

A Question of Compensation: The Social Life of the Fundamental Dimensions of Social Perception

Nicolas Kervyn

Université catholique de Louvain and Fonds de la recherche scientifique, Belgium

Vincent Y. Yzerbyt

Université catholique de Louvain, Louvain-la-Neuve

Charles M. Judd and Ana Nunes

University of Colorado at Boulder

This research examines the impact of the compensation effect between the fundamental dimensions of warmth and competence on behavioral confirmation. In Experiment 1, participants were presented with 2 groups that varied on 1 of the 2 dimensions and asked to select the questions that they wanted to pose to learn more about the groups. Participants preferred to ask negative (positive) questions about the unmanipulated dimension to the high (low) group. In Experiment 2, participants rated the 2 groups on the basis of naïve people answers to those questions. As predicted, compensation emerged. Experiment 3 involved interactions among 3 participants, 1 interviewing the other 2 using the questions selected in Experiment 1. Ratings of targets' reactions again showed compensation.

Keywords: stereotypes, compensation, warmth, competence, impression confirmation

“Business people are rich and smart, but they are arrogant and calculating.” “Women are warm and caring, but they can’t perform in demanding jobs.” “Southerners are welcoming and know how to party, but they are indolent and unorganized.” These are but a few illustrations of views that are commonly held about various groups. Interestingly enough, all of these examples oppose two fundamental dimensions (i.e., warmth and competence) that have been found to organize social perception (Fiske, Cuddy, & Glick, 2007). Women and Southerners are said to be high on warmth but low on competence, whereas business people are viewed as high on competence but low on warmth. Judd, James-Hawkins, Yzerbyt, and Kashima (2005) have provided experimental evidence for a compensation effect (i.e., a negative relationship) between warmth and competence in impression formation. A key remaining question and the focus of the present research efforts is whether this effect persists beyond impression formation and can affect the behavior of the perceiver toward the targets. Even more intriguing is whether compensatory expectations may also affect the behavior of the targets of perception, showing behavioral confirmation effects (Snyder, 1984). Building on work by Judd et al. (2005), the specific ambition of the present research is thus to show that beyond mere impression formation, the compensation effect influences social perception and behavior further downstream.

Warmth and competence have been shown to be fundamental dimensions in person perception as well as group perception (for recent contributions, see Abele & Wojciszke, 2007; Fiske et al., 2007; Judd et al., 2005; Wojciszke, 2005). The relative importance of those two dimensions in the perception of the social world has been examined in some research. For instance, Wojciszke (1994, 2005) found evidence for the primacy of the warmth relative to the competence dimension (see also Fiske et al., 2007). Our specific interest here resides in the relationship between these two dimensions. Early studies in impression formation provide some evidence concerning this relationship. For instance, the multidimensional scaling study of Rosenberg, Nelson, and Vivekananthan (1968) revealed the presence of a positive correlation between the social and the intellectual dimensions. This moderately positive correlation may be seen as one more manifestation of the well-known halo effect (e.g., Thorndike, 1920), that is, the tendency to “think of a person in general as rather good or rather inferior and to color the judgment of the separate qualities by this feeling” (p. 25). Drawing on earlier findings in the intergroup domain (Phalet & Poppe, 1997), Fiske, Cuddy, Glick, and colleagues (Cuddy et al., in press; Fiske, Cuddy, Glick, & Xu, 2002; Fisk, Xu, Cuddy, & Glick, 1999) examined how social groups are perceived along the two dimensions, grouping them into four quadrants created by crossing the two dimensions. In the context of their stereotype content model, Fiske and colleagues found that whereas some groups were perceived as high on both dimensions and others as low on both, a substantial number of groups were judged to be high on one but low on the other. Data such as these suggest that the two dimensions may in fact be negatively related in group stereotypes.

More recently, Yzerbyt, Provost, and Corneille (2005) explicitly proposed that warmth and competence would often manifest a negative relationship, a pattern that they referred to as “compensation.” To examine this conjecture, these authors relied on a fully crossed design in which members of two groups rated their in-

Nicolas Kervyn, Department of Social and Organizational Psychology, Université catholique de Louvain, Louvain, Belgium and Fonds de la Recherche Scientifique, Belgium; Vincent Y. Yzerbyt, Department of Social and Organizational Psychology, Université catholique de Louvain, Louvain-la-Neuve; Charles M. Judd and Ana Nunes, Department of Psychology, University of Colorado at Boulder.

Correspondence concerning this article should be addressed to Nicolas Kervyn, Department of Social and Organizational Psychology, Université catholique de Louvain, Place du Cardinal Mercier 10, B-1348 Louvain-la-Neuve, Belgium. E-mail: nicolas.o.kervyn@uclouvain.be

group as well as the outgroup. Specifically, Yzerbyt et al. (2005) asked French and Belgian respondents to indicate how they perceived their own and the other group in terms of competence and warmth. In line with the authors' compensation hypothesis, both groups of respondents described one of the two groups as more competent than warm and the other group as less competent than warm. Support for the compensation hypothesis has also been found with other social targets. For instance, Cuddy, Fiske, and Glick (2004) reported that participants saw working women as warmer and also less competent upon learning that they became mothers. Similarly, Cuddy, Norton, and Fiske (2005) showed that an elderly person was seen as warmer when also described as lower in competence (i.e., experienced memory losses).

Judd et al. (2005) were the first to experimentally test the compensatory nature of the relationship between warmth and competence in social perception. These authors used artificial groups in order to control all the information provided to participants. Participants learned about two groups by reading about behaviors that group members had engaged in. Whereas some saw two groups that differed in competence (one high and one low), others saw two groups differing in warmth (again one high and one low). Participants then judged the two groups both on the dimension that was manipulated and on the other, unmanipulated, dimension. Results showed clear compensation effects. On the manipulated dimension, naturally, the high group was judged higher than the low group. On the unmanipulated dimension, be it warmth or competence, this difference was reversed, the high group being judged lower than the low group.

Interestingly, Judd et al. (2005) reported empirical evidence that this compensation effect between competence and warmth applies not only when two groups are judged but also when the targets of judgment are two individuals. This apparent inconsistency with the halo effect (Rosenberg et al., 1968; Thorndike, 1920) was resolved, however, by a further study that Judd et al. (2005) conducted in which participants judged a single target, either described as high or low on the manipulated dimension. In this single-target condition, the two dimensions manifested a clear halo effect. The target that was high (low) on the manipulated dimension was judged as also high (low) on the unmanipulated one. Thus, it would appear that the negative compensatory relationship between the two dimensions emerges only when targets are in a comparative context. Finally, recent work by Yzerbyt, Kervyn, and Judd (2008) has shown that the compensation effect does not work with just any pair of dimensions but would seem to be unique to the two fundamental dimensions. That is, when one social target is perceived to be more competent (warm) than the other, it is also seen as less warm (competent) than the other and the other way around. And this seems to be the case uniquely for these two fundamental dimensions.

In light of the above empirical evidence, we define the *compensation effect* as the tendency to differentiate two social targets in a comparative context on the two fundamental dimensions by contrasting them in a compensatory direction. Specifically, when observers see one of two groups as being higher than the other on one dimension, as a result of compensation, they will also see it as lower than the other on the other fundamental dimension. Why would such a compensation effect emerge in social perception? The various boundary conditions identified by Judd et al. (2005) and by Yzerbyt et al. (2008) suggest that compensation stems not

only from a concern for distributive justice (Kay & Jost, 2003) but may also derive from the structural and functional relations between the two fundamental dimensions of social judgment. In other words, a justice motivation may indeed encourage social perceivers to even things out as far as their social world is concerned and to distribute a few positive characteristics and a few negative characteristics to all groups in a given comparative context. At the same time, if compensation emerges only or more easily when the fundamental dimensions of competence and warmth are involved, then one would question a strong version of the motivational interpretation that predicts compensation for any pair of valenced dimensions. Instead, a more sophisticated version of the justice account would seem to be appropriate, in which the specific constraints that render these two dimensions fundamental constitute a key determinant of compensation.

In the present research we sought to show that this compensation effect not only leads to biases in impression formation but also influences social perception and behavior further downstream. Specifically, we conducted three experiments in order to investigate the emergence of both hypothesis confirmation and behavioral confirmation (Snyder, 1984) processes ensuing from compensation. Hypothesis confirmation has been the focus of an impressive number of studies (for a review, see Kunda, 1999). In the social domain, research has provided abundant evidence that perceivers are particularly adept at verifying their prior views of others. Specifically, people not only selectively pay attention to information that proves consistent with their initial beliefs but also interpret new evidence in line with their preferred hypothesis (e.g., Darley & Gross, 1983). Finally, perceivers often manage to shape other people's behavior in such a way as to have them support their favored conclusions (Snyder, 1984).

