

# Fundamental Dimensions of Social Judgment: Understanding the Relations Between Judgments of Competence and Warmth

Charles M. Judd and Laurie James-Hawkins  
University of Colorado at Boulder

Vincent Yzerbyt  
Catholic University of Louvain at Louvain-la-Neuve

Yoshihisa Kashima  
University of Melbourne

In seems there are two dimensions that underlie most judgments of traits, people, groups, and cultures. Although the definitions vary, the first makes reference to attributes such as competence, agency, and individualism, and the second to warmth, communality, and collectivism. But the relationship between the two dimensions seems unclear. In trait and person judgment, they are often positively related; in group and cultural stereotypes, they are often negatively related. The authors report 4 studies that examine the dynamic relationship between these two dimensions, experimentally manipulating the location of a target of judgment on one and examining the consequences for the other. In general, the authors' data suggest a negative dynamic relationship between the two, moderated by factors the impact of which they explore.

*Keywords:* social judgment, stereotypes, person perception

There is a remarkable consensus in the literature about the fundamental dimensions that seem to underlie social judgment regardless of whether one is talking about lay judgments or those informed by the social sciences. When making judgments of people, trait terms, behaviors, groups, or cultures, it seems that target objects are seen to differ along the same two fundamental dimensions again and again, albeit with different names and slightly different interpretations.

In the person-perception domain, a fundamental insight of the classic work of Asch (1946) was that attributes such as cold or warm conveyed very different information from attributes such as intelligent, industrious, and critical. Rosenberg, Nelson, and Vivekananthan (1968) formalized this insight by systematically examining the structure underlying traits and person judgments. Although they explored a number of alternative spatial representations, they ultimately argued that a two-dimensional representation was the most appropriate, defining the resulting dimensions as *intellectual good/bad* and *social good/bad*.<sup>1</sup> At one end of the first dimension were traits such as intelligent and

determined, contrasting with foolish and irresponsible at the other end, whereas at one end of the second dimension were traits such as sociable and helpful, contrasting with unpopular and irritable.

This result was well anticipated by work in the domain of personality psychology, where very similar two-dimensional representations had previously been thought to underlie actual differences in personalities. Early models of personality posited a circumplex structure in which personality traits were arranged in a two-dimensional space (e.g., Freedman, Leary, Ossorio, & Coffrey, 1951; Wiggins, 1979). A common interpretation of the dimensions thought to underlie this space was that personalities differed in dominance (vs. submissiveness) and friendliness (vs. hostility). Subsequent elaborations suggested that additional dimensions were important in personality assessment (i.e., the Big Five; Goldberg, 1981; McCrae & Costa, 1987; Norman, 1963), but these elaborations did not ultimately question the explanatory power of the two fundamental dimensions of the circumplex model (e.g., McCrae & Costa, 1989). Although the terms that were used for these dimensions were not the same as those that Rosenberg et al. (1968) coined, the distinctions were very similar, and the attributes that defined the endpoints of the dimensions were virtually identical.

Regardless of whether the intent is to describe how persons are perceived (e.g., the intent of Rosenberg et al., 1968), the actual differences among personalities (e.g., the circumplex model of

---

Charles M. Judd and Laurie James-Hawkins, Department of Psychology, University of Colorado at Boulder; Vincent Yzerbyt, Department of Psychology, Catholic University of Louvain, Louvain-la-Neuve, Belgium; Yoshihisa Kashima, Department of Psychology, University of Melbourne, Parkville, Victoria, Australia.

This work was partially supported by National Institute of Mental Health Grant R01-45049 to Charles M. Judd. We thank Elizabeth Towner for her considerable help in data collection and coding. We also thank the members of the University of Colorado Stereotyping and Prejudice Lab for their comments on this research and on the manuscript.

Correspondence concerning this article should be addressed to Charles M. Judd, Department of Psychology, University of Colorado, Boulder, CO 80309-0345. E-mail: charles.judd@colorado.edu

---

<sup>1</sup> Rosenberg et al. (1968) explored, for instance, whether these two dimensions better described the resulting trait space than the widely accepted three dimensions at the core of the semantic differential: evaluation, potency, and agency (Osgood, Suci, & Tannenbaum, 1957). They concluded that their two-dimensional definition was a more adequate representation of the trait (and person-perception) space.

personality; Wiggins, 1979), or the semantic meaning of trait terms implicit in language (e.g., Brown, 1986; Schweder, 1975), the point is that underlying dimensions in judgments of target individuals and trait terms seem to revolve around very similar two-dimensional structures. Additionally, the anthropological evidence suggests that these two dimensions seem fundamental to social judgment across cultures and languages (White, 1980). It is important to note that a great deal of subsequent research in the domain of person perception and interpersonal behavior has simply taken these two dimensions for granted (e.g., Hastie & Kumar, 1979, in person memory; Reeder, Pryor, & Wojciszke, 1992, and Wojciszke, 1994, in inferences from behavior; and Strong et al., 1988, in the reciprocity of behaviors).

In addition to this work in person and trait judgments, there is a long line of work that has attempted to understand the ways in which people think about significant social groups, that is, group stereotypes. Much of the early research on the content of group stereotypes was largely descriptive (e.g., Devine & Elliot, 1995; Gilbert, 1951; Karlins, Coffman, & Walters, 1969; Katz & Braly, 1933), but more recently, attention has been focused on systematic regularities in the content of group stereotypes. Exploring stereotypes about European nationalities, both Linssen and Hagendoorn (1994) and Phalet and Poppe (1997) have argued that group stereotypes vary along two dimensions that are remarkably similar to those that have been of interest in the person-perception and personality literatures. Similarly, Fiske and her colleagues have put together an impressive line of work exploring the systematic regularities in the content of group stereotypes (Fiske, Cuddy, Glick, & Xu, 2002; Fiske, Xu, Cuddy, & Glick, 1999; Glick & Fiske, 2001a, 2001b). Taking their cue from Rosenberg et al.'s (1968) work in person perception, they have argued that the same two dimensions underlie the content of most group stereotypes. They have referred to these as *competence* and *warmth*, but they are essentially identical to the Rosenberg et al. dimensions of intellectual good/bad and social good/bad. They have further suggested that out-groups have the capacity to elicit particular emotional responses depending on the stereotypic beliefs about those groups along these two dimensions. Out-groups that are seen to be high on competence but low on warmth elicit envy, whereas those that are high on warmth but low on competence are likely to be pitied.

A particularly productive application of this two-dimensional approach to group stereotypes has been in the domain of gender stereotyping and prejudice, where Glick and Fiske (2001a, 2001b) have argued that many gender beliefs are fundamentally ambivalent. On the one hand, traditional women (e.g., housewives) may be judged to be warm but not particularly competent, a benevolent sexism response. On the other hand, nontraditional women (e.g., feminists) may be seen as competent but hostile, a hostile sexism response. In fact, there is long tradition in gender stereotyping that has focused on exactly these two dimensions in explaining how gender differences are perceived and how they affect patterns of gender discrimination and bias in society (e.g., Altermatt, DeWall, & Leskinen, 2003; Bakan, 1966; Cross & Madson, 1997; Eagly & Mladinic, 1989; Eagly, Wood, & Diekmann, 2000; Glick & Fiske, 1996; Glick, Fiske, & Mladinic, 2000).

Finally, in both social scientists' and lay people's descriptions of entire societies and cultures, the same two fundamental dimensions seem to recur.<sup>2</sup> In Tönnies's (1887/1955) and Durkheim's (1893/

1964) classical writings, modern and traditional societies were contrasted by constructs arguably analogous to competence and warmth. In traditional societies, people were seen to be embedded in closely knit communities in which they were suspended in close interpersonal relationships, whereas, in modern societies, individuals were seen to engage in calculated exchange relationships, where individual competence can be expressed more freely. More modern reincarnations of this contrast can be found in the constructs of individualism and collectivism (e.g., Hofstede, 1980; Triandis, 1995) and independent and interdependent self-construals (Markus & Kitayama, 1991). Western cultures are seen to value agentic and individualist qualities, whereas Eastern cultures have been described as valuing more communal and collectivist qualities (for recent reviews, see Kashima, 2001; Oyserman, Coon, & Kimmelmeier, 2002). In terms of lay people's perceptions of societies, both Eagly and Kite (1987) and Poppe and Linssen (1999) have shown that the dimensions analogous to individual competence and interpersonal warmth, or agency and communality, can characterize national stereotypes. In sum, it seems that regardless of the target object of social judgment, the same fundamental underlying dimensions of judgment seem to recur, albeit with exact definitional variations. Like Fiske and her colleagues (1999, 2002), we refer to these dimensions of social judgment as competence and warmth, recognizing that these are very similar to the dimensions Rosenberg et al. (1968) labeled as intellectual good/bad and social good/bad. If we assume that these are fundamental dimensions of social judgment (however labeled), an important next question is whether and how they are related to each other. It was the intent of the present research to explore this question. Although the past literature has largely been in agreement about the importance of these two dimensions, there has been considerably less consensus regarding the relationship between the two.

