, 2017). Indeed, when individuals are put in a position of superiority on one dimension, they often concede superiority on the other dimension to the outgroup as a result of a pressure toward nondiscrimination (Cambon et al., 2015; Cambon & Yzerbyt, 2018; Yzerbyt & Cambon, 2017; Yzerbyt et al., 2008). Interestingly enough,

this *noblesse-oblige* pattern on the part of questioners emerged in the two experiments that included external observers (Experiments 2 and 3). This public setting may have made more salient the normative pressures not to appear judgmental, and encouraged questioners toward magnanimity. At any rate, it would be important for future research to address this intriguing result. One way to go about this would be to manipulate the salience of benevolence norms in the Quiz Master paradigm.

The above rationale stresses a most fascinating aspect of the Quiz Master paradigm. Clearly, because of the ostensibly random assignment to roles, this setting should definitely encourage all participants to doubt the existence of any difference between the contestants. The present data from the observers show that, whereas being in a position of power (asking questions) leads to the attribution of higher competence, being in a position of submissiveness (answering the questions) is conducive to the attribution of higher warmth. This illustrates how the fundamental attribution error intrudes the present setting in order to produce compensatory perceptions. Notably, the questioner was traditionally thought to be immune to this error. The fact that we collected data on both dimensions of judgment allows us to suggest an alternative interpretation – namely that all parties involved are in fact likely to fall prey to a biased perception of the questioners and answerers along compensatory lines, although both contestants appear hesitant to manifest their superiority in the present context.

One potential limitation of the present findings stems from the nature of our sample, consisting mostly of first and second year female psychology students. To be sure, future research should be conducted with more diverse and balanced samples.

Another possible limitation of the present endeavor may reside in the fact that the quiz context created a set of expectations with respect to the behavior of the participants. Specifically, questioners were to ask questions and, by way of cooperation, answerers were to answer them. As such, one might consider that the prediction of compensation has little to do with the cooperative behavior shown by the answerers. Tempting as this interpretation may seem, it overlooks the fact that the work on compensation never argued that the behaviors of the two social targets, be they groups or individuals, are expected to be the same or that compensation is in no way related to actual differences in conduct. Power, prestige, status and resources do shape behaviors as surely as lack of these features does. In their initial demonstration, Yzerbyt et al. (2005) took as a starting point the fact that the characteristics of French, the more competent group, and Belgians, the warmer group, are indeed not entirely similar, a situation largely supported in the ratings of uninterested observers, i.e. the Swiss. Subsequent work by Kervyn et al. (2009b) showed that the compensation effect materialized in the fact that the parties asked their counterpart different questions. This ended up creating different behaviors in the respondents because they conformed to the expected pattern and showed a form of self-fulfilling

prophecy. In sum, whenever there is a social hierarchy, different behaviors emerge and different perceptions follow. As shown by Judd et al. (2005) and Yzerbyt et al. (2008), the fact that people embrace such compensatory perceptions even when the actual information provided to them is (artificially made) neutral only shows the power of social structural constraints in shaping everyday compensatory judgments. In sum, compensation does not emerge in spite of reality but precisely because of reality. This means that the cooperative behavior on the part of the answerers is precisely the evidence that fuels compensation. Still, it would be highly interesting if future research were to manipulate competence in ways that impose as little constraints as possible on the behavior of the less competent person in terms of warmth. In situations such as these, manifestations of higher levels of warmth among the less competent people would constitute even stronger evidence for our compensation hypothesis.

To sum up, our three experiments not only replicate and extend the classic Quiz Master paradigm (Ross et al., 1977), but they also shed light on the dynamics of compensation in most interesting ways. In addition to opening new lines of research in the field of causal attribution and the fundamental attribution error, investigating these various issues should also allow addressing important questions regarding social comparison and compensation in interpersonal relations. This is our agenda for future research.

Notes
[1] We were unable to conduct a proper power analysis based on Ross et al.'s (1977) experiment because too few indicators were available in the original article. We used the PANGEA web app (jakewestfall.shinyapps.io/pangea/) to determine that the effect size that could be detected with the current sample size to achieve a power of 0.80 was $d = 0.52$.

[2] Questioners first had to ask the question, and if the answerer encountered difficulties, to list the four possible answers. They specified the right answer before moving on to the next question.

[3] General knowledge questions were taken from online Quiz websites or made up by the experimenter.