In what has become a classic demonstration of behavioral confirmation and self-fulfilling prophecies, Snyder, Tanke, and Berscheid (1977) asked male participants to interact with female participants over the phone. The perceived physical attractiveness of the female participant was manipulated by showing the men different photographs of their alleged conversational partner. In actuality, the female participants remained blind to this manipulation and were randomly assigned to experimental conditions. Male participants were found to form more positive expectations when shown the attractive as opposed to the unattractive picture. The responses of the women were taped and evaluated by independent judges who were also blind to the picture manipulation. These judges rated the women as being more sociable in the attractive than in the unattractive picture condition. Clearly, male participants' expectations, based on the pictures they had seen, elicited behaviors from the female targets that were not only different in the two conditions but also in line with expectations.

For our purpose, it is important to note that these behavioral confirmation results are consistent with a halo effect. That is, more attractive targets were expected to be more sociable, and this expectation led to behavioral confirmation on the sociability dimension. To our knowledge, all prior work on behavioral confirmation (Snyder, 1984) in which confirmation from one dimension to another has been examined has found what we would call a "halo" confirmation. Another objective of the present work was to demonstrate an important new sort of behavioral confirmation, one that is consistent with a compensatory relationship between competence and warmth in comparative judgments. Specifically, we

intend to show that when a social group is praised compared with another group on one of the two fundamental dimensions, observers may behaviorally damn this same group on the second fundamental dimension.

We conducted three experiments. In Experiment 1, we presented participants with two groups, one high and the other low on one dimension and both ambiguous on the second dimension. Consistent with earlier work, we expected compensation on the unmanipulated dimension in judged impressions of the two groups. More importantly, we also asked participants to select questions that they might like to ask members of the two groups to get to know them further (Snyder & Swann, 1978). These questions were written to elicit further information about the targets that confirmed that they were either high or low on either the dimension of competence or the dimension of warmth. We expected that questions would be selected that were confirmatory on the manipulated dimension and compensatory on the unmanipulated one. In Experiment 2, we asked a new sample of participants to answer all the questions selected in Experiment 1. We then created subsets of answers corresponding to the questions selected most frequently in Experiment 1 for each one of the two groups and presented these answers to naïve participants. Our aim was to check whether differential impressions of the two groups would emerge simply on the basis of their answers to the most frequently selected questions, both confirmatory on the dimension that differentiated the groups for whom the questions were selected and compensatory on the unmanipulated dimension. Importantly, although these naïve participants thought they were reading answers from two different groups, these answers had actually been collected from the exact same set of naïve respondents. Any difference in the resulting group impressions is thus entirely due to the different questions asked of the same targets rather than to any target differences. We went one step further in Experiment 3 by bringing 3 participants to the laboratory at the same time. Of the participants, 1 was randomly assigned to be the interviewer, whereas the other 2 were told that they were members of one or the other group allegedly on the basis of the results of a personality questionnaire. The interviewer then asked different questions to the two interviewees, either those previously selected for the high- and low-competence groups or those previously selected for the high- and low-warmth groups. We expected that subsequent impressions of the two interviewees, based on their answers to the different questions, would be confirmatory on the dimension used for question selection but compensatory on the other dimension.

Experiment 1

In our Experiment 1, Belgian participants were shown behaviors allegedly performed by the members of two groups. One group was described with behaviors high on one of the two dimensions (warmth or competence), and the other group was described with behaviors relatively low on that same dimension. The information given about the other dimension was minimal and counterbalanced for the two groups. We measured the impressions participants formed about the two groups on the two dimensions and expected a compensation effect to emerge, thereby replicating Judd et al.'s (2005) compensation pattern with a Belgian rather than with an American participant population. But our major goal in this study was not simply a replication. Rather, we wanted to use this study

to begin the process of examining hypothesis confirmation processes of the compensation effect. Accordingly, building on a well-established procedure in the research on hypothesis confirmation (Dumont et al., 2003; Snyder & Swann, 1978), we asked participants to select from among a set of questions that they would find most useful in gaining further information about the described group members. Our hypothesis was that the questions selected by participants would reflect the compensation effect. Specifically, for the questions pertaining to the manipulated dimension, those implying the positive (negative) end of the dimension should be selected to be asked of the high (low) group. More important, we also expected participants to select questions that would manifest a compensatory pattern on the unmanipulated dimension. That is, the positive questions should be selected for the low group and the negative ones for the high group.

Method

Pretest

Behaviors that were relatively diagnostic on one dimension (either high or low) and nondiagnostic (or neutral) on the other needed to be created. Therefore, 79 behaviors pretested by Judd et al. (2005) were translated and presented to 22 students at the Catholic University of Louvain, approached in the University's libraries, reversing the order of presentation for half of them. For each behavior, participants answered two questions using 9-point scales ranging from 1 (*not at all*) to 9 (*absolutely*), namely, "to what extent would you say that a person displaying this behavior is competent, intelligent, dynamic, and organized" and "to what extent would you say that a person displaying this behavior is sociable, warm, friendly, and caring." The mean perceived competence and warmth for each behavior were computed, and behaviors that were significantly different from the midpoint of the scale on one dimension but not on the other were selected. This criterion yielded a total set of 17 negative competence, 15 positive competence, 8 negative warmth, and 12 positive warmth behaviors. Because only eight behaviors were needed for each valence on each dimension, those behaviors that best fit the criterion were selected. As was the case in Judd et al.'s (2005) data, the pretest revealed the presence of a positive correlation between the competence and warmth dimensions across all 79 behaviors ($r = .34$, $p < .01$). This correlation was not significant, although it was still positive, for the 32 behaviors retained in the final set ($r = .07$, $p > .70$). These positive correlations in the mean behavior ratings are consistent with the halo relationship found by Rosenberg et al. (1968).

Participants and Design

Fifty-three students at the Catholic University of Louvain were each paid 5 Euros (approximately \$7 U.S.) for taking part in the experiment. Due to temporary constraints in the subject pool, all participants were women. Each participant saw behaviors from two groups, one high on the manipulated dimension and one low. The manipulated dimension varied between participants, and group (high vs. low) varied within them. Two factors that simply counterbalanced order (high group or low group presented and measured first) and names given to the high and the low group

(blue and green) were also included. Because each group was also described by means of a few behaviors on the unmanipulated dimension, a third counterbalancing factor varied behavior sets. Specifically, the behaviors on the unmanipulated dimension attributed to the two groups were counterbalanced across participants so that on average, the same set of behaviors on the unmanipulated dimension was attributed to the high group for half the participants and to the low group for the other half. Participants were randomly assigned to the 16 different conditions (2 manipulated dimensions \times 3 counterbalancing factors). The counterbalancing factors were not included in the analyses. The two dependent variables were the judged warmth and competence of each of the two groups and the questions that participants selected to ask the members of the two groups.

Procedure

Upon entering the lab, participants were informed that they were about to take part in an impression formation task. They were going to learn about two groups, the blue group and the green group, by reading a deck of cards one by one. The cards were in a random order. On each card, participants read the name of one of the two groups followed by a behavior attributed to a member of this group. As they read, participants were asked to sort the cards into two piles, one for each group. Twelve behaviors for each group were presented. Of these, 8 behaviors were diagnostic of the manipulated dimension and were high for the high group and low for the low group. Additionally, for each group, participants were shown 4 behaviors diagnostic of the unmanipulated dimension, with 2 being high on that dimension and 2 low. As described previously, the set of behaviors used for the unmanipulated dimension was counterbalanced across participants so that any particular behavior diagnostic of the unmanipulated dimension was associated equally often with the high group and with the low group. Next, participants were asked to take one pile at a time and to read the behaviors a second time. When participants had finished reading all of the cards twice, the experimenter took them back, and participants were asked to write down their impression of each group so that someone who had not read the cards would know what the groups were like. Once done, participants were then asked to read what they had written in order to make sure that their text conveyed their impressions. The written impressions were not a dependent variable but were used to help consolidate the group impressions in participants' minds. Next, participants were given the trait-rating questionnaire followed by the question-selection questionnaire. Participants were then thanked and their questions answered. All participants received a written debriefing at the end of data collection.

Dependent Variables

Trait ratings. Participants rated each group on 20 traits, 5 for each one of the four poles constituted by crossing valence (positive vs. negative) and dimension (competence vs. warmth). The traits used are listed in Appendix A. The questions read: "On average, how much would you say that members of the green/blue group are" Participants rated each trait on a 9-point scale ranging from 1 (*not at all*) to 9 (*totally*).