In their multidimensional analysis, Rosenberg et al. (1968) reported a substantial positive correlation between the two dimensions across the 64 traits they scaled ( $r = .42$ ). Because these trait locations came from the sorting of traits into clusters describing individual targets, an interpretation of this correlation is that individuals who are seen as possessing more positive intellectual qualities are also seen as possessing more positive social qualities, that is, a halo effect. There are those people one likes and others one likes less, and the people one likes tend to be both competent and warm.

In the literature on the content of group stereotypes, Fiske and her colleagues (2002) have suggested that many out-group stereotypes may have a mixed evaluative content, with positive regard on one of the two dimensions of warmth and competence but negative regard on the other. Stereotypes about groups that are judged to be warm but incompetent, such as housewives and the elderly, are called *paternalistic stereotypes*, whereas those of groups that are seen as competent but unfriendly, such as feminists and rich people, are labeled *envious stereotypes*. This distinction is funda-

<sup>2</sup> Some might argue that a culture is just another kind of group. However, Fiske et al. (2002) clearly talked about groups as they are perceived from within a given culture. Accordingly, one might talk about groups nested under cultures and about cultures as distinct, perhaps superordinate, groups.

mental to the notion of ambivalent sexism, as explained earlier (Glick & Fiske, 2001a, 2001b).

The mixed content model of out-group stereotypes suggests that in the judgment of groups, there may exist a negative relationship between the two dimensions of competence and warmth. Of course, Fiske et al. (2002) did not suggest that all group stereotypes are of this mixed nature. For instance, they found that some groups (such as welfare recipients) are negatively regarded on both dimensions, whereas in-groups tend to be regarded positively on both. So, the suggested negative correlation in the judgment of stereotyped groups is far from perfect. Nevertheless, the suggestion in the stereotyping domain is rather different from that obtained in the research of Rosenberg et al. (1968).

Finally, in descriptions of cultural differences, the prevailing wisdom suggests a strong negative correlation between the two dimensions. In Hofstede's (1980) research on work values, he extracted a single dimension whose bipolar opposites he characterized as *individualism* and *collectivism* (clearly related to competence/agency and warmth/communality; see Kashima, 2001, for a discussion). It has often been assumed that cultures that tend to be solidarity oriented (valuing community and kinship) tend not to be those cultures that are individualistic (valuing individual achievement and efficiency). Although the cross-cultural data do not always show a negative correlation between individualism and collectivism across actual cultures (e.g., Oyserman et al., 2002), perceptions of a negative correlation persist among social scientists.

The obvious question from this quick summary of the relevant literature is why the two dimensions seem to be positively related in the judgment of traits and individuals, whereas the suggestion is that they tend to be negatively related to each other in the judgment of groups and cultures. Perhaps the difference lies in the target objects of judgment, or perhaps the difference is to be found someplace else, in some other difference in the research on this issue to date.

Given the centrality of these two dimensions to social judgment, it seems surprising that no one has addressed the issue of how they are related in a systematic way. All the work to date on the relations between these dimensions has been purely correlational, reporting how they covary across traits, across persons, across stereotyped groups, and across cultures. In light of this, it seemed important to begin to understand the dynamic relation between the two dimensions, approaching the question from an experimental point of view. If one were to manipulate a target's position on one of the two dimensions, what consequences would there be for how that target is seen on the other? How would this dynamic experimental relationship between the two dimensions depend on the nature of the target entity being judged and the conditions under which those judgments were made?

We started by addressing the issue in the context of group stereotypes. In our first two studies, we presented participants with descriptions of two novel groups. These were described with behaviors that individual members of those groups had supposedly engaged in. As described in the *Methods* section for Studies 1 and 2, in the first study, one group was described with largely high-competence behaviors, and the other was described with largely low-competence behaviors. In the second study, the two groups differed along the warmth dimension. Additionally, for both groups in both studies, equivalent and ambiguous information was given about the two groups on the second unmanipulated dimen-

sion (i.e., warmth in Study 1 and competence in Study 2). Our fundamental question was how these two groups would be judged on this second dimension. Following the mixed stereotype content model of Fiske et al. (2002), we thought one might reasonably expect a negative relationship between the two dimensions: The group that was described more positively on the manipulated dimension would be judged less positively on the second, regardless of whether we were going from competence to warmth (Study 1) or the other way around (Study 2). On the other hand, it also seemed possible that we would find an overall halo effect consistent with Rosenberg et al. (1968), such that a group judged positively on the manipulated dimension would be also judged positively on the unmanipulated second dimension.

Following these first studies, we explored various factors that could moderate the relationship we obtained between the two dimensions. For instance, we varied whether the target entities were described as groups versus individuals. As we show below, the factors we explored in this regard ultimately proved fruitful in helping us to understand the inconsistencies between the positive relationship found by Rosenberg et al. (1968) and the negative relationship suggested by Fiske et al. (2002) in the group-stereotyping literature.

Before we began, we conducted pretests to scale the behaviors to be used as stimulus materials in our studies, establishing their locations on the relevant dimensions of competence and warmth. We first report these pretest results.

### Pretests

Seventy-nine behaviors were generated and pretested for use in the following studies. Because we wanted to use these behaviors to manipulate one dimension without affecting the other, the primary goal in generating these behaviors was to identify behaviors that were diagnostic on one of the two dimensions (either high or low) while being nondiagnostic on the other dimension. Two pretests were completed. In the first, participants ( $n = 22$ ) were asked to judge each behavior first on one dimension and then on the other. The two rating questions were "How motivated, intelligent, energetic, and organized do you think the person who did this behavior is?" and "How sociable, warm, friendly, and caring do you think the person who did this behavior is?" Responses were given on  $-4$  (*not at all*) to  $4$  (*very much*) scales.

We calculated means for each behavior on each question and then correlated them across behaviors. We had generated behaviors with the hope that the obtained correlation would be close to zero. Instead, the obtained correlation in this first pretest was large and positive ( $r = .52$ ), indicating that behaviors seen as high on one dimension were also judged to be high on the second.

It seemed possible that this positive correlation might have been due to the fact that participants made both ratings for each behavior sequentially. To examine this possibility, we conducted a second pretest in which the rating dimension varied between participants ( $n = 14$  in each case). Even so, the obtained correlation remained the same ( $r = .54$ ), indicating that, regardless of the design of the pretest, behaviors that were seen positively on one dimension were also seen positively on the second (replicating the results found by Rosenberg et al., 1968).

From the pretest means, we attempted to select behaviors that were diagnostic of one of the two dimensions (either high or low)

but relatively neutral on the second. In spite of our best efforts, the two mean ratings were still correlated .49 across the behaviors actually used in the following studies (these behaviors are given in the Appendix, along with their mean pretest ratings on the two dimensions). The continuing presence of a positive correlation in the actual behaviors used as stimuli in our studies suggests that our test of whether the two dimensions would be negatively related in the perception of groups is a particularly strong one.

### Studies 1 and 2

The first two studies we report shared a common goal: to examine how novel groups described as relatively high or low on one of the two dimensions would be judged on the second. Participants were provided with behavioral information for the two groups. In the first study, one of the two groups was described with largely high-competence behaviors, whereas the second was described with largely low-competence behaviors. In the second study, the behaviors attributed to the two groups varied in warmth rather than competence. Following the presentation of the behaviors, each group was rated on several trait dimensions related to the two underlying dimensions. In addition to these trait ratings, we also asked participants to complete within-group histograms for each group estimating the relative proportions of group members at five different levels along the trait dimensions. From these, we could measure the perceived central tendencies of the groups (analogous to the trait ratings) and the perceived within-group variability. Although we had no specific hypothesis, we thought that our manipulations might affect perceived within-group variability as well as mean group judgments. No consistent effects in this regard were found. Hence, we report only results using the perceived group means from these histogram judgments.

### Method

#### Participants and Design

Participants were 96 undergraduate students ( $n_s = 32$  in Study 1 and 64 in Study 2) at the University of Colorado who participated in partial fulfillment of an introductory psychology course requirement. All participants were told that the study concerned how people formed impressions of groups. The design of each study included only a single within-subject factor: target group. In Study 1, the manipulated target group difference was along the competence dimension, whereas, in Study 2, it was along the warmth dimension. Because the two studies were completed at different times and involved somewhat different measures (as described below), we preserve them here as separate studies rather than analyzing their data together.