[4] For these two specific items taken from Ross et al. (1977) the anchoring was 1 (worse than most people) to 9 (better than most people).

Acknowledgements

Competing Interests
The authors have no competing interests to declare.

References
Abele, A. E., Cuddy, A. J. C., Judd, C. M., & Yzerbyt, V. (2008). Fundamental dimensions of social judgment. *European Journal of Social Psychology*, *38*(7), 1063–1065. DOI: https://doi.org/10.1002/ejsp.574

Abele, A. E., Ellemers, N., Fiske, S. T., Koch, A., & Yzerbyt, V. (2019). *Navigating the social world: Shared horizontal and vertical dimensions for evaluating self, individuals, and groups*. Manuscript submitted for publication. University of Erlangen.

Abele, A. E., Hauke, N., Peters, K., Louvet, E., Szymkow, A., & Duan, Y. (2016). Facets of the fundamental content dimensions: Agency with competence and assertiveness-communion with warmth and morality. *Frontiers in Psychology*, *7*, 1–17. DOI: https://doi.org/10.3389/fpsyg.2016.01810

Abele, A. E., & Wojciszke, B. (2007). Agency and communion from the perspective of self versus others. *Journal of Personality and Social Psychology*, *93*(5), 751–763. DOI: https://doi.org/10.1037/0022-3514.93.5.751

Abele, A. E., & Wojciszke, B. (2014). Communal and agentic content in social cognition: A dual perspective model. *Advances in Experimental Social Psychology*, *50*, 195–255. DOI: https://doi.org/10.1016/B978-0-12-800284-1.00004-7

Abele, A. E., & Wojciszke, B. (2018). *Agency and Communion in Social Psychology*. London, UK: Routledge. DOI: https://doi.org/10.4324/9780203703663

Abelson, R. P., Kinder, D. R., Peters, M. D., & Fiske, S. T. (1982). Affective and semantic components in political person perception. *Journal of personality and social psychology*, *42*(4), 619–630. DOI: https://doi.org/10.1037/0022-3514.42.4.619

Alicke, M. D., Klotz, M. L., Breitenbecher, D. L., Yurak, T. J., & Vredenburg, D. S. (1995). Personal contact, individuation, and the better-than-average effect. *Journal of Personality and Social Psychology*, *68*(5), 804–825. DOI: https://doi.org/10.1037/0022-3514.68.5.804

Butera, F., & Mugny, G. (1995). Conflict between incompetences and influence of a low-expertise source in hypothesis testing. *European Journal of Social Psychology*, *25*, 457–462. DOI: https://doi.org/10.1002/ejsp.2420250408

Cambon, L., & Yzerbyt, V. (2016). Compensation is for real: Evidence from existing groups in the context of actual relations. *Group Processes & Intergroup Relations* (pp. 1–12). DOI: https://doi.org/10.1177/1368430215625782

Cambon, L., & Yzerbyt, V. (2018). Two routes toward compensation: An investigation into the mechanisms of compensation for high- and low-status groups. *Journal of Experimental Social Psychology*, *77*, 24–35. DOI: https://doi.org/10.1016/j.jesp.2018.04.002

Cambon, L., Yzerbyt, V., & Yakimova, S. (2015). Compensation in intergroup relations: An investigation of its structural and strategic foundations. *British Journal of Social Psychology*, *54*(1), 140–158. DOI: https://doi.org/10.1111/bjso.12067

Carlsson, R., & Björklund, F. (2010). Implicit stereotype content. *Social Psychology*, *41*(4), 213–222. DOI: https://doi.org/10.1027/1864-9335/a000029

Carrier, A., Louvet, E., Chauvin, B., & Rohmer, O. (2014). The primacy of agency over competence in status perception. *Social Psychology*, *45*(5), 347–356. DOI: https://doi.org/10.1027/1864-9335/a000176

Carrier, A., Mierop, A., Corneille, O., & Yzerbyt, V. (2019). *The pragmatics of social perception: How morality and sociability dimensions address relational threats and opportunities.* Manuscript submitted for publication. Catholic University of Louvain.

Cohen, J. (1988). The effect size index: d. *Statistical power analysis for the behavioral sciences*, *2*, 284–288.