Questions. Participants learned that they had to create two questionnaires, one for the blue group and the other for the green

group (for a similar procedure, see Dumont et al., 2003). The purpose of these questionnaires was ostensibly to gather further information about the two groups. Participants were told that group members would later be given interviews on the basis of the questions selected, and these would be presented to other participants who would then form impressions of the group via these interviews. Therefore, questions should be selected that would gather the most diagnostic information from the ostensible interviews. Participants were then shown a list of 40 questions that could be included in the interview and asked to indicate for each the extent to which it ought to be asked in the interview of each group on a 6-point scale ranging from 1 (*definitely ask of the blue group*) to 6 (*definitely ask of the green group*). There were 10 questions for each of the four combinations of dimension and valence. These questions were written on an a priori basis by the authors and are given in Appendix B. The 40 questions were presented to participants in a random fixed order.

Results

One participant failed to comply with the instructions and was dropped from the analyses, leaving a total of 52 participants. Twenty-seven participants were assigned to the competence-manipulated condition and 25 to the warmth-manipulated condition. Our primary goal in this study was to examine the questions that participants selected on the basis of their impressions. However, we first examined the trait ratings to verify compensation in the impressions of the groups that were formed.

Trait Ratings

For each group, we averaged the ratings of the 10 traits (reversing the ratings of the negative traits) on each dimension. Reliability was high for all four scales (Cronbach's α s > 0.85). We then analyzed these composite scores separately for the manipulated and unmanipulated dimension because effects on the first are manipulation checks, and those on the second constitute effects of primary interest. Each analysis involved a 2 (condition: manipulation of warmth vs. competence) \times 2 (group: high vs. low on the manipulated dimension) mixed model analysis of variance (ANOVA), with the first factor varying between participants, and the second varying within participants (see Table 1).

The analysis of the ratings on the manipulated dimension revealed the presence of a very strong group effect, $F(1, 50) = 594.29, p < .001$, confirming the success of our manipulation. As expected, participants' ratings were higher for the high group ($M = 8.03$) than for the low group ($M = 2.88$). The condition

Table 1
Mean Ratings of High and Low Groups on Manipulated and Unmanipulated Dimensions in Experiment 1

Group	Dimension			
	Manipulated		Unmanipulated	
	Low	High	Low	High
Competence	3.70	7.83	5.85	4.14
Warmth	1.99	8.24	6.63	5.54

effect was also significant, $F(1, 50) = 35.24, p < .001$. On average, both groups were rated more positively on the manipulated dimension when competence ($M = 5.76$) rather than warmth ($M = 5.12$) was manipulated. Finally, there was a significant Group \times Condition interaction, $F(1, 50) = 24.91, p < .001$. This interaction reflected the fact that the difference between the high and the low group was larger when warmth was manipulated (and judged), $t(24) = 23.33, p < .001$, than when competence was manipulated (and judged), $t(26) = 12.66, p < .001$.

Turning to the ratings on the unmanipulated dimensions, the group effect was significant, $F(1, 50) = 22.98, p < .001$. In line with our prediction that a compensation effect would emerge in the data, the group that had been presented as higher on the manipulated dimension was rated lower on the unmanipulated one ($M = 4.81$) than the group that had been presented as lower on the manipulated dimension ($M = 6.23$). We also found a significant condition effect, $F(1, 50) = 42.74, p < .001$. In general, groups were rated more positively on the unmanipulated dimension when warmth ($M = 6.09$) rather than competence ($M = 4.99$) was manipulated (i.e., when rating competence rather than warmth). The Group \times Condition interaction was not significant, $F(1, 50) = 1.14, p = .29$, meaning that the group effect did not differ as a function of which dimension was manipulated.

Questions

We examined the ratings given to the questions by computing four scores, averaging the ratings given to questions of the same valence and on the same dimension. Again, we analyzed the two scores given to questions on the manipulated dimensions and those given to questions on the unmanipulated dimension separately, conducting a 2 (condition: manipulation of warmth vs. competence) \times 2 (valence: positive vs. negative questions) mixed model ANOVA, with the first factor varying between participants, and the second varying within them (see Figure 1).

The analysis of the ratings given to questions that pertained to the manipulated dimension revealed the presence of a significant valence effect, $F(1, 50) = 15.52, p < .001$. Not surprisingly,

participants indicated that the positive questions should be directed to the high group ($M = 4.34$) more than the negative questions ($M = 3.25$). The Condition \times Valence interaction was not significant.

Turning to the ratings given to questions pertaining to the unmanipulated dimension, the valence effect was significant, $F(1, 50) = 5.36, p < .03$. Confirming a compensation effect in the ratings of the questions, participants indicated that the negative questions should be directed to the high group ($M = 3.48$) more than the positive questions ($M = 3.04$). There was no Condition \times Valence interaction, $F(1, 50) = 0.24, p = .63$, meaning that this valence effect did not differ as a function of which dimension was manipulated.

Because the questions were presented to participants with the instruction that this would allow us to create a questionnaire for each one of the two groups (and because subsequent experiments actually did use these questions for interviews), we sorted the 40 questions on the basis of their mean scores, computed across participants, separately for each dimension. From each of these two rankings, we took the 10 questions that had the lowest score and the 10 questions that had the highest score. This allowed us to create four virtual questionnaires with 10 questions each, one for the high-competence group, one for the low-competence group, one for the high-warmth group, and one for the low-warmth group. Interestingly, only a total of 29 different questions were selected (i.e., 11 questions ended up in the list in both the competence and the warmth manipulation). These four lists of questions were those that participants, on average, had indicated they would most like to ask each of the four groups. As can be seen in Appendix B, the distribution of the questions selected on this basis supported our compensation hypothesis. Except for the questions selected for the high-competence group, those for the other three groups consisted only of positive questions on the manipulated dimension and negative ones on the unmanipulated dimension for the high groups and of negative questions on the manipulated dimension and of positive ones on the unmanipulated for the low groups. Interestingly, participants tended to avoid choosing the negative questions, as only 12 of the 40 selected questions are negative ones.

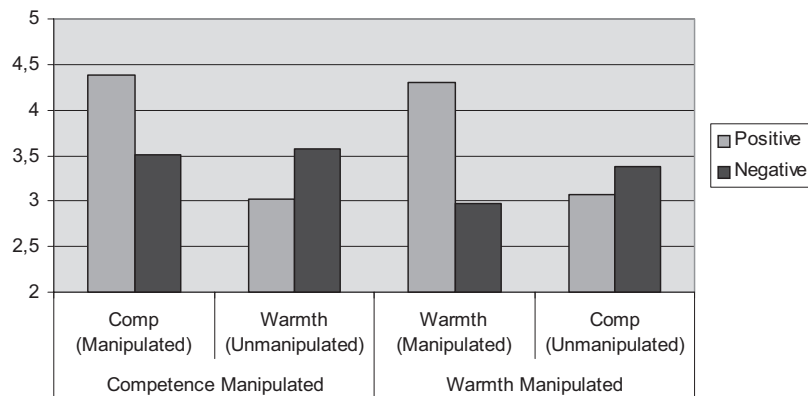


Figure 1. Choice of group for the questions as a function of the dimension manipulated and the questions' valence and dimension in Experiment 1. Comp = Competence.

Discussion

Our data lend strong support to our hypotheses. First, we replicated the compensation effect found by Judd et al. (2005) in participants' trait ratings of the groups they learned about. More importantly, we obtained a compensation effect whether competence or warmth was the manipulated dimension. This is important inasmuch as Judd et al. (2005) manipulated only one dimension at a time in any given experiment. Although all participants were women, it is doubtful that this feature influenced the present results. Previous research (Judd et al., 2005) consistently found compensation in the judgment of both male and female participants.

More importantly, participants' choice of questions for subsequent interviews confirmed and extended earlier findings. When asked to select questions to be used to gather more information about the two groups, participants selected questions that were likely to yield answers that would confirm the impressions that participants had formed of the groups. This confirmation showed a compensatory pattern on the unmanipulated dimension.

This represents an important first step in demonstrating compensation effects in hypothesis confirmation. Experiments 2 and 3 now used these chosen questions to demonstrate that they would be sufficient to generate compensatory impressions of targets simply on the basis of the responses of actual participants to these chosen questions.

Experiment 2

In Experiment 1, participants were told that they were to select questions to be used in an interview of members of each group to gather further information about them. In Experiment 2, we did just that. First, we collected answers from naïve respondents to all of the questions selected on the basis of Experiment 1. Then we selectively presented answers to these questions provided by the respondents to new participants. Specifically, some participants read answers to the 10 questions selected for the high-competence group and answers to the 10 questions selected for the low-competence group. Others read answers to questions selected for the high- and low-warmth groups. Our interest was whether participants would form different impressions of the target individuals who responded to the questions on competence and warmth as a function of the set of questions they read.

What makes this experiment particularly interesting is that the answers to all questions were in fact provided by the same respondents. Specifically, naïve respondents answered the full set of 29 different questions (that were the most frequently chosen for the four groups in Experiment 1), and their answers were later rearranged into four different interviews of 10 questions each. Any difference in the way participants rated the two groups would thus be due to the questions selected for those interviews and not to the respondents who answered them.