#### Procedure

Between 1 and 3 participants were run at a time. Participants arrived at the laboratory and filled out an informed-consent form. They were told that they would be asked to form impressions of two groups, the "Greens" and the "Blues," based on behaviors that individual group members told us they had performed. Participants were then given a card set that contained 32 behaviors, one per card. Sixteen of these were attributed to the Greens and 16 to the Blues. The cards were presented in a random order, with behaviors from the two groups intermixed. Participants were paced through the cards by the experimenter to ensure that the behaviors were closely read, and as they went through them, they were asked to sort them into two

piles, one for each group. They were then asked to read the behaviors a second time, one group at a time, individually at their own pace.

After reading the behaviors of each group, participants were asked to write a paragraph-long impression of each. This was done to solidify their impression of the group as a whole. They then rated each group on eight 9-point trait-rating scales asking participants to indicate the extent to which they thought people in the group possessed each trait, with response options going from  $-4$  (*not at all*) to  $4$  (*very much*). The rated traits included two high and two low traits from each dimension (high competence: capable and skilled; low competence: lazy and disorganized; high warmth: sociable and caring; low warmth: unfriendly and insensitive).

Following the trait ratings, participants were asked to complete a histogram task for each group on the same traits used in the trait-rating task. The histogram consisted of five columns for each trait, and participants were asked to fill in the height of the five columns to indicate the relative numbers of group members at each level of the trait. The height of the area filled in was then measured. From these measurements, we computed the mean of the histogram distribution, assigning scale values of  $-2$  to  $2$  to the five columns. Histograms were completed for both groups on one trait before proceeding to the next trait.

Study 2 included an additional task that was not included in Study 1. Because we wanted to examine recognition memory for behaviors that had not been presented, particularly on the trait dimension that had not been manipulated, we gave participants a list of behaviors, some of which they had seen and some of which they had not, and asked them to indicate for each whether they had seen it and, if so, with which group. Participants were then fully debriefed and dismissed.

### Materials

In Study 1, competence was the manipulated dimension, with our interest being whether differences along the warmth dimension would subsequently be perceived. In Study 2, we reversed the role of these two dimensions, manipulating warmth between the two groups.

In both studies, 16 behaviors were used to describe each of the two groups. Of these, 6 were taken from one end of the manipulated dimension and 2 from the other, 2 were from each end of the unmanipulated dimension, and 4 were irrelevant to the two dimensions. Thus, in Study 1, the high group had 6 high-competence behaviors and 2 low, and the low group the reverse. In Study 2, the high group had 6 high-warmth behaviors and 2 low, and the low group had the reverse.

In the case of each group, 2 high and 2 low behaviors along the unmanipulated dimension (warmth in Study 1 and competence in Study 2) were included so that participants would feel they had some information relevant to that dimension for each group. Which specific behaviors were associated with each group was counterbalanced across participants so that each high and low behavior on the unmanipulated dimension occurred equally often with both groups. Thus, on average across participants, both groups were described with the identical behaviors on the unmanipulated dimension.<sup>3</sup> In addition to the counterbalancing of behaviors, we also counterbalanced whether the high or low group was the Blue group or the Green, as well as the order in which participants gave their impressions of the two groups.

### Results

#### Trait Ratings

To analyze the trait ratings given the two groups, we collapsed across the four ratings given each group on each dimension. This

<sup>3</sup> We also counterbalanced the behaviors on the manipulated dimension as well as the neutral filler behaviors. As a result, each presented behavior was seen with both the high group and the low group across participants.

was done by averaging the ratings on the two traits that were positively valenced on a dimension and the two traits that were negatively valenced on that dimension and then taking the difference between these two averages, resulting in possible scores of -8 to 8 for each trait. More positive scores on these differences thus meant that a group was judged more competent and more warm. In each study, we analyzed these scores treating target group (high vs. low on manipulated dimension) and trait dimension (competence ratings vs. warmth ratings) as within-subjects factors.

*Study 1.* The analysis revealed a significant main effect for target group,  $F(1, 31) = 10.67, p < .005$ ; a significant main effect for trait dimension,  $F(1, 31) = 4.46, p < .05$ ; and a significant Target Group  $\times$  Trait Dimension interaction,  $F(1, 31) = 28.69, p < .0001$ . The relevant means are presented in Table 1.

The target-group main effect suggests that, on average, the high-competence target group was rated more positively on both dimensions than the low-competence target group. The trait-dimension main effect suggests that more positive competence than warmth ratings were given. Of most interest was the obtained Target Group  $\times$  Trait Dimension interaction. The interaction itself is not diagnostic of our predicted results as it would likely have emerged simply if the trait difference on the manipulated competence dimension were greater than on warmth. However, the two associated simple effects are directly relevant. The significant simple target-group difference on competence ratings,  $F(1, 31) = 42.97, p < .0001$ , represents a manipulation check. Most diagnostic is the simple difference on warmth, the unmanipulated dimension. Here, the significant difference,  $F(1, 31) = 5.16, p < .03$ , is in the opposite direction: The high-competence group was judged to be less warm than the low-competence group.

*Study 2.* As in Study 1, the target-group main effect,  $F(1, 61) = 57.08, p < .0001$ ; the trait-dimension main effect,  $F(1, 61) = 48.01, p < .0001$ ; and the interaction,  $F(1, 61) = 57.08, p < .0001$ , were all significant. The relevant means are contained in Table 2. The high-warmth group was rated higher than the low group, on average, on both dimensions. Higher ratings on average were given on the manipulated warmth dimension. The decomposition of the interaction revealed both the unsurprising difference on the manipulated warmth dimension,  $F(1, 61) = 121.53, p < .0001$ , and a theoretically important target-group difference on the unmanipulated competence dimension,  $F(1, 61) = 5.52, p < .03$ , in the reverse direction: The high-warmth group was judged as lower in competence than the low-warmth group.

Table 1  
*Means for High- and Low-Competence Groups on Competence and Warmth (Study 1)*

Trait	Trait-rating Ms		Histogram Ms	
	High target	Low target	High target	Low target
Competence	4.99	-1.02	0.92	-0.16
Warmth	0.06	2.34	-0.10	0.68

*Note.* Trait-rating means presented are on a -8 to 8 scale. Histogram means presented are on a -2 to 2 scale.

Table 2  
*Means for High- and Low-Warmth Groups on Competence and Warmth (Study 2)*

Trait	Trait-rating Ms		Histogram Ms	
	High target	Low target	High target	Low target
Competence	1.94	3.42	0.47	0.60
Warmth	4.94	-2.53	0.95	-0.54

*Note.* Trait-rating means presented are on a -8 to 8 scale. Histogram means presented are on a -2 to 2 scale.

### Histogram Means

*Study 1.* Parallel analyses were done on the means from the generated histograms. Mean values are given in Table 1. This analysis revealed a significant Target Group  $\times$  Trait Dimension interaction,  $F(1, 31) = 35.65, p < .0001$ . As in the trait-rating data, the high-competence target group was perceived as higher in competence and lower in warmth, whereas the low-competence group showed the reverse pattern. Both simple target-group differences were significant. The high-competence target group was seen as significantly higher in competence than the low group,  $F(1, 31) = 37.36, p < .0001$ , and significantly lower in warmth than the low-competence target group,  $F(1, 31) = 14.43, p < .001$ .

*Study 2.* The analysis of the histogram means from Study 2 (given in Table 2) revealed significant main effects of both target group,  $F(1, 60) = 61.35, p < .0001$ , and trait dimension,  $F(1, 60) = 34.12, p < .0001$ , as well as a significant interaction between them,  $F(1, 60) = 67.33, p < .0001$ . The simple difference on the manipulated warmth dimension was of course highly significant,  $F(1, 60) = 124.39, p < .0001$ . However, the difference on the competence dimension, such that the high-warmth group was judged less competent, was not significant,  $F(1, 60) = 1.01, p > .25$ .

### Recognition

At the end of Study 2, we gave participants a list of 16 behaviors, all related to the unmanipulated dimension of competence. Half were high competence, and half were low; crossed with that, half had been seen before, and half had not. Participants were asked to indicate for each whether it had been seen before and, if so, with which group it had been seen. This could yield five types of recognition response: (a) previously seen behaviors recognized and attributed to the correct target, (b) previously seen behaviors recognized but attributed to the wrong target, (c) previously seen behaviors not recognized, (d) previously not seen behaviors identified as such, and (e) previously not seen behaviors incorrectly attributed to one target or the other. Of these, the third type is a perfect function of the sum of the first and second, and the fourth and fifth are also perfectly related to each other. Accordingly, we computed and analyzed three recognition proportions: (a) the proportion of previously seen behaviors recognized and attributed to the correct target, (b) the proportion of previously seen behaviors recognized but attributed to the wrong target, and (c) the proportion of previously not seen behaviors incorrectly attributed to one target or the other. Each of these was analyzed as a function

of behavior type (high competence vs. low) and target (the target to which the behavior was either correctly or incorrectly attributed by the participant), both within participants.