Collange, J., Fiske, S. T., & Sanitioso, R. (2014). Maintaining a positive self-image by stereotyping others: Self-threat and the stereotype content model. *Social Cognition*, *27*(1), 1–12. DOI: https://doi.org/10.1521/soco.2009.27.1.138

Dienes, Z. (2011). Bayesian versus orthodox statistics: Which side are you on? *Perspectives on Psychological Science*, *6*(3), 274–290. DOI: https://doi.org/10.1177/1745691611406920

Dunning, D., Meyerowitz, J. A., & Holzberg, A. D. (1989). Ambiguity and self-evaluation: The role of idiosyncratic trait definitions in self-serving assessments of ability. *Journal of Personality and Social Psychology*, *57*(6), 1082–1090. DOI: https://doi.org/10.1037/0022-3514.57.6.1082

Etz, A., Gronau, Q. F., Dablander, F., Edelsbrunner, P. A., & Baribault, B. (2018). How to become a Bayesian in eight easy steps: An annotated reading list. *Psychonomic Bulletin & Review*, *25*, 219–234. DOI: https://doi.org/10.3758/s13423-017-1317-5

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

Fiske, S. T. (2018). Warmth and competence are parallels to communion and agency: Stereotype Content Model. In A. Abele & B. Wojciszke (Eds.), *Agency and Communion in Social Psychology* (pp. 126–141). London, UK: Routledge. DOI: https://doi.org/10.4324/9780203703663-4

Fiske, S. T., Cuddy, A. J., & Glick, P. (2007). Universal dimensions of social cognition: Warmth and competence. *Trends in cognitive sciences*, *11*(2), 77–83. DOI: https://doi.org/10.1016/j.tics.2006.11.005

Fiske, S. T., Cuddy, A. J. C., 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*(6), 878–902. DOI: https://doi.org/10.1037/0022-3514.82.6.878

Fiske, S. T., Xu, J., Cuddy, A. J. C., & Glick, P. (1999). (Dis)respecting versus (dis)liking: Status and interdependence predict ambivalent stereotypes of competence and warmth. *Journal of Social Issues*, *55*(3), 473–489. DOI: https://doi.org/10.1111/0022-4537.00128

Gerber, J. P., Wheeler, L., & Suls, J. (2017). A social comparison theory meta-analysis 60+ years. *Psychological Bulletin*, *144*(2), 177–197. DOI: https://doi.org/10.1037/bul0000127

Glick, P., & Fiske, S. T. (1996). The ambivalent sexisms inventory: Differentiating hostile and benevolent sexism. *Journal of Personality and Social Psychology*, *70*(3), 491–512. DOI: https://doi.org/10.1037/0022-3514.70.3.491

Harris, L. T., & Fiske, S. T. (2006). Dehumanizing the lowest of the low: Neuroimaging responses to extreme outgroups. *Psychological Science*, *17*(10), 847–853. DOI: https://doi.org/10.1111/j.1467-9280.2006.01793.x

Helm, F., Abele, A. E., Müller-Kalthoff, H., & Möller, J. (2017). Applying dimensional comparison theory to the fundamental dimensions of social judgment – Agency and communion. *Learning and Individual Differences*, *54*, 116–125. DOI: https://doi.org/10.1016/j.lindif.2017.01.011

Holoien, D. S., & Fiske, S. T. (2013). Downplaying positive impressions: Compensation between warmth and competence in impression management. *Journal of Experimental Social Psychology*, *49*(1), 33–41. DOI: https://doi.org/10.1016/j.jesp.2012.09.001

Hoorens, V., Pandelaere, M., Oldersma, F., & Sedikides, C. (2012). The hubris hypothesis: You can self-enhance, but you'd better not show it. *Journal of Personality*, *80*(5), 1237–1274. DOI: https://doi.org/10.1111/j.1467-6494.2011.00759.x

Jeffreys, H. (1961). *Theory of Probability* (3rd ed.). Oxford: Oxford University Press, Clarendon Press.