Method

Interview Construction

Forty-four undergraduate students from the Catholic University of Louvain were each paid 5 Euros (approximately \$7 U.S.) in exchange for their written answers to the full set of 29 questions

selected most frequently for the four groups in Experiment 1. They were recruited through posters placed around the University campus. The questions were presented in a randomly determined fixed order, and this order was reversed for half the respondents. The answers of 4 respondents had to be discarded either because they failed to answer all the questions or because several answers were clearly inappropriate. We randomly distributed the remaining 40 respondents into eight sets of 5 respondents. With each set of 5 respondents, four interviews of 10 questions with five different answers to each question was created. As in Experiment 1, the goal was to have participants form impressions of two groups, one high and one low on the manipulated dimension. Using the answers of the 40 respondents, 16 pairs of interviews, that is, 8 pairs of interviews for the manipulation of warmth and 8 for the manipulation of competence, were created. In each pair, there was one set of interview questions and answers for the high group on a given dimension (from 5 respondents) and another set for the low group on the same dimension (from the same 5 respondents). It is important to stress the fact that the manipulation resides in the questions that were selected to create the interviews and that this selection was based on the question selection task done in Experiment 1.

Participants and Design

Eighty female psychology students of the Catholic University of Louvain took part in the experiment in exchange for partial course credit. They were randomly assigned to 1 of 16 conditions (2 manipulated dimensions \times 8 pairs of interviews). The design was a 2 (condition: manipulation of warmth vs. competence) \times 2 (group: high vs. low on the manipulated dimension) mixed design, with the first factor varying between participants and the second varying within them. For half the participants, the high-group interview was presented first and that group was first in the impression writing and traits rating task; for the other half, the low group came first. The name of the group was also counterbalanced such that the high group was called the blue group and the low group the green group or the other way around.

Procedure

Upon participants' arrival at the lab, they were given the alleged interviews of members of the two groups. After reading the interviews twice, participants were asked to write their impressions and to fill in a series of trait ratings. These two tasks were the same as in Experiment 1.

Results

We averaged the ratings of the 10 traits (reversing the ratings of the negative traits) on each dimension. Reliability was high for all four scales (all Cronbach's α s $>$ 0.87). This experiment relied on a yoked design. For each manipulated dimension, each one of eight pairs of interviews was presented to 5 participants. Because of possible dependence in the data due to this yoking, we combined the five answers given to each pair of interviews and treated these as the units of analysis. Concretely, we averaged the ratings given by the 5 participants who read a given pair of interviews.

As in Experiment 1, we analyzed the scores on the manipulated and the unmanipulated dimension separately because effects on the first are manipulation checks, and those on the second constitute effects of primary interest. We analyzed scores by means of a 2 (condition: manipulation of warmth vs. competence) \times 2 (group: high vs. low on the manipulated dimension) mixed model ANOVA, with the first factor varying between and the second varying within interview pairs. For the manipulated dimension, the group effect was significant in the predicted direction, $F(1, 14) = 55.98, p < .001$ (see Table 2). The Group \times Condition interaction was also significant $F(1, 14) = 17.76, p < .001$, revealing that the group effect was larger when warmth was manipulated, $F(1, 7) = 81.61, p < .001$, than when competence was manipulated, $F(1, 7) = 4.59, p = .069$.

Turning to the unmanipulated dimension, the group effect was significant, $F(1, 14) = 48.98, p < .001$, confirming the predicted compensatory effect. Additionally, the condition effect and the Group \times Condition interaction were significant, $F(1, 14) = 12.39, p < .01$; and, $F(1, 14) = 6.32, p < .05$, respectively. Ratings were higher on average when competence rather than warmth was the unmanipulated dimension. Additionally, the compensatory effect on the unmanipulated dimension was larger when that dimension was warmth rather than when it was competence. Follow-up tests revealed, however, that the compensatory group effect was significant on the unmanipulated dimension regardless of whether it was competence, $F(1, 7) = 33.07, p < .001$, or warmth, $F(1, 7) = 26.68, p < .002$.

Discussion

In line with the compensation hypothesis, the low-competence group was perceived as warmer than the high-competence group, and the low-warmth group was seen as more competent than the high-warmth group. More important, this difference in impression results from the particular questions included in each interview rather than from any differences between respondents who were interviewed. Within a given pair of interviews, the respondents to both sets of 10 questions—the set pertaining to the high group and the set pertaining to the low group—were actually the same people. Thus, our participants were actually judging the same respondents (responding to different questions) when they thought they were judging the members of two different groups. Ours is thus a very powerful demonstration that the compensation effect observed in the question selection task from Experiment 1 does indeed have an impact beyond impression formation through confirmation biases. Even when forming an impression of *the exact*

same person, the questions that derive from prior compensatory impressions of group membership lead to compensatory effects on the unmanipulated dimension.

Clearly, Experiment 1 and Experiment 2 should in fact be looked at as two steps of one larger process. Experiment 1 showed that when participants read about two groups, one high and one low on one dimension, both neutral on the second dimension, the impressions that are formed show a compensation effect on the second dimension. When these same participants were given a chance to collect additional information, they selectively chose those questions that would likely confirm their initial impression. Experiment 2 established the presence of this confirmation bias as participants formed very different impressions of the two groups, depending on which set of questions, and thus which answers, were associated with these groups. This happened despite the fact that all questions had in fact been answered by the same naïve respondents.

Contrary to what was observed in Experiment 1, impressions of the groups collected in Experiment 2 differed more on warmth than on competence, regardless of whether it was the manipulated or the unmanipulated dimension (see Table 2). We interpret this finding as being due to the specific materials used for the presentation of the groups, that is, to the four sets of 10 questions used to create the interviews. Although the analysis of the question selection data in Experiment 1 revealed no difference for the two manipulations, there are differences between the number of questions selected on warmth and those selected on competence in the 4 \times 10 questions that were selected. For three of the four questions sets, there was a majority of warmth questions (see Appendix B). The question sets used were thus more effective at creating a clear and contrasting impression of the two groups on warmth than on competence. Another way to interpret the fact that participants saw a greater difference on warmth is that warmth is perhaps more initially important in impression formation (Wojciszke, Bazinska, & Jaworski, 1998) and possibly constrains compensation more than competence (Yzerbyt et al., in press). Therefore, it would make sense that participants rely more on this dimension to differentiate between the two groups. Still, compensatory impression formation was observed on the unmanipulated dimension, regardless of this difference.

Experiment 3

The results of Experiment 2 constitute strong evidence for the hypothesized confirmatory process of compensation between warmth and competence. Nevertheless, one potential limitation is that we relied on a rather sterile environment, devoid of actual social interactions in which participants could spontaneously interact and get to know each other. In Experiment 3, we sought to examine whether these same confirmation effects would occur in an actual interaction between individuals who get to know each other for the first time. Therefore, in Experiment 3, we developed an interview scenario in which 2 participants were led to believe they were members of either the blue or green group and were then interviewed by a third participant using the questions sets selected in Experiment 1.

Three previously unacquainted participants were brought into the laboratory. One of them was randomly assigned to be the interviewer; the other 2 were told they would be interviewed, using

Table 2
Mean Ratings of High and Low Groups on Manipulated and Unmanipulated Dimensions in Experiment 2

Group	Dimension			
	Manipulated		Unmanipulated	
	Low	High	Low	High
Competence	5.18	5.96	6.23	4.10
Warmth	3.88	6.69	6.17	5.16

questions that we had developed in order to examine whether responses to them were diagnostic of blue/green group membership. These 2 participants were then administered a bogus personality test, revealing that 1 of them was a member of the blue group and 1 was a member of the green group. After this, the interviewer conducted the interviews of the 2 target people, posing the predetermined questions to the 2 target individuals in a randomly determined order. At the end of the interview, all 3 participants gave impression ratings of each target group (i.e., green and blue) as well as of each person who had been interviewed. We expected to find target group differences on the manipulated dimension and, more importantly, compensatory target group differences on the second unmanipulated dimension. We predicted that these effects would also emerge in judgments of the interviewer and of the interviewees when rating the other interviewee. We were less confident in our predictions of how the interviewees would rate themselves.

Method

Participants and Design

One hundred twenty-nine female undergraduates from the University of Colorado participated in partial fulfillment of an introductory psychology course requirement. Because of their wider availability in the population, participation was limited to women in order to eliminate the influence of gender on the dynamics of the interview process. This should not pose a problem given that earlier findings (Judd et al., 2005; Yzerbyt et al., 2008) showed that men and women were equally susceptible to compensation effects. Moreover, Wood and Karten (1986) have shown that when one of the members of a group is given high status, group members react similarly irrespective of gender. In light of the intimate link between status and competence (Fiske et al., 1999), confidence that the gender of the participants would not affect the findings was established. The participants came to the laboratory in previously unacquainted triads. Each triad was randomly assigned to one of two conditions: Either the competence questions or warmth questions would be used to conduct the interview. Additionally, within triads, role was randomly assigned. First, the interviewer was randomly selected from the three. Next, following a bogus personality test given to the two interviewees, one of them was randomly told she was a member of the blue group, and the other was told she was a member of the green group. An additional randomization was done at this point, with either the blue or the green group member designated as the target receiving the "high" competence or warmth questions and the other receiving the "low" questions.