Overall, recognition memory was good, with, on average, 78% of all previously seen behaviors recognized as such and attributed to the correct group. The only significant effect emerging from the analysis of correctly recognized behaviors was a main effect of behavior type,  $F(1, 63) = 4.55, p < .05$ . The previously seen high-competence behaviors were better recognized than the low behaviors. No significant effects emerged for previously seen behaviors that were attributed to the incorrect group. The analysis of the previously unseen behaviors that were incorrectly recognized revealed both a significant main effect of behavior type,  $F(1, 63) = 28.97, p < .0001$ , and a significant Behavior Type  $\times$  Target interaction,  $F(1, 63) = 14.40, p < .001$ . The relevant mean recognition proportions are given in Table 3. As these make clear, on average, new behaviors were more likely to be falsely recognized if high competence than if low. However, this difference was importantly moderated by target group: A previously unseen high-competence behavior was much more likely to be falsely attributed to the low-warmth group,  $F(1, 63) = 9.93, p < .001$ , whereas the opposite was the case for the low-competence behaviors,  $F(1, 63) = 17.67, p < .001$ .<sup>4</sup>

### Correlations

At the level of the mean differences between the target groups, the data from both studies so far suggested a negative relationship between warmth and competence dimensions. When groups differed on one of these, they were seen to differ on the other in the opposite direction. An alternative way to examine this same negative relationship was to examine the correlation between the two judged trait dimensions. To do this, we computed for each participant two indices from the trait ratings: the perceived competence difference between the two groups and the perceived warmth difference between them (in both, taking the difference between the high group on the manipulated dimension and the low group). In the first study, where competence was manipulated, the correlation between these two indices was  $-.32 (df = 30, p < .07)$ . In the second study, manipulating warmth, the correlation was  $-.27 (df = 60, p < .05)$ . Thus, in both studies, the larger the perceived difference between the two groups on the manipulated dimension, the larger the perceived difference between them on the other dimension in the opposite direction.

### Discussion

The results from these first two studies are quite consistent. When one manipulates one of the two dimensions underlying

group stereotypes (competence or warmth) and provides ambiguous (and equivalent) information about the two groups on the other dimension, one finds a consistent negative relationship between the two dimensions. A high-competence group is seen as lower in warmth than a low-competence group, and a high-warmth group is judged to be lower in competence than a low-warmth group. Additionally, those who see larger differences between the two groups on one dimension tend to see larger differences between them on the other, in the opposite direction. Also, in the second study, we found that this negative relationship bleeds over into false memories: A group that is high on one of the two dimensions is falsely remembered as having shown behaviors that are low on the other dimension and vice versa. It is important to note that these results were found even though our own pretest data suggested that the two dimensions would be positively related when the individual behaviors were rated on the two dimensions (consistent with Rosenberg et al.'s, 1968, results). Although more competent behaviors are judged to be warmer (a halo effect in the judgment of individual behaviors), groups that are seen to be higher on one dimension are judged to be lower on the other.

These results are largely supportive of the mixed content stereotype model of Fiske et al. (2002), but importantly, they demonstrate the mixed evaluative content of group stereotypes in an experimental context. Rather than examining the covariation between the two fundamental dimensions of social perception in judging real-world groups, we have shown that groups that are experimentally constructed to be high or low on one of the two dimensions but equivalent on the second are nevertheless judged differently on that second dimension. Additionally, the two studies in combination demonstrate that the links between the two dimensions are reciprocal: Negative inferences are made both from competence to warmth and from warmth to competence.

Having said this, it is also interesting to note that the judged group differences in the second study, although stronger on the manipulated dimension (i.e., perhaps a more extreme manipulation), were weaker on the second (and in fact were not significant on the histogram means). Apparently, for our participants, products of a Western individualistic culture, competence seemed to matter more for inferences of warmth than the other way around.

Having shown in these studies that the two dimensions do share a dynamic negative relationship in the perception of groups, we conducted three additional studies to explore the limiting conditions under which the effect might be observed. First, it seemed likely that the direction of the relationship between the two dimensions might depend on whether individuals (and individual behaviors) were judged or whether groups were judged. The positive correlation between the two dimensions that Rosenberg et al. (1968) reported was found when the individuals were the targets of judgments, whereas the mixed content model of stereotyping was specifically developed to account for the content of group stereo-

Table 3  
False Recognition of Previously Unseen Competence Behaviors  
(Study 2)

Attributed to	Behavior type	
	High competence	Low competence
High-warmth group	0.18	0.30
Low-warmth group	0.40	0.09

<sup>4</sup> All recognition data were also analyzed using signal-detection procedures, treating seen behaviors that were recognized and correctly attributed as hits and unseen behaviors that were falsely attributed to a target as false alarms. Signal-detection analyses were entirely consistent with those presented in the text, suggesting that there was a bias toward attributing low-competence behaviors to the high-warmth group and high-competence behaviors to the low-warmth group.

types. Accordingly, our third study used exactly the same stimulus materials and design as the first study, but additionally, we manipulated the target entities that participants were asked to evaluate. Half of the participants were given the group instructions of Study 1. The other half were told that two individuals were being described and that their job was to form impressions of these two people. In this study (and the subsequent ones), for the sake of simplicity, we limited ourselves to inferences from manipulations of competence to judgments of warmth. We had already shown that the negative relations are bidirectional; now, we took the stronger of the two effects and explored its limiting conditions.

### Study 3

#### Method

##### Participants and Design

Participants were 64 undergraduate students at the University of Colorado who participated in partial fulfillment of an introductory psychology course requirement. Participants were shown behaviors that purportedly described either two groups or two individuals; one of these was described with predominantly high-competence behaviors and the other with low-competence behaviors. Hence, the design involved one between-subjects factor (entity: groups vs. individuals) and one within-subject factor (target: high competence vs. low).

##### Procedure and Materials

Between 1 and 8 participants were run at one time. They were told that they would be asked to form impressions of either two groups or two individuals. They were then given the same set of 32 cards as in Study 1 and paced through these in the same way as in Study 1. In the case of participants forming impressions of groups, the behaviors were attributed to either the Greens or the Blues, as in Study 1. For those forming impressions of two people, the behaviors were attributed to either "Person Green" or "Person Blue." Counterbalancing procedures were used identical to those in Study 1, thereby assuring that the warmth behaviors used to describe the two groups or individuals were identical across participants.

After they went through the behaviors, participants completed the same trait-rating and histogram tasks as in Study 1 for both targets. In the case of the histogram task, the instructions were modified for participants who had been asked to form impressions of two individuals. They were asked to think about different situations in which a person might find him or herself and to fill in the histogram to indicate the relative proportions of those situations in which he or she would manifest the given traits.

After both groups or individuals had been rated, participants were given a surprise recognition test. They were shown a list of 16 behaviors, half high on warmth and half low. Additionally, half had been seen before and

half not. Participants were asked to recall those behaviors they had seen and to indicate with which group (or individual) the behaviors had been seen.

## Results

### Trait Ratings

As in Studies 1 and 2, a positive-trait-minus-negative-trait mean-difference score for each group on each dimension was computed from the eight trait-rating scales. These were analyzed as a function of target (high competence vs. low), trait dimension (competence vs. warmth), and entity (individual or group). The first two varied within participants and the last between them. The relevant means are given in Table 4. The main effect for entity was not significant,  $F(1, 62) = 1.08, p > .30$ , suggesting that, on average, on the two dimensions, groups and individuals were rated similarly. As in Study 1, there was a significant main effect for target,  $F(1, 62) = 11.73, p < .001$ ; a significant main effect of trait dimension,  $F(1, 62) = 14.79, p < .0003$ ; and a significant interaction between these two factors,  $F(1, 62) = 41.81, p < .0001$ . The high target was given higher ratings, on average, collapsing across the two dimensions; higher ratings were given on competence than warmth; and most importantly, the competence difference was reversed on the warmth dimension. It is important to note that this interaction was not qualified by entity,  $F(1, 62) = 0.00, p > .95$ , thus indicating that the entity manipulation (group vs. individual) had little effect on the magnitude of the predicted interaction.

As before, simple analyses were conducted on each trait dimension. The competence difference was unsurprisingly large and significant,  $F(1, 62) = 55.74, p < .0001$ . On warmth, the simple target difference was also significant but in the opposite direction,  $F(1, 62) = 6.91, p < .01$ . Neither of these simple differences was moderated by entity, both  $F_s < 1.00$ .