Judd, C. M., James-Hawkins, L., Yzerbyt, V., & Kashima, Y. (2005). Fundamental dimensions of social judgment: Understanding the relations between judgments of competence and warmth. *Journal of Personality and Social Psychology*, *89*(6), 899–913. DOI: https://doi.org/10.1037/0022-3514.89.6.899

Kervyn, N., Bergsieker, H. B., Grignard, F., & Yzerbyt, V. (2016). An advantage of appearing mean or lazy: Amplified impressions of competence or warmth after mixed descriptions. *Journal of Experimental Social Psychology*, *62*, 17–23. DOI: https://doi.org/10.1016/j.jesp.2015.09.004

Kervyn, N., Fiske, S. T., & Yzerbyt, V. (2015). Forecasting the primary dimension of social perception: Symbolic and realistic threats together predict warmth in the stereotype content model. *Social Psychology*, *46*(1), 36–45. DOI: https://doi.org/10.1027/1864-9335/a000219

Kervyn, N., Judd, C. M., & Yzerbyt, V. (2009a). You want to appear competent? Be mean! You want to appear sociable? Be lazy! Group differentiation and the compensation effect. *Journal of Experimental Social Psychology*, *45*(2), 363–367. DOI: https://doi.org/10.1016/j.jesp.2008.08.006

Kervyn, N., Yzerbyt, V., Demoulin, S., & Judd, C. M. (2008). Competence and warmth in context: The compensatory nature of stereotypic views of national groups. *European Journal of Social Psychology*, *38*(7), 1175–1183. DOI: https://doi.org/10.1002/ejsp.526

Kervyn, N., Yzerbyt, V., & Judd, C. M. (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*(1), 155–187. DOI: https://doi.org/10.1080/13546805.2010.517997

Kervyn, N., Yzerbyt, V., & Judd, C. M. (2011). When compensation guides inferences: Indirect and implicit measures of the compensation effect. *European Journal of Social Psychology*, *41*(2), 144–150. DOI: https://doi.org/10.1002/ejsp.748

Kervyn, N., Yzerbyt, V., Judd, C. M., & Nunes, A. (2009b). A question of compensation: The social life of the fundamental dimensions of social perception. *Journal of Personality and Social Psychology*, *96*(4), 828–842. DOI: https://doi.org/10.1037/a0013320

Klein, W. M. (2001). Post hoc construction of self-performance and other performance in self-serving social comparison. *Personality and Social Psychology Bulletin*, *27*(6), 744–754. DOI: https://doi.org/10.1177/0146167201276009

Leach, C. W., Ellemers, N., & Barreto, M. (2007). Group virtue: the importance of morality (vs. competence and sociability) in the positive evaluation of in-groups. *Journal of Personality and Social Psychology*, *93*(2), 234–249. DOI: https://doi.org/10.1037/0022-3514.93.2.234

Lindholm, T., & Yzerbyt, V. (2018). When being nice or being smart could bring you down: Compensatory dynamics in strategic self-presentation. *International Review of Social Psychology*, *31*(16), 1–11. DOI: https://doi.org/10.5334/irsp.136

Locke, K. D. (2014). Agency and communion in social comparisons. In Z. Krizan & F. X. Gibbons (Eds.), *Communal Functions of Social Comparison* (pp. 11–38). Cambridge University Press. DOI: https://doi.org/10.1017/CBO9781139035583.003

Locke, K. D., & Nekich, J. C. (2000). Agency and communion in naturalistic social comparison. *Personality and Social Psychology Bulletin*, *26*(7), 864–874. DOI: https://doi.org/10.1177/0146167200269011

Muller, D., & Butera, F. (2004). On being concerned about bragging. *Cahiers de Psychologie Cognitive*, *22*(2), 159–179.

Perloff, L. S., & Fetzer, B. K. (1986). Self–other judgments and perceived vulnerability to victimization. *Journal of Personality and Social Psychology*, *50*(3), 502–510. DOI: https://doi.org/10.1037/0022-3514.50.3.502

Rosenberg, S., Nelson, C., & Vivekananthan, P. S. (1968). A multidimensional approach to the structure of personality impressions. *Journal of Personality and Social Psychology*, *9*(4), 283–294. DOI: https://doi.org/10.1037/h0026086

Ross, L. D., Amabile, T. M., & Steinmetz, J. L. (1977). Social roles, social control, and biases in social-perception processes. *Journal of Personality and Social Psychology*, *35*(7), 485–494. DOI: https://doi.org/10.1037/0022-3514.35.7.485

Rouder, J. N., & Morey, R. D. (2011). A Bayes factor meta-analysis of Bem's ESP claim. *Psychonomic Bulletin & Review*, *18*(4), 682–689. DOI: https://doi.org/10.3758/s13423-011-0088-7

Röver, C. (in press). Bayesian random-effects meta-analysis using the bayesmeta R package. *Journal of Statistical Software* (pp. 1–51).