After the conclusion of the interview, each of the 3 participants gave their impressions of both groups as well as of both target individuals who had been interviewed, using trait rating scales. Notice that whereas the interviewer rated two groups and two target individuals who did not include the self, each interviewee rated her own group and herself in addition to the other interviewee and the other interviewee's group.

Procedure and Materials

Participants were told that a new and fundamental personality dimension had been discovered according to which people could

be categorized into one of two groups that would be referred to, for the purposes of this experiment, as the green group and blue group. They were also told that the experiment was intended to test interview questions that had been developed to be diagnostic of group (blue or green) membership. No other information regarding the two groups was provided.

The 3 participants were then assigned a role, as interviewer or as one of two interviewees, by picking from three slips of paper (one read *Interviewer*, and two read *Respondent*). The two interviewees were then given a "test" to determine whether each was a member of the blue or green group. This test was the same one that was used to manipulate group membership in earlier experiments (see Judd et al., 2005) and involved a variation on the well-known "dot estimation" task. The experimenter collected the answers, led the interviewees to believe she scored their "tests," and declared one to be a member of the blue group and the other a member of the green group. The interviewee's group membership was in fact randomly assigned.

A sheet of blue or green paper was placed in front of each interviewee to designate group membership. The interviewer was then given a set of 20 questions from Experiments 1 and 2 (10 questions selected for the high group in Experiment 1, and 10 questions selected for the low group in Experiment 1) to pose to the respondents. These were the same questions that were chosen for each group in Experiment 1 and that were used in Experiment 2. The "high" and "low" questions were listed in an alternating order, such that the first question was asked of one interviewee and then the next question was asked of the other interviewee. Given this alternating constraint, the order of the questions was randomly determined. Additionally, it was randomly determined whether the blue person was asked the "high" questions and the green the "low" ones or vice versa. Finally, the target (green vs. blue) given the first question was counterbalanced.

After the interview was complete, the 3 participants were asked to rate each group by answering 16 trait questions using a 9-point rating scale ranging from 1 (*not at all*) to 9 (*very much*). These traits included eight competence-related traits (four positive and four negative) and eight warmth-related traits (four positive and four negative); they are listed in Appendix A. Participants were then asked to complete the same rating scales a second time, this time rating the two individual interviewees rather than the groups to which they supposedly belonged. The participants were asked about suspicions and were then debriefed, thanked, and dismissed.

Results

In total, 43 triad sessions were conducted, with 22 being assigned to the competence condition and 21 to the warmth condition. For both the group and the individual target ratings, we formed composite scores on each dimension by averaging the ratings on the eight traits (reversing the negative ones first). The reliabilities of these composites were all high (all Cronbach's α s > .77).

We first conducted analyses of the ratings of the two groups (green and blue) as judged by the interviewer and then as judged by the interviewees. In the case of the interviewer's group ratings, these were analyzed by whether judgments were made of the group that was asked the "high" or "low" questions on the manipulated dimension and whether that manipulated dimension was compe-

tence or warmth. As in the previous experiments, separate analyses were conducted for ratings on the manipulated and unmanipulated dimensions. The expected difference between the high and low targets on the manipulated dimension constitutes simply a manipulation check. On the unmanipulated dimension, we expected to find the compensation effect. In the case of the interviewee's ratings of the two groups, we separately analyzed the ratings of the other interviewee's group and those of the rater's own group. And again, these analyses were done separately for the manipulated and unmanipulated dimensions, with group varying within (high vs. low) and condition (manipulation of competence vs. warmth) varying between triads. Once these group ratings were analyzed, we redid these analyses, focusing on the ratings of the 2 target individuals who had been interviewed.

Ratings of Groups

Judgments by interviewer. On the manipulated dimension, the target group (high vs. low) effect was significant, $F(1, 39) = 21.56, p < .0001$ (see top portion of Table 3 for means). Unsurprisingly, the high group was judged higher on the trait questions than the group assigned to the low questions. The condition (manipulation of competence vs. warmth) effect was not significant, $F(1, 39) = 2.29, p > .10$, nor was the Group \times Condition interaction, $F(1, 39) = 0.03, p > .50$. On the unmanipulated dimension, the target group main effect was significant, $F(1, 39) = 21.94, p < .0001$. Consistent with the predicted compensation effect, the group given the high questions on the manipulated dimension was rated lower on the unmanipulated dimension than the group given the low questions on the manipulated dimension. The condition effect was significant, $F(1, 39) = 6.08, p < .02$, such that the groups were rated more positively on the unmanipulated dimension when warmth rather than competence was manipulated, meaning that higher ratings on competence than warmth were given when this was the unmanipulated dimension. The Group \times Condition interaction was not significant, $F(1, 39) = 0.35, p > .50$.

Judgments of other group. We now examine how the two interviewees judged the group to which the other interviewee

belonged. On the manipulated dimension, the group effect was significant, $F(1, 39) = 22.57, p < .0001$. As the means in the middle portion of Table 3 show, the high group was unsurprisingly given higher trait ratings on the manipulated dimension than the low group. The condition main effect was not significant, $F(1, 39) = 1.10, p > .30$. The Group \times Condition interaction approached significance, $F(1, 39) = 4.07, p < .06$, indicating that a greater difference between low and high group ratings was found when competence rather than warmth was manipulated. On the unmanipulated dimension, the group effect was significant, $F(1, 39) = 20.69, p < .0001$, and as predicted, it was in a compensatory direction. The Group \times Condition interaction was also significant, $F(1, 39) = 10.76, p < .005$. There was a significant difference between the low and high group ratings when competence was the manipulated dimension, $t(1, 39) = 5.67, p < .0001$, but not when warmth was the manipulated dimension, $t(1, 39) = 0.88, ns$. The condition effect was not significant, $F(1, 39) = 0.11, ns$.

Own group judgments. Each interviewee rated the group to which she belonged (for means, see the bottom portion of Table 3). On the manipulated dimension, the group (high vs. low) effect was significant, $F(1, 39) = 8.87, p < .005$, such that when one's own group was the high group, it was rated higher than the low group. The condition effect was not significant, $F(1, 39) = 0.02, ns$, nor was the Group \times Condition interaction, $F(1, 39) = 0.33, ns$. On the unmanipulated dimension, the group effect was significant, $F(1, 39) = 14.03, p < .001$, in the predicted compensatory direction: One's own group was rated as lower on the unmanipulated dimension when one had responded to the high questions on the manipulated dimension. Given that interviewees are here making judgments about a group to which they supposedly belong, this compensatory effect is particularly interesting. The Group \times Condition interaction was also significant, $F(1, 39) = 5.45, p < .03$. Again, the difference between low and high group ratings on the unmanipulated dimension was significant when competence was manipulated, $t(1, 39) = 4.41, p < .0001$, but not when warmth was manipulated, $t(1, 39) = 0.97, ns$. The condition effect was not significant.

Comparing other and own group judgments. To examine whether different effects were obtained when the interviewees rated their own versus the other group, we conducted additional analyses, including both sets of ratings with own/other as an additional within-group factor. For ratings on the manipulated dimension, the group effect was highly significant, $F(1, 39) = 24.16, p < .0001$, and this was not moderated by whether ratings were of one's own group or the other interviewee's group, $F(1, 39) = 0.41, ns$. For ratings on the unmanipulated dimension, again the group effect, in a compensatory direction, was highly significant, $F(1, 39) = 28.28, p < .0001$. Consistent with the results already reported, this compensatory effect was significantly stronger when competence rather than warmth was manipulated, $F(1, 39) = 13.16, p < .01$. More important, however, the magnitude of the compensatory effect did not depend on whether interviewees were rating their own group or the other interviewee's group, $F(1, 39) = 1.79, ns$.

Rating of Individuals

We then analyzed the ratings of the individuals who had been interviewed. Again, the interviewer made judgments about the two

Table 3
Mean Interviewer Ratings of High and Low Groups on Manipulated and Unmanipulated Dimensions in Experiment 3

Manipulated dimension	Dimension			
	Manipulated		Unmanipulated	
	Low	High	Low	High
	Judgments by interviewer			
Competence	5.98	7.07	6.97	5.82
Warmth	6.42	7.43	7.44	6.54
	Judgments of other group			
Competence	5.99	7.32	7.44	5.68
Warmth	6.63	7.17	6.78	6.50
	Own group judgments			
Competence	6.12	7.01	7.27	6.16
Warmth	6.29	6.40	6.87	6.61

interviewees. Interviewees, of course, rated both themselves and the other person.