### Histogram Means

Similar analyses were conducted on the means from the histogram task. The relevant means are given in Table 4. As before, there emerged a significant Target  $\times$  Trait Dimension interaction,  $F(1, 62) = 29.57, p < .0001$ . This was not qualified by entity,  $F(1, 62) = 0.02, p > .88$ . Simple effects tests within trait dimension showed the unsurprising target difference on competence,  $F(1, 62) = 30.42, p < .0001$ , and the more interesting significant target difference on warmth,  $F(1, 62) = 4.37, p < .04$ , in the opposite

Table 4  
Means for High- and Low-Competence Groups and Individuals on Competence and Warmth (Study 3)

Trait	Trait-rating Ms				Histogram Ms			
	Group target		Individual target		Group target		Individual target	
	High	Low	High	Low	High	Low	High	Low
Competence	4.95	-0.98	4.66	-0.11	0.81	0.04	0.52	0.06
Warmth	0.27	1.80	0.08	2.64	0.07	0.23	0.04	0.46

Note. Trait-rating means presented are on a -8 to 8 scale. Histogram means presented are on a -2 to 2 scale.

direction. Neither of these simple target effects was moderated by entity, both  $F_s < 1.00$ .

### Recognition

All behaviors included on the recognition task were relevant to the warmth dimension only. Half were high warmth, and half were low; crossed with this, half had been seen before, and half had not. As in Study 2, we computed and analyzed three recognition proportions: (a) the proportion of previously seen behaviors recognized and attributed to the correct target, (b) the proportion of previously seen behaviors recognized but attributed to the wrong target, and (c) the proportion of previously not seen behaviors incorrectly attributed to one target or the other. Each of these was analyzed as a function of behavior type (high warmth vs. low) and target (the target to which the behavior was either correctly or incorrectly attributed by the participant), both within participants, and entity, between them.

Overall, memory for behaviors that had been seen was very high. On average, participants recognized and correctly attributed 65% of the behaviors that they had seen. The only effect to emerge from the analyses of these seen and correctly attributed behaviors was a main effect of entity (group vs. individual),  $F(1, 62) = 28.92, p < .0001$ , such that behaviors attributed to individuals were more likely to be correctly remembered than the same behaviors attributed to groups (83% vs. 48%).

For behaviors that had been seen but were attributed to the wrong target, there was a parallel effect for entity,  $F(1, 62) = 38.46, p < .0001$ , such that participants were more likely to misattribute behaviors to the wrong target in the case of groups than in the case of individuals.

For the previously unseen behaviors that were falsely attributed to one target or the other, we replicated the interesting results of Study 2. The analysis revealed a marginally significant Target  $\times$  Behavior Type interaction,  $F(1, 62) = 3.14, p = .08$ . Previously not seen high-warmth behaviors were nonsignificantly more likely to be falsely attributed to the low-competence target group than the high-competence one (19% vs. 11%),  $F(1, 62) = 2.46, p = .12$ , whereas low-warmth behaviors were marginally more likely to be falsely attributed to the high-competence group than the low one (22% vs. 13%),  $F(1, 62) = 3.01, p < .10$ . This interaction was not moderated by entity,  $F(1, 62) = .21, p > .50$ .

### Correlations

As in the earlier studies, we computed the trait-rating difference between the high and low targets on competence and warmth for each participant. These were then correlated. There was a negative correlation between the two in the case of participants thinking about the targets as groups,  $r(30) = -.31, p < .08$ , replicating our earlier results. However, for participants in the individual condition, the two variables were uncorrelated,  $r(30) = .01, p > .95$ .

### Discussion

The results of Study 3 replicated and extended those from the earlier studies. The data from the group condition amounted to a full replication of Study 1. Additionally, the memory results confirmed what we found in Study 2, showing false memories on the

unmanipulated dimension (in this case, warmth) such that high-warmth behaviors were mistakenly attributed to the low-competence target and low-warmth behaviors to the high-competence target. Of most interest in this study was the fact that in the analysis of the mean judgments (from both traits and histogram) and in the recognition data, the results were identical regardless of whether participants were judging groups or individual targets. The high-competence target was seen as lower in warmth than the low-competence target. The only suggestion of a difference due to entity emerged in the correlational results, where the negative correlation between the judged trait differences between the two targets was found only for groups. For individual targets, the correlation was near zero.

As discussed in the introductory section above, previous work in the person-perception domain has suggested that when individuals are judged, a positive relationship emerges between the two dimensions, reflective of a halo effect. Indeed, in our own pretest data, when individual behaviors were judged, this positive relationship between the two dimensions was obtained. It therefore was somewhat surprising to us that the entity manipulation undertaken in this study had so little impact. This suggested that the discrepancy between the results reported in this article and those found earlier are not attributable to differences in the objects of perception and judgment.

An additional difference between that earlier work and the present work, beyond whether groups versus individuals (or individual behaviors) were the object of perception and judgment, is that our experimental approach involved the comparative judgment of two entities, whereas, in the earlier person-perception work and in our own pretest data, a large number of individual targets (either people or behaviors) were judged one at a time, with perhaps less pressure to make comparisons between particular target pairs. It is possible that this difference was responsible in part for the discrepancy between the results we found and the positive correlation between the two dimensions that has been found in other rating data. To examine this possibility, we replicated the procedures of Study 1 again, but this time, we made target group a between-participants factor. That is, rather than judging two groups, one described in high-competence terms and the other described as low on that dimension, participants judged only the high- or the low-competence group. The question was whether we would continue to observe the negative relation between the two dimensions of competence and warmth in this one-group situation.

## Study 4

### Method

#### Participants and Design

Participants were 64 undergraduate students at the University of Colorado who participated in partial fulfillment of an introductory psychology course requirement. The design manipulated target group (high competence vs. low) between participants.

#### Procedure and Materials

All procedures and materials were identical to those used in Study 1 except that each participant was told he or she would be forming an

impression of one group, either the Greens or the Blues, rather than two, and each participant received a stack of 16 cards containing the behaviors that described that one group, rather than the 32 cards given to participants in Study 1. In fact, participants were not even told that there were two groups; they were simply told that they were to form an impression of one group based on the behaviors they would read. The behaviors that were used (across participants) were the same 32 behaviors that had been used in Studies 1 and 3. Half of the participants received the set with 6 high-competence behaviors (and 2 low), and the other half received the set with 6 low-competence behaviors (and 2 high). All sets of behaviors also contained 2 high-warmth, 2 low-warmth, and 4 irrelevant behaviors, with the specific behaviors counterbalanced across participants as before. As in the previous studies, participants were asked to write out their impression of the group and complete a series of trait ratings and histograms, followed by the memory task.

*Results*

*Trait Ratings*

The trait-rating differences, one on competence and one on warmth, were analyzed with target group (high or low competence) between participants and trait dimension within. The resulting means are given in Table 5. This analysis revealed a significant main effect for target group,  $F(1, 62) = 19.03, p < .0001$ , and a significant Target Group  $\times$  Trait Dimension interaction,  $F(1, 62) = 31.22, p < .0001$ . Collapsing across the two dimensions, the high target group was rated more positively than the low one, as in the previous studies. To examine the interaction, we again tested simple target-group differences along each trait dimension separately. Unsurprisingly, there was a significant difference on competence,  $F(1, 62) = 52.57, p < .0001$ . However, the simple test on the warmth dimension showed no difference between the high-competence group and the low one,  $F(1, 62) = 0.01, p > .90$ , and in fact, the mean difference on this dimension was in the opposite direction from that found in the previous studies: The high-competence group was judged slightly higher on warmth than the low group.

*Histogram Means*

A parallel analysis was conducted on the means from the histogram task. The relevant means are in Table 5. Again, we observed a main effect of target group,  $F(1, 61) = 13.29, p < .001$ , and a significant Target Group  $\times$  Trait Dimension interaction,  $F(1, 61) = 14.89, p < .001$ . Just as in the trait ratings, there was a significant simple target-group difference on competence,  $F(1, 61) = 29.47, p < .0001$ , but no difference on warmth,  $F(1, 61) =$

$0.60, p > .40$ . Again, the mean warmth rating for the high-competence group was actually slightly higher than that for the low group.

*Recognition*

In this study, there were four kinds of items included in the recognition task: high- and low-warmth behaviors that either had been previously seen or had not. Unsurprisingly, in this one group study (with only 16 encountered behaviors), recognition memory was excellent. Of the previously seen behaviors, on average, 86% of them were correctly recognized, whereas only 4% of unseen behaviors were incorrectly given false positives. There were no significant differences in these recognition rates as a function of either target group or whether the behavior was high warmth or low.

*Correlations*

In the two earlier studies, we computed the correlation between the target-group difference on two dimensions, finding that the greater the perceived difference on competence, the greater the perceived difference on warmth in the opposite direction. In this study, with target group varying between participants, we computed separate correlations for each group between participants' competence and warmth trait ratings of the group they saw. In the case of the high-competence group, there was a marginally significant positive relationship between the two dimensions,  $r(30) = .31, p = .08$ ; for the low-competence target group, the correlation was of roughly the same magnitude but not significant,  $r(30) = .26, p > .10$ . Contrary to the results in the previous two studies, these positive correlations between the two dimensions confirmed what we found in what Rosenberg et al. (1968) reported: Individuals who perceive a group to be higher on competence perceive that same group to be higher on warmth.