Sevillano, V., & Fiske, S. T. (2016). Animals as social objects: Groups, stereotypes, and intergroup threats. *European Psychologist*, *21*(3), 206–217. DOI: https://doi.org/10.1027/1016-9040/a000268

Steele, C. M. (1988). The psychology of self-affirmation: Sustaining the integrity of the self. In Berkowitz, L. (Ed.), *Advances in Experimental Social Psychology*, *21*, 261–302. Academic Press. DOI: https://doi.org/10.1016/S0065-2601(08)60229-4

Swencionis, J. K., & Fiske, S. T. (2016). Promote up, ingratiate down: Status comparisons drive warmth-competence tradeoffs in impression management. *Journal of Experimental Social Psychology*, *64*, 27–34. DOI: https://doi.org/10.1016/j.jesp.2016.01.004

Tafarodi, R. W., & Swann, W. B. (2001). Two-dimensional self-esteem: Theory and measurement. *Personality and Individual Differences*, *31*(5), 653–673. DOI: https://doi.org/10.1016/S0191-8869(00)00169-0

Tajfel, H., & Turner, J. (1979). An integrative theory of intergroup conflict. In: W. Austin & S. Worchel (Eds.), *The Social Psychology of Intergroup Relations* (pp. 56–65). Montery, CA: Brookes-Cole.

Tesser, A. (2000). On the confluence of self-esteem maintenance mechanisms. *Personality and Social Psychology Review*, *4*(4), 290–299. DOI: https://doi.org/10.1207/S15327957PSPR0404_1

Tesser, A. (2001). On the plasticity of self-defense. *Current Directions in Psychological Science*, *10*(2), 66–69. DOI: https://doi.org/10.1111/1467-8721.00117

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*(6), 685–699. DOI: https://doi.org/10.3200/SOCP.146.6.685-699

Van Damme, C., Hoorens, V., & Sedikides, C. (2016). Why self-enhancement provokes dislike: The hubris hypothesis and the aversiveness of explicit self-superiority claims. *Self and Identity*, *15*(2), 173–190. DOI: https://doi.org/10.1080/15298868.2015.1095232

Willis, J., & Todorov, A. (2006). First impressions: Making up your mind after a 100-ms exposure to a face. *Psychological Science*, *17*(7), 592–598. DOI: https://doi.org/10.1111/j.1467-9280.2006.01750.x

Wojciszke, B. (1994). Multiple meanings of behavior: Construing actions in terms of competence or morality. *Journal of Personality and Social Psychology*, *67*(2), 222–232. DOI: https://doi.org/10.1037/0022-3514.67.2.222

Wojciszke, B., Baryla, W., Parzuchowski, M., Szymkow, A., & Abele, A. E. (2011). Self-esteem is dominated

by agentic over communal information. *European Journal of Social Psychology*, *41*(5), 617–627. DOI: https://doi.org/10.1002/ejsp.791

Wojciszke, B., & **Klusek, B.** (1996). Moral and competence-related traits in political perception. *Polish Psychological Bulletin*, *27*, 319–324.

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

Yzerbyt, V. (2018). The dimensional compensation model: Reality and strategic constraints on warmth and competence in intergroup perceptions. In A. Abele & B. Wojciszke (Eds.), *Agency and Communion in Social Psychology* (pp. 126–141). London, UK: Routledge. DOI: https://doi.org/10.4324/9780203703663-11

Yzerbyt, V., & **Cambon, L.** (2017). The dynamics of compensation: When ingroup favoritism paves the way for outgroup praise. *Personality and Social Psychology Bulletin*, *43*(5) 587–600. DOI: https://doi.org/10.1177/0146167216689066

Yzerbyt, V., Kervyn, N., & **Judd, C. M.** (2008). Compensation versus halo: The unique relations between the fundamental dimensions of social judgment. *Personality and Social Psychology Bulletin*, *34*(8), 1110–1123. DOI: https://doi.org/10.1177/0146167208318602

Yzerbyt, V., Provost, V., & **Corneille, O.** (2005). Not competent but warm… really? Compensatory stereotypes in the French-speaking world. *Group Processes & Intergroup Relations*, *8*(3), 291–308. DOI: https://doi.org/10.1177/1368430205053944