Judgments by interviewer. On the manipulated dimension, we found a significant difference in the ratings of the 2 individuals (high or low) in the expected direction, $F(1, 39) = 13.07, p < .001$ (see means in the top portion of Table 4). The condition (manipulation of competence vs. warmth) effect was not significant, $F(1, 39) = 1.20, ns$, nor was the Individual \times Condition interaction, $F(1, 39) = 0.71, ns$. On the unmanipulated dimension, the ratings of the 2 individuals were also significantly different in a compensatory direction, $F(1, 39) = 11.93, p < .002$. Again, the predicted compensation effect was observed such that individuals who were asked the high questions on the manipulated dimension were rated lower by the interviewer than individuals who were asked the low questions on the manipulated dimension. The condition effect was not significant, $F(1, 39) = 1.74, ns$, nor was the Individual \times Condition interaction, $F(1, 39) = 1.04, ns$.

Judgments of other interviewee. On the manipulated dimension, ratings of the 2 individuals (high vs. low) differed significantly, $F(1, 39) = 5.06, p < .05$. Interviewees who were asked high questions received higher trait ratings from the other interviewee than those who were asked low questions (see means in the middle portion of Table 4). The condition effect was not significant, $F(1, 39) = 0.12, ns$, nor was the Individual \times Condition interaction, $F(1, 39) = 0.58, ns$. On the unmanipulated dimension, the ratings of the 2 individuals were not significantly different, $F(1, 39) = 2.37, ns$, although the means were in the predicted (compensatory) direction. The Individual \times Condition interaction was significant, $F(1, 39) = 9.53, p < .004$. Whereas we get a clear compensatory effect when competence was manipulated, $t(1, 39) = 3.32, p < .005$, no compensation emerged when warmth was manipulated, $t(1, 39) = 1.08, ns$. The dimension main effect was not significant, $F(1, 39) = 0.16, ns$.

Judgments of self. Perhaps unsurprisingly, the ratings of self on the manipulated, $F(1, 39) = 0.06, ns$, and unmanipulated, $F(1, 39) = 0.83, ns$, dimensions did not differ as a function of high versus low questions, manipulated dimension, or their interaction: manipulated, $F(1, 39) = 0.04, ns$; unmanipulated, $F(1, 39) = 0.21,$

ns. Clearly, self-judgments are affected by extensive knowledge of oneself that renders inconsequential the incremental effects of our interview manipulation. Nevertheless, it is interesting that the direction of the mean differences on the unmanipulated dimension (see the bottom portion of Table 4) is compensatory, consistent with all of the results presented to this point.

Comparing other interviewee and self-ratings. Again, analyses were conducted that included each interviewee's ratings of self and the other in order to examine whether effects significantly differed as a function of whether one was rating oneself or the other interviewee. On the manipulated dimension, the effect of target person (ratings of the target who was asked the high vs. low questions, regardless of whether that target was oneself or the other interviewee) approached significance, $F(1, 39) = 3.18, p = .083$, and this difference did not depend on whether one was rating oneself or the other interviewee, $F(1, 39) = 0.46, ns$. On the unmanipulated dimension, on average across self and other ratings, there was a difference in ratings for the target who was asked the high questions compared with the target who was asked the low questions that approached significance, $F(1, 39) = 3.22, p = .081$, in a compensatory direction. Although this difference did not depend significantly on whether one was rating oneself or the other interviewee, $F(1, 39) = 0.31, ns$, it was significantly stronger when competence was manipulated rather than warmth, $F(1, 39) = 4.31, p < .05$. Additionally, the triple interaction of individual, condition, and whether one was rating oneself or the other interviewee was significant, $F(1, 39) = 6.72, p < .05$. As the means in Table 4 indicate, compensation emerged most forcefully in ratings of the other interviewee when competence had been manipulated.

Discussion

As predicted, a compensatory pattern was observed on the unmanipulated dimension for many of the ratings in Experiment 3. All of the judgments that were made by the interviewer manifested this predicted pattern regardless of which dimension was manipulated and regardless of whether the interviewer was rating the two groups or the two interviewees.

When looking at the ratings completed by the 2 interviewees, compensation clearly emerged in the ratings of the two groups to which they belonged, although this difference depended on which dimension was manipulated. It was always significant when competence was manipulated (and warmth rated), but was significantly less strong when warmth was manipulated (and competence rated). This difference in the strength of the compensation effect as a function of which dimension was manipulated is consistent with prior work (Judd et al., 2005; Yzerbyt et al., in press). More important, the strength of the compensation effect did not depend on whether the interviewees were rating the group to which they belonged or the group to which the other interviewee belonged. In terms of the interviewees' ratings of themselves and the other person, compensation was clearly weaker and emerged as significant only when the other person was rated and competence was manipulated (and warmth rated). Although compensation was not significant in ratings of oneself, it was also not significantly weaker in ratings of oneself than in ratings of the other interviewee.

The fact that compensation was found in ratings of groups to which the raters themselves belonged, at least when competence

Table 4
Mean Interviewer Ratings of High and Low Interviewees on Manipulated and Unmanipulated Dimensions in Experiment 3

Manipulated dimension	Dimension			
	Manipulated		Unmanipulated	
	Low	High	Low	High
Judgments by interviewer				
Competence	6.18	7.28	7.15	6.20
Warmth	6.73	7.41	7.31	5.78
Judgments of other interviewee				
Competence	6.64	7.36	7.62	6.56
Warmth	6.89	7.22	7.18	6.82
Judgments of self				
Competence	7.12	7.21	7.43	7.35
Warmth	7.23	7.24	7.34	7.07

was manipulated, complements earlier work in which similar effects have been reported (see Judd et al., 2005, Experiment 4). Because of the minimal group paradigm that was used in this and related studies, it is true that participants did not have a lot of information about their own group and hence compensation might be larger in this case than if we were dealing with real groups. Nevertheless, people do have extensive knowledge about themselves as individuals and should base their impressions of their own group, at least in part, on this knowledge (Cadinu & Rothbart, 1996; Otten & Wentura, 2001). Even with this information base, however, compensation emerged.

In summary, the data from this third experiment provide consistent support for the hypothesis that the questions that were chosen by respondents in Experiment 1 create a sort of confirmation spiral, eliciting responses from target individuals that evoke confirmatory impressions on the dimension on which they are most informative and compensatory effects on the other fundamental dimension of social judgment. What makes this demonstration particularly compelling is that the impressions formed in this third experiment were based on actual live interactions between previously unacquainted individuals rather than on responses to target individuals who were the same across all respondents. Admittedly, the effects we have documented reside in the questions that were asked. But these questions elicited particular responses, and in combination, these gave rise to compensatory impressions in real social interactions.

General Discussion

The goal of the present set of experiments was to test whether the compensation effect that has been shown in impression formation (Judd et al., 2005) also affects other important social perception processes. In the question selection task of Experiment 1, our results clearly support a hypothesis-confirmation strategy, in which questions are selected to confirm impressions in a compensatory manner. And in Experiments 2 and 3, those biases in question selection actually elicited responses that gave rise to compensatory impressions of targets who were known only by their responses. This set of experiments offers a comprehensive illustration of one of the ways in which compensatory impressions of individuals on competence and warmth and mixed stereotypes of groups on these dimensions are built and maintained.

An additional and intriguing feature of the present set of experiments is that they are the first to demonstrate that the confirmation of prior impressions may actually diverge from the classical halo effect. To this point, studies that have examined hypothesis confirmation processes and self-fulfilling prophecies have always been conducted in such a way that the positive standing of a target on one dimension (e.g., attractiveness) triggered a positive impression among perceivers and indeed confirmatory behaviors from targets on a second dimension (e.g., sociability). In the present case, this virtuous circle is definitely not observed. In contrast, the target who happened to be more positive on one dimension (e.g., competence) ended up confirming a less positive impression on a second dimension (e.g., warmth). We think that the present demonstration of a compensatory confirmation of hypotheses is not only unique but also theoretically provocative of the dynamics involved in impression formation.

It may also be noted that the results in Experiments 2 and 3 showed stronger compensation effects in the competence manipulation, in which warmth is the unmanipulated dimension, than in the warmth manipulation, in which competence is the unmanipulated dimension. To be sure, this difference between the two manipulations might be due to the specific materials we used. As a matter of fact, three of the four questions sets used in Experiment 2 and 3 had majority of warmth questions (see Appendix B). Still, to the extent that this pattern is reminiscent of stronger results found in the competence manipulation in Judd et al. (2005) and Yzerbyt et al. (2008), we believe that there may also be an underlying theoretical reason for this difference. Our interpretation of this pattern is that our participants are likely to be more motivated to differentiate the two targets on the warmth dimension simply because it is the primary dimension of social perception (Wojciszke, 1994, 2005; Wojciszke et al., 1998; Ybarra, Chan, & Park, 2001). Clearly, this difference between the two fundamental dimensions in terms of their primacy and its consequences on the way they relate to each other deserves further research.