*Discussion*

The consistent negative relationship between the dimensions of competence and warmth found in the earlier studies was not replicated here, when participants were exposed to only one of the two target groups. In fact, there was a nonsignificant tendency in the data from this study to see the target group that was higher in competence as also higher in warmth. Additionally, and importantly, the correlation between the two dimensions across participants was positive, replicating the positive relationship between the dimensions found by Rosenberg et al. (1968) and reported in our own pilot data.

It would thus appear that the process of comparing two targets on these two dimensions is a necessary one if the negative relationship between the dimensions of competence and warmth is to be found. The realization that one group (or individual) is better than the other on one of the two dimensions perhaps leads to a compensation on the other dimension, saying that the second group (or individual) is better on that other dimension. In other words, there is nothing in the absolute location of a target group on one of these dimensions that determines how it is judged on the other. We would suggest also that when two groups (or individuals) are compared, there operates a norm of justice or a

Table 5  
*Means for High- and Low-Competence Groups on Competence and Warmth (Study 4)*

Trait	Trait-rating <i>M</i> s		Histogram <i>M</i> s	
	High target	Low target	High target	Low target
Competence	3.53	-0.78	0.80	0.05
Warmth	1.09	1.02	0.29	0.17

*Note.* Trait-rating means presented are on a -8 to 8 scale. Histogram means presented are on a -2 to 2 scale.

Table 6  
*Means by Membership Condition for High- and Low-Competence Groups on Competence and Warmth (Study 5)*

Trait	Trait-rating <i>M</i> s				Histogram <i>M</i> s			
	Membership condition: High		Membership condition: Low		Membership condition: High		Membership condition: Low	
	High target	Low target	High target	Low target	High target	Low target	High target	Low target
Competence	5.02	-0.75	4.31	0.20	0.63	-0.08	0.48	0.14
Warmth	1.19	2.26	0.05	3.34	0.23	0.38	0.01	0.52

*Note.* Trait-rating means presented are on a -8 to 8 scale. Histogram means presented are on a -2 to 2 scale.

sense that there must be good qualities to everyone, so that this sort of comparison and compensation process operates (e.g., Kay & Jost, 2003).

If this explanation is correct, then there must be a strong countervailing motivation that operates simultaneously when one is a member of one of the two groups being compared. On the one hand, in comparing two groups, a compensatory norm may operate, motivating one to compensate on one dimension for shortcomings on another relative to the comparison group. This, we think, is the process that leads to the effects we have so far demonstrated. On the other hand, there is the countervailing motivation, if one is a member of one of the groups, to assert that the in-group is better than the out-group in every way on both dimensions. Ultimately, if this motivation is sufficiently powerful, it would mean that one simply denies that one's group is inferior on either of the dimensions of social judgment: "My in-group is both warm and competent, and they, that is, the out-group, do not compare on either dimension."

To explore these ideas, we conducted a fifth study that was an exact replication of our first study, presenting participants with two groups that differed in competence and examining the warmth inferences they made. This time, however, we manipulated the supposed group memberships of the participants. Using a quasi-minimal group manipulation, we led participants to believe that they were a member of either the high-competence or the low-competence group. We expected the members of the low-competence group to attempt to minimize the superiority of the other group on competence, whereas, on warmth, they should be especially likely to see their own group as more warm than the other group. As for members of the high-competence group, the motivation to see the in-group more positively on all dimensions would lead them not only to affirm the superiority of their group with respect to competence but also to regard their group as warmer than the low-competence group. The question was whether the motivation for in-group bias would be sufficient in this case to overcome the comparative process that would lead participants to compensate a group for relative virtues on one dimension by seeing it less positively on the second.

## Study 5

### *Method*

#### *Participants and Design*

Participants were 64 undergraduate students at the University of Colorado who participated in partial fulfillment of an introductory psychology

course requirement. As in Studies 1 and 3, target group was manipulated within participants, with each participant forming impressions of both the high- and low-competence target groups. Additionally, in this study, participants' purported membership was manipulated, thus varying group membership (member of high vs. low group) between participants.

#### *Procedure and Materials*

Upon arrival in the laboratory, participants were informed that psychologists had recently discovered a new psychological test that classified people into one of two groups, which we called the Greens or the Blues, and that participants would be asked to learn about both groups and give us their impressions. Participants were then asked to complete this test so that we could inform them about their own group membership. The test presented participants with six cards each containing random dots. They were asked a series of questions about each of these (e.g., How many dots were there? Were they random or ordered?). The experimenter then gathered the participants' answers and took some time to supposedly score their performance. Subsequently, all participants were informed that they were members of the Green group.<sup>5</sup>

After participants learned of their group membership, they were then exposed to the same 32 cards used to describe the high- and low-competence groups in the earlier studies. They went through the same procedures as in the earlier two-group studies, half of them learning that the Green group was the high group and half learning that it was the low group.

## *Results*

### *Trait Ratings*

The competence and warmth trait-rating differences were analyzed as a function of target group (high or low), trait dimension, and group membership (high member or low), with this last factor varying between participants. The resulting means are given in Table 6. This analysis revealed a target-group main effect,  $F(1, 61) = 16.23, p < .001$ ; a trait-dimension main effect,  $F(1, 61) = 8.36, p < .01$ ; a Target Group  $\times$  Group Membership interaction,  $F(1, 61) = 8.18, p < .01$ ; and a Target Group  $\times$  Trait Dimension interaction,  $F(1, 61) = 72.59, p < .0001$ . The first of these effects suggested that the high-competence group was judged more positively on average on both trait dimensions, as found in previous studies. However, this was moderated by group membership such that it was less true among members of the low-competence group:

<sup>5</sup> We identified the groups as Green and Blue to keep things consistent with the earlier studies. Additionally, we extended the usual dot estimation procedure typically used in minimal group research by adding additional questions, and we chose not to identify participants as overestimators or underestimators because we felt that those terms might have undesired associations with the competence and warmth trait dimensions.

for high-group members, simple target difference,  $F(1, 30) = 149.64, p < .001$ ; for low-group members,  $F(1, 31) = 0.85, p > .20$ . This result was consistent with the idea that each group was evaluated more positively on both dimensions by participants who were members of that group than by participants who were not.

At the same time, the Target Group  $\times$  Trait Dimension interaction was significant, consistent with our earlier results. To examine this, we again conducted simple analyses within each dimension, this time expecting that group membership might moderate the target-group difference on each dimension.<sup>6</sup> In the case of competence, we expected everyone would agree that the high-competence group was higher but that members of the low-competence group might minimize the magnitude of this difference, relative to members of the high-competence group. As expected, on competence, there was a significant target-group difference,  $F(1, 61) = 98.22, p < .0001$ . This difference tended to be somewhat larger among members of the high-competence group than among members of the low group,  $F(1, 61) = 2.79, p < .10$ . On the warmth dimension, the high-competence group was evaluated as less warm than the low-competence group,  $F(1, 61) = 14.36, p < .001$ , consistent with our earlier results. Members of both groups judged this warmth difference in the same direction, although it was marginally larger among the members of the low-competence group than among the high group members,  $F(1, 61) = 3.75, p < .06$ .

### Histogram Means

The histogram means are in Table 6. Their analysis revealed a significant Target Group  $\times$  Group Membership interaction,  $F(1, 60) = 7.48, p < .01$ , and a Target Group  $\times$  Trait Dimension interaction,  $F(1, 60) = 34.98, p < .0001$ . The first of these effects showed that each group was evaluated more positively on average on both dimensions by participants who were members of it than by participants who were not. To examine the second of these interactions, we again conducted simple effects analysis within each trait dimension. On the competence ratings, there was a large target-group difference,  $F(1, 60) = 34.52, p < .0001$ . Although both membership groups agreed that the high-competence group was higher, this difference did significantly depend on group membership,  $F(1, 60) = 4.24, p < .05$ , with members of the low group seeing less of a difference between the two groups. On warmth, there was a significant target-group difference,  $F(1, 60) = 4.94, p < .04$ , with the high-competence group seen as less warm than the low group. Although true for all participants, this difference also depended on group membership,  $F(1, 60) = 4.04, p < .05$ . Members of the high-competence group saw less of a difference between the two groups on warmth.

### Recognition

From the recognition data, the only significant effects that emerged were those found previously for recognition errors in response to previously unseen warmth behaviors. On these, there was a significant Target Group  $\times$  Trait Dimension interaction,  $F(1, 61) = 10.18, p < .005$ , such that previously unseen high-warmth behaviors were more likely to be incorrectly attributed to the low-competence target group than the high one ( $M = .40$  vs.  $M = .18$ ), whereas the reverse was true for the low-warmth

behaviors ( $M = .18$  vs.  $M = .33$ ). This difference depended marginally on group membership,  $F(1, 61) = 3.04, p < .09$ , with members of the low-competence target group showing the interaction somewhat more strongly.