In line with previous research (Cuddy et al., 2004, 2005; Judd et al., 2005; Yzerbyt et al., 2005), the present findings emphasize the special relation between competence and warmth (see also Kay et al., 2007). Whenever several social targets are compared, perceivers seem to embrace some sort of compensatory perception. Also consistent with Judd et al.'s (2005) findings is the fact that in Experiment 3, the compensation effect is observed on the ratings of the groups and also of the individuals. This is a further indication that although person and group perception have long been studied separately, the fundamental dimensions of warmth and competence and the way they interact are likely to be common to both kinds of targets.

The processes that we have illustrated offer interesting new ways to explain how mixed stereotypes are built and maintained. Abundant research on real social categories (Fiske et al., 1999; Pennebaker, Rime, & Blankenship, 1996; Phalet & Poppe, 1997; Yzerbyt et al., 2005) has identified pairs of social groups in which one is seen as high on competence (or some other competence-related characteristic), whereas the other is reputed to excel on warmth (or some other warmth-related trait). Perhaps the best-known illustration of this peculiar compensatory relationship can be found in the various personality traits commonly attributed to women and men (Abele, 2003; Eagly, 1987; Glick & Fiske, 1996). According to Eagly (1987; Eagly & Steffen, 1984) and her colleagues, the perception of men as being agentic and less communal and of women as being communal and less agentic stems from the repeated observation of these two social categories in different settings and social roles. Because these different settings and roles mandate specific behaviors on the part of the people occupying them and because the distribution of women and men in these roles is unbalanced, observers come to associate certain characteristics with one group as opposed to another (see also Hoffman & Hurst, 1990). This association of positive and negative aspects within gender stereotypes can be linked to the work undertaken to develop the benevolent and hostile sexism scales (Glick & Fiske, 1996; Glick, Fiske, & Mladinic, 2000). According to Glick and colleagues, benevolent sexism blame communal women for their lack of agency, whereas hostile sexism blame agentic women for their lack of communality. We consider this distinction between benevolent and hostile sexism to be a very nice illustration of the

perverse effect that the compensation effect can have on intergroup relations.

Along similar lines, system justification theorists (Jost & Banaji, 1994; Jost, Pelham, Sheldon, & Sullivan, 2003; Kay et al. 2007) have argued that a compensatory view of groups is likely to promote harmonious relations among their members. In a series of experiments designed to test this idea, Kay and Jost (2003) showed that when exposed to the complementary stereotypes of “poor but happy/honest” and “rich but miserable/dishonest”, participants scored higher on a measure of system justification than when they were exposed to noncomplementary exemplars. We agree with these authors that people expect and are more comfortable with groups that manifest a compensatory relationship (see also Mummeny & Schreiber, 1983).

In yet another field of research, substantial work has been done in an attempt to identify the antecedents of national stereotypes (Linszen & Hagendoorn, 1994; Phalet & Poppe, 1997; Poppe & Linszen, 1999). Factors such as a north-south axis, economic power, conflicts, and geographical size have been identified as predicting either the warmth or the competence that is attributed to a country (see also Pennebaker et al., 1996). What our research on the compensation effect claims is that the perceived warmth of a country is likely to influence its perceived competence and that the perceived competence of a country is likely to influence its perceived warmth. Taking the comparative nature of compensation into account, a more accurate description of the phenomenon is that warmth and competence are related in such a way as to compensate each other.

It is interesting to speculate about the potential consequences of this compensation mechanism in the context of intergroup contact. A long-standing assumption of intergroup contact research is that contact will bring about an overall improvement in the perception of the outgroup in the eyes of the ingroup. The present research suggests that people who get to know another group are not very likely to acknowledge the strengths of the outgroup on both competence and warmth. It is more likely the case that positivity or even excellence in one dimension goes hand in hand with the affirmation of the ingroup's superiority on the other dimension. Clearly, this view is highly compatible with the dual-category model (Hewstone & Brown, 1986).

Although we would argue that compensation influences a wide variety of aspects in social life, still very little is known about its exact determinants, its boundary conditions, and all of its consequences. Several factors that may lead people to compensate more or less remain to be identified, but those that have been evidenced in our own research as well as in other people's work, such as the presence of a comparative context (Judd et al., 2005) or the nature of the dimensions at stake (Yzerbyt et al., 2008), suggest that compensation results from a concern for distributive justice, but one that takes into account the structural and functional constraints inherent in the fundamental dimensions of social perception.

The present findings emphasize that the compensatory relationship between warmth and competence has a clear impact on the sort of evidence that perceivers examine and test in their social interactions. And these compensatory expectations actually give rise to their own fulfillment. Our work suggests that someone who interacts with some people whom they believe to be quite competent and others whom they believe to be com-

paratively less competent, for instance, will actually elicit responses and behavior from those individuals that confirm that the former are rather cold and unfriendly, whereas the latter are comparatively more communal. This is indeed a strange social dynamic, leading to rather ambivalent impressions of others, be they individuals or groups, and certainly affecting the pattern of subsequent interactions.

References

- Abele, A. E. (2003). The dynamics of masculine-agentic and feminine-communal traits: Findings from a prospective study. *Journal of Personality and Social Psychology, 85*, 768–776.
- Abele, A. E., & Wojciszke, B. (2007). Agency and communion from the perspective of self versus others. *Journal of Personality and Social Psychology, 93*, 751–763.
- Cadinu, M. R., & Rothbart, M. (1996). Self-anchoring and differentiation processes in the minimal group setting. *Journal of Personality and Social Psychology, 70*, 661–677.
- Cuddy, A. J. C., Fiske, S. T., & Glick, P. (2004). When professionals become mothers, warmth doesn't cut the ice. *Journal of Social Issues, 60*, 701–718.
- Cuddy, A. J. C., Fiske, S. T., Kwan, V. S. Y., Glick, P., Demoulin, S., Leyens, J.-P., & Bond, M. H., et al. (in press). Stereotype content model across cultures: Towards universal similarities and some differences. *British Journal of Social Psychology*.
- Cuddy, A. J. C., Norton, M., & Fiske, S. T. (2005). This old stereotype: The stubbornness and pervasiveness of the elderly stereotype. *Journal of Social Issues, 61*, 267–285.
- Darley, J. M., & Gross, P. H. (1983). A hypothesis-confirming bias in labeling effects. *Journal of Personality and Social Psychology, 44*, 20–33.
- Dumont, M., Yzerbyt, V., Snyder, M., Mathieu, B., Comblain, C., & Scaillet, N. (2003). Suppression and hypothesis testing: Does suppressing stereotypes during interactions help to avoid confirmation biases? *European Journal of Social Psychology, 33*, 659–677.
- Eagly, A. H. (1987). *Sex differences in social behavior: A social-role interpretation*. Hillsdale, NJ: Erlbaum.
- Eagly, A. H., & Steffen, V. J. (1984). Gender stereotypes stem from the distribution of women and men into social roles. *Journal of Personality and Social Psychology, 46*, 735–754.
- Fiske, S. T., Cuddy, A. J. C., & Glick, P. (2007). Universal dimensions of social perception: Warmth, then competence. *Trends in Cognitive Sciences, 11*, 77–83.
- 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 the perceived status and competition. *Journal of Personality and Social Psychology, 82*, 878–902.
- 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*, 473–491.
- Glick, P., & Fiske, S. T. (1996). The Ambivalent Sexism Inventory: Differentiating hostile and benevolent sexism. *Journal of Personality and Social Psychology, 70*, 491–512.
- Glick, P., Fiske, S. T., & Mladinic, A. (2000). Beyond prejudice as simple antipathy: Hostile and benevolent sexism across cultures. *Journal of Personality and Social Psychology, 79*, 763–775.
- Hewstone, M., & Brown, R. (1986). Contact is not enough: An intergroup perspective on the “contact hypothesis.” In M. Hewstone & R. Brown (Eds.), *Contact and conflict in intergroup encounters* (pp. 1–44). Oxford, England: Blackwell Publishers.
- Hoffman, C., & Hurst, N. (1990). Gender stereotypes: Perception or