### Correlations

As in the earlier two-group studies, we correlated the judged competence difference with the judged warmth difference. Across all participants, the two dimensions were negatively, although not significantly, correlated,  $r(60) = -.20, p < .11$  (partialing out group membership). Although in the same negative direction, this correlation was somewhat reduced in magnitude compared with those found in Studies 1 and 3. The magnitude of this relationship did not depend on whether participants were members of the high- or low-competence group.

### Discussion

This study showed the operation of two separate effects in the judgment of a pair of target groups known to differ on one of the two dimensions of social judgment. On the one hand, there was a clear competence difference between the two groups, and because of this, participants were motivated, as in the earlier studies, to claim that the low-competence group must have other redeeming features, in this case, meaning that they were judged to be more warm in the absence of any relevant diagnostic information to that effect. On the other hand, because participants had been told that they were in fact members of one of the two groups, there was an effort to enhance the positive regard in which they held their own group relative to their regard for the other group. Because they could not completely deny the clear competence difference that had been manipulated, all that the low-competence group members were able to do was to minimize this difference and accentuate the positive regard they had for their group, in compensation, on the warmth dimension. The members of the high-competence group were clearly happy to admit that they were the more competent of the two groups. Not surprisingly, they saw less of a warmth difference favoring the low-competence group than did members of that low-competence group. However, perhaps surprisingly, even they were unable to deny the strength of the comparative compensation motivation. Although showing less of a warmth difference between the two groups, they nevertheless continued to acknowledge that their own group might have been less warm than the other.

It is certainly the case that the minimal group membership manipulation did not instill a strong identification with the group.

<sup>6</sup> It is important to appreciate that this moderation of the target-group difference on both dimensions (in different directions) by group membership is not the triple interaction (which was not significant) of Target Group  $\times$  Trait Dimension  $\times$  Group Membership. This is because the target-group differences on the two dimensions went in opposite directions, and we expected that target-group difference to be greater for the members of the high-competence group on competence but greater for the members of the low-competence group on warmth. In fact, the predicted moderation of the target-group differences on both dimensions by membership group was equivalent to the significant Target Group  $\times$  Group Membership interaction already discussed.

It is also certainly the case that the strength of our competence manipulation did not allow the competence difference to be denied by those who would have been motivated to do so (Cadinu & Cerchioni, 2001). In more real-world situations, where group identifications are potent and objective differences on one or the other dimensions more easily denied, it is likely that the motivation to enhance the standing of one's own group on both dimensions may take precedence over any compensatory motivation to attribute something of value to both groups. Still, the present results and others (Yzerbyt, Provost, & Corneille, 2005) attest to the strength of this compensatory process even in the face of group memberships and the motivation to enhance the value of one's own group.

### General Discussion

Our research started with the observation that there seem to be two fundamental dimensions of social judgment that emerge, albeit with some variation, across the objects that are the focus of social judgment. Names for the first of these two have included intellectual good/bad, dominance, competence, agency, and individualism. Names for the second have included social good/bad, friendliness, warmth, communality, and interdependence. Although we would not want to argue that these are all identical terms (within dimension) or that there are no meaningful differences in how the two dimensions are defined, we were struck by the similarity of the meanings of these dimensions across objects of judgment. They certainly seem more similar than different.

What does seem to vary is how they are related to each other. Are objects of judgments that are positively regarded on one of these dimensions generally positively regarded on the other? Or do objects of judgment tend to be more positively regarded on one dimension but less on the other? Not only does the past literature sometimes suggest a positive and sometimes a negative relationship between the two but so too do the data we report in this article. When we asked participants to judge behaviors designed to be diagnostic of one dimension but not of the other, a strong positive correlation between the two was found, similar to that reported by Rosenberg et al. (1968) in the judgment of traits. We also found such a positive relationship between the two judgmental dimensions when a single social group was judged. On the other hand, when two groups or individuals were comparatively judged, we consistently found that the one judged more positively on one dimension was judged less positively on the other. Additionally, the judgments along the two dimensions were negatively related to each other such that participants who saw a bigger difference between two groups on one dimension tended to see a bigger difference between them on the other, in the opposite direction. Finally, false memories went in the same direction.

How can we now make sense of these apparently conflicting results and those that we have noted in the past literature? We have suggested that in some judgment circumstances, there is a comparative and compensatory process at work such that if one group or person is judged more positively than another, one attempts to rectify this disparity by asserting that the situation must be reversed on the other dimension of social judgment. What are the situations that apparently trigger such a compensatory process?

We suggest that a number of conditions seem necessary. First, it would seem that the objects of judgment need to be people, groups,

or cultures about which there are social norms acting against uniform disparagement. In the judgment of behaviors or traits, it seems less likely that a compensatory motive ought to operate.

Second, such a comparative compensatory process seems likely to operate especially if a comparison is being made between only two objects of judgment. With more than two objects, no particular comparisons may be necessitated, and hence, compensatory processes may be less potent. That said, however, there may be implicit standards of comparison that mandate a comparison of one social object with another particular one even in the context of numerous potential comparison objects. So, in the group-stereotyping literature, there are often spontaneous standards of comparison that exist: African Americans are contrasted with Whites, housewives with professional women, Republicans with Democrats. It is for this reason, we suspect, that the mixed stereotype content model of group stereotypes (Fiske et al., 2002) suggests a somewhat negative relationship between the two dimensions in the judgment of many, but not all, social groups. When it comes to entire cultures, although many cultures may potentially be the objects of judgments, it typically comes down to overly simple comparisons between traditional and modern societies or between Eastern and Western ways of life.

A third condition seems necessary. Our fifth study demonstrated that when one is a member of one of the groups being judged, the tendency toward in-group bias (enhancing the positive regard of the in-group on both dimensions at the expense of the out-group) is likely to conflict with the comparative compensatory process that underlies the negative relationship between the dimensions of competence and warmth that we have shown. When in-group identification is particularly strong and/or intergroup conflict particularly intense, the motivation toward enhancement of the in-group may eliminate any evidence of the compensatory comparative process. That said, however, we suspect that the power of the compensatory process in real-world group comparisons should not be underestimated (Jost & Banaji, 1994; Kay & Jost, 2003; Yzerbyt et al., 2005).

Of course, much of the above is speculative, beyond the realm of anything that our data are capable of confirming. Future studies are necessary to nail down these conditions more precisely. Regardless of that need, it seems to us that the present work has at least laid out the relevant issues in a way that can guide more extended research. Additionally, it seems important to note that the comparative compensatory process we are arguing for underlines the fact that these two dimensions are the most significant dimensions of social judgment when it comes to social evaluation or at least that all others pale in comparison. If other dimensions of evaluative social judgment are seen as important, then multiple compensatory processes become possible. One might be perfectly content to maintain a negative view of a group on both competence and warmth if there exists a third dimension in which compensation is possible (Mummendey & Schreiber, 1983, 1984; Mummendey & Simon, 1989). Thus, our work not only underlines the fundamental nature of these two dimensions to social judgment but also makes clear that the questions of how they are related needs a dynamic and complex answer, made so at least in part by the complex web of motivations that underlie human judgment.