- rationalization? *Journal of Personality and Social Psychology*, 58, 197–208.
- Jost, J. T., & Banaji, M. (1994). The role of stereotyping in system-justification and the production of false consciousness. *British Journal of Social Psychology*, 33, 1–27.
- Jost, J. T., Pelham, B. W., Sheldon, O., & Sullivan, B. N. (2003). Social inequality and the reduction of ideological dissonance on behalf of the system: Evidence of enhanced system justification among the disadvantaged. *European Journal of Social Psychology*, 33, 13–36.
- 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.
- Kay, A., & Jost, J. (2003). Complementary justice: Effects of “poor but happy” and “poor but honest” stereotype exemplars on system justification and implicit activation of the justice motive. *Journal of Personality and Social Psychology*, 85, 823–837.
- Kay, A., Jost, J., Mandisodza, A., Sherman, S., Petrocelli, J., & Johnson, A. (2007). Panglossian ideology in the service of system justification: How complementary stereotypes help us to rationalize inequality. In M. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 38, pp. 305–358). Thousand Oaks, CA: Academic Press.
- Kunda, Z. (1999). *Social cognition: Making sense of people*. Cambridge, MA: MIT Press.
- Linssen, H., & Hagendoorn, L. (1994). Social and geographical factors in the explanation of the content of European nationality stereotypes. *British Journal of Social Psychology*, 33, 165–182.
- Mummendey, A., & Schreiber, H. J. (1983). Better or different? Positive social identity by discrimination against or differentiation from out-groups. *European Journal of Social Psychology*, 13, 389–397.
- Otten, S., & Wentura, D. (2001). Self-anchoring and in-group favoritism: An individual profile analysis. *Journal of Experimental Social Psychology*, 37, 525–532.
- Pennebaker, J. W., Rime, B., & Blankenship, V. E. (1996). Stereotypes of emotional expressiveness of Northerners and Southerners: A cross-cultural test of Montesquieu’s hypotheses. *Journal of Personality and Social Psychology*, 70, 372–380.
- Phalet, K., & Poppe, E. (1997). Competence and morality dimensions in national and ethnic stereotypes: A study in six eastern-European countries. *European Journal of Social Psychology*, 27, 703–723.
- Poppe, E., & Linssen, H. (1999). In-group favouritism and the reflection of realistic dimensions of difference between national states in Central and Eastern European nationality stereotypes. *British Journal of Social Psychology*, 38, 85–102.
- Rosenberg, S., Nelson, C., & Vivekananthan, P. S. (1968). A multidimensional approach to the structure of personality impressions. *Journal of Personality and Social Psychology*, 9, 283–294.
- Snyder, M. (1984). When belief creates reality. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 18, pp. 247–305). Thousand Oaks, CA: Academic Press.
- Snyder, M., & Swann, W. (1978). Hypothesis-testing processes in social interaction. *Journal of Personality and Social Psychology*, 36, 1202–1212.
- Snyder, M., Tanke, E., & Berscheid, E. (1977). Social perception and interpersonal behavior: On the self-fulfilling nature of social stereotypes. *Journal of Personality and Social Psychology*, 35, 656–666.
- Thorndike, E. L. (1920). A constant error on psychological rating. *Journal of Applied Psychology*, 4, 25–29.
- Wojciszke, B. (1994). Multiple meanings of behavior: Construing actions in terms of competence or morality. *Journal of Personality and Social Psychology*, 67, 222–232.
- Wojciszke, B. (2005). Morality and competence in person- and self-perception. *European Review of Social Psychology*, 16, 155–188.
- Wojciszke, B., Bazinska, R., & Jaworski, M. (1998). On the dominance of moral categories in impression formation. *Personality and Social Psychology Bulletin*, 24, 1251–1263.
- Wood, W., & Kerten, S. J. (1986). Sex differences in interaction style as a product of perceived sex differences in competence. *Journal of Personality and Social Psychology*, 50, 341–347.
- Ybarra, O., Chan, E., & Park, D. (2001). Young and old adults’ concerns about morality and competence. *Motivation and Emotion*, 25, 85–100.
- Yzerbyt, V. Y., 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, 1110–1123.
- Yzerbyt, V. Y., 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.

Appendix A

Personality Traits

Valence	Dimension			
	Warmth		Competence	
	Negative	Positive	Negative	Positive
Traits used in Experiment 1 and 2	Cold Derogating Irritable Selfish Unpleasant	Caring Nice Sociable Tolerant Warm	Disorganized Lazy Messy Negligent Unintelligent	Capable Competent Conscientious Motivated Skilled
Traits used in Experiment 3	Cold Hostile Insensitive Unfriendly	Caring Popular Sociable Warm	Disorganized Incompetent Lazy Unreliable	Capable Competent Determined Skilled

Appendix B

Questions Rated in Experiment 1 and Used in Experiments 2 and 3

Experiment 1	Experiments 2 & 3			
	Competence manipulation		Warmth manipulation	
	High group	Low group	High group	Low group
High-competence questions				
<i>When you have to compete with others for something, what do you do to motivate yourself to do as well as you can?</i>	X			X
<i>When you are in class, do you find it better to take really detailed notes or do you prefer to get the global picture and pick up the details from readings?</i>	X			X
<i>When you really want to do well on a test you are going to take, do you find it better to study at home or in the library? Why?</i>	X			X
<i>When you read the newspaper, do you tend to spend more time reading international news or more national and local news? What makes one or the other more interesting to you?</i>	X			X
<i>What do you do to organize your planning and time in preparation for exams at the end of the semester?</i>	X			
<i>To get informed about what's going on in the world, do you go to newspapers, radio? TV? The Web? And why do you prefer some of these to others?</i>	X			
<i>In what areas would you say you are particularly strong or even gifted? Can you give some examples?</i>	X			
<i>Please describe one of your important achievements that you got through hard work and determination.</i>	X			
<i>What kind of novels do you prefer and why?</i>				
<i>When in a museum, do you like to have a guide or do you prefer to go at your own pace?</i>				
Low-competence questions				
<i>Everyone has strong and weak points. In thinking about yourself, what would you characterize as your weak point?</i>	X			
<i>In what parts of your life do you feel you are the least organized or the least on top of things?</i>	X			
<i>When you decide to cut class or skip a lecture, what kinds of things are you likely to do instead?</i>		X		
<i>Can you remember a time when your naiveté or your ignorance caused you problems? What happened?</i>		X	X	
<i>What do you do when you have forgotten to complete a class assignment or your part of a group assignment?</i>		X		
<i>How do you feel when you come out of an exam that you feel you have failed?</i>				
<i>Everybody knows people that are way better than them in some areas. Please describe the areas in which you wish you were more competent.</i>				
<i>In your studies, what subjects are the most problematic and why?</i>				
<i>What kind of show do you enjoy watching when you feel like relaxing?</i>				
<i>All of us have doubts about our abilities sometimes. What makes you doubt your own abilities?</i>				
High-warmth questions				
<i>What kinds of things do you enjoy doing when you spend an evening with your friends?</i>		X	X	
<i>When you have friends from out of the area come and visit you, where would you take them and what would you show them?</i>		X	X	
<i>Where do you like to go when you go out at night? What do you like about these places?</i>		X	X	
<i>Can you tell us about a recent instance or situation in which you helped someone out?</i>		X		
<i>What kinds of things are you likely to do to cheer up a friend who is depressed or having personal problems?</i>		X	X	
<i>When you meet someone new, what kinds of things are you likely to ask them to get to know them better?</i>		X	X	

(Appendix continues)

Appendix (continued)

Experiment 1	Experiments 2 & 3			
	Competence manipulation		Warmth manipulation	
	High group	Low group	High group	Low group
<i>What do you do when you haven't made plans for the evening, but you still want to do something with your friends?</i>		X	X	
<i>What types of gifts do you like to give to your friends on their birthdays?</i>			X	
<i>What do you do to stay in touch with friends who live far away?</i>			X	
<i>Can you tell me about the last donation you made or the last petition you signed? What was it for specifically? What made you donate or sign?</i>			X	
Low-warmth questions				
<i>What are the kinds of things that people do around you that make you angry?</i>				X
<i>When you've been invited to a party that you really don't want to go to, what kinds of excuses might you use?</i>				X
<i>When you eat by yourself, do you prefer to eat while watching TV or reading something? What do you prefer to watch or read while eating?</i>				X
<i>When someone really annoys you, what kinds of things might you think about that you would like to do to him or her?</i>				X
<i>Try to remember the last time you acted unfairly toward a friend or a fellow student. What made you act that way?</i>				X
<i>There are some parties or family events where you may go to because you feel an obligation to go rather than a true desire to go. Can you describe the last time you went to a party or family event due to a feeling of obligation?</i>				X
<i>When you are in a bad mood, do you tend to be sad and withdrawn or cold and insensitive?</i>				
<i>When you feel like spending some time alone, how do you manage to get away from everyone?</i>				
<i>What kind of situations make you feel jealous?</i>				
<i>Last time you had a fight with someone, what was it about?</i>				

Received August 23, 2007
Revision received June 20, 2008
Accepted June 20, 2008 ■