## References

- Altermatt, T. W., DeWall, C. N., & Leskinen, E. (2003). Agency and virtue: Dimensions underlying subgroups of women. *Sex Roles, 49*, 631–641.
- Asch, S. E. (1946). Forming impressions of personality. *Journal of Abnormal and Social Psychology, 41*, 258–290.
- Bakan, D. (1966). *The duality of human existence An essay on psychology and religion*. Chicago: Rand McNally.
- Brown, R. (1986). *Social psychology*. New York: Free Press.
- Cadinu, M. R., & Cerchioni, M. (2001). Compensatory biases after ingroup threat: “Yeah, but we have a good personality.” *European Journal of Social Psychology, 31*, 353–367.
- Cross, S. E., & Madson, L. (1997). Models of the self: Self-construals and gender. *Psychological Bulletin, 122*, 5–37.
- Devine, P. G., & Elliot, A. J. (1995). Are racial stereotypes really fading? The Princeton trilogy revisited. *Personality and Social Psychology Bulletin, 11*, 1139–1150.
- Durkheim, E. (1964). *The division of labor in society* (G. Simpson, Trans.). New York: Free Press. (Original work published 1893)
- Eagly, A. H., & Kite, M. E. (1987). Are stereotypes of nationalities applied to both women and men? *Journal of Personality and Social Psychology, 53*, 451–462.
- Eagly, A. H., & Mladinic, A. (1989). Gender stereotypes and attitudes towards women and men. *Personality and Social Psychology Bulletin, 15*, 543–558.
- Eagly, A. H., Wood, W., & Diekmann, A. H. (2000). Social role theory of sex differences and similarities: A current appraisal. In T. Eckes & H. M. Trautner (Eds.), *The developmental social psychology of gender* (pp. 123–174). Mahwah, NJ: Erlbaum.
- 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*, 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.
- Freedman, M. B., Leary, T. F., Ossorio, A. G., & Coffrey, H. S. (1951). The interpersonal dimension of personality. *Journal of Personality, 20*, 143–161.
- Gilbert, G. M. (1951). Stereotype persistence and change among college students. *Journal of Abnormal and Social Psychology, 46*, 245–254.
- 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. (2001a). Ambivalent sexism. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 33, pp. 115–188). Thousand Oaks, CA: Academic Press.
- Glick, P., & Fiske, S. T. (2001b). Ambivalent stereotypes as legitimizing ideologies: Differentiating paternalistic and envious prejudice. In J. Jost & B. Major (Eds.), *The psychology of legitimacy* (pp. 278–306). Cambridge, England: Cambridge University Press.
- 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.
- Goldberg, L. R. (1981). Unconfounding situational attributions from uncertain, neutral, and ambiguous ones: A psychometric analysis of descriptions of oneself and various types of others. *Journal of Personality and Social Psychology, 41*, 517–552.
- Hastie, R., & Kumar, P. A. (1979). Person memory: Personality traits as organizing principles in memory for behaviors. *Journal of Personality and Social Psychology, 37*, 25–38.
- Hofstede, G. (1980). *Culture's consequences*. Beverly Hills, CA: Sage.
- Jost, J. T., & Banaji, M. R. (1994). The role of stereotyping in system-justification and the production of false consciousness. *British Journal of Social Psychology, 33*, 1–27.
- Karlins, M., Coffman, T. L., & Walters, G. (1969). On the fading of social stereotypes: Studies in three generations of college students. *Journal of Personality and Social Psychology, 13*, 1–16.
- Kashima, Y. (2001). Culture and social cognition: Towards a social cognition of cultural dynamics. In D. Matsumoto (Ed.), *Handbook of culture and psychology* (pp. 325–360). New York: Oxford University Press.
- Katz, D., & Braly, K. (1933). Racial stereotypes of one hundred college students. *Journal of Abnormal and Social Psychology, 28*, 280–290.
- Kay, A. C., & Jost, J. T. (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.
- 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.
- Markus, H. R., & Kitayama, S. (1991). Culture and the self: Implications for cognition, emotion, and motivation. *Psychological Review, 98*, 224–253.
- McCrae, R. R., & Costa, P. T. (1987). Validation of the five-factor model of personality across instruments and observers. *Journal of Personality and Social Psychology, 52*, 81–90.
- McCrae, R. R., & Costa, P. T. (1989). The structure of interpersonal traits: Wiggins' circumplex and the five-factor model. *Journal of Personality and Social Psychology, 56*, 586–595.
- Mummendey, A., & Schreiber, H.-J. (1983). Better or just different? Positive social identity by discrimination against, or by differentiation from outgroups. *European Journal of Social Psychology, 13*, 389–397.
- Mummendey, A., & Schreiber, H.-J. (1984). “Different” just means “better”: Some obvious and some hidden pathways to in-group favoritism. *British Journal of Social Psychology, 23*, 363–368.
- Mummendey, A., & Simon, B. (1989). Better or just different? III. The impact of importance of comparison dimension and relative in-group size upon intergroup discrimination. *British Journal of Social Psychology, 28*, 1–16.
- Norman, W. T. (1963). Toward an adequate taxonomy of personality attributes: Replicated factor structure in peer nomination personality ratings. *Journal of Abnormal and Social Psychology, 66*, 574–583.
- Osgood, C. E., Suci, G. J., & Tannenbaum, P. H. (1957). *The measurement of meaning*. Urbana: University of Illinois Press.
- Oyserman, D., Coon, H. M., & Kimmelmeier, M. (2002). How American is individualism? Relational Americans and other lessons from cultural and cross-cultural research. *Psychological Bulletin, 128*, 3–72.
- 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 favoritism 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.
- Reeder, G. D., Pryor, J. B., & Wojciszke, B. (1992). Trait-behavior relationships in social information processing. In G. R. Semin & K. Fiedler (Eds.), *Language, interaction and social cognition* (pp. 37–57). London: Sage.
- 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.
- Schweder, R. A. (1975). How relevant is an individual difference theory on personality? *Journal of Personality, 43*, 455–483.
- Strong, S. R., Hills, H. I., Kilmartin, C. T., DeVries, H., Lanier, K., Nelson, B. N., et al. (1988). The dynamic relations among interpersonal behaviors: A test of complementarity and anticomplementarity. *Journal of Personality and Social Psychology, 54*, 798–810.

- Tönnies, F. (1955). *Community and association* (C. P. Loomis, Trans.). London, England: Routledge & Kegan Paul. (Original work published 1887)
- Triandis, H. C. (1995). *Individualism and collectivism*. Boulder, CO: Westview Press.
- White, G. M. (1980). Conceptual universals in interpersonal language. *American Anthropologist*, 82, 759–781.
- Wiggins, J. S. (1979). A psychological taxonomy of trait-descriptive terms: The interpersonal domain. *Journal of Personality and Social Psychology*, 37, 395–412.
- Wojciszke, B. (1994). Multiple meanings of behavior: Construing actions in terms of competence or morality. *Journal of Personality and Social Psychology*, 67, 222–232.
- Yzerbyt, V., Provost, V., & Corneille, O. (2005). Not so competent but warm . . . Really? Compensatory stereotypes in the French-speaking world. *Group Processes and Intergroup Relations*, 8, 291–308.

## Appendix

## Pretest Mean Ratings of Behavioral Stimuli

Behavior	Competence		Warmth	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
High competence				
X worked hard on the extra-credit assignment in linear algebra. <sup>a</sup>	2.75	1.03	0.08	0.81
X is very careful when it comes to savings so that buying that first house will be possible. <sup>a, b</sup>	2.61	0.99	0.28	0.74
X organized a student group to give feedback to the university administration. <sup>a, b</sup>	2.64	1.13	1.25	1.11
X practiced the violin piece 20 times a day. After a month, X felt he/she had it right. <sup>a, b</sup>	2.72	1.16	0.00	1.12
X published a short story in a literary magazine while still in college. <sup>a</sup>	2.42	0.87	0.17	0.61
X travels extensively in Europe and speaks several languages. <sup>a, b</sup>	2.67	1.12	0.97	1.28
X won the yearly award for the employee who contributes most to the company's profits. <sup>a</sup>	2.75	0.91	0.56	0.91
X wrote a little computer program that solved a tough calculus integration problem. <sup>a</sup>	2.92	1.03	0.11	1.12
Low competence				
X's electricity was turned off because the bill hadn't been paid. <sup>a, b</sup>	-2.47	1.40	-0.19	0.71
When called upon by the professor, X was confused and unable to answer the question in a coherent way. <sup>a</sup>	-0.97	1.30	-0.19	0.92
X considered dropping out of school because of failing introductory psychology. <sup>a, b</sup>	-2.33	1.37	-0.19	0.82
X's bicycle was stolen several times because he/she forgot to set the lock. <sup>a, b</sup>	-2.16	1.21	-0.14	0.90
X did poorly on the exam because of mixing up the chapters that needed to be studied. <sup>a</sup>	-1.31	1.69	0.00	0.63
X had trouble finding work because he/she was always late for job interviews. <sup>a</sup>	-2.40	1.09	-0.72	1.16
Coworkers have learned not to ask X to organize projects since he/she rarely gets things done on time. <sup>a</sup>	-2.44	1.11	-0.61	1.05
X took almost an hour to find his/her car after parking it in a huge shopping mall. <sup>a, b</sup>	-1.11	1.49	-0.08	0.84
High warmth				
X volunteered to take care of the neighbor's pet when she was out of town. <sup>a</sup>	1.03	0.88	2.11	1.24
X loves to be with other people. <sup>b</sup>	0.76	1.49	2.99	1.02
X always greets friends with a big hug. <sup>a, b</sup>	0.81	1.21	2.39	1.23
X drove friends to the airport at 5:00 in the morning, even though he/she really could have used the sleep. <sup>a</sup>	1.17	1.23	2.75	1.11
X helped a blind woman cross the street. <sup>b</sup>	0.59	1.60	3.04	0.85
X always smiles at strangers on the street just to make their day better. <sup>b</sup>	0.79	1.63	2.55	1.22
X loves to hold hands while walking. <sup>b</sup>	-0.18	1.18	2.17	1.18
X spent hours with a friend after the friend's dog died. <sup>b</sup>	0.41	1.40	3.15	0.78
X enjoys having long conversations with friends. <sup>b</sup>	0.73	1.51	2.74	0.77
X gave up his/her seat on the crowded bus when an elderly woman got on. <sup>a, b</sup>	0.81	0.89	2.67	1.15

## Appendix (continued)

Behavior	Competence		Warmth	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Low warmth				
X rarely talked to the other people in the house that a bunch of them shared. <sup>a,b</sup>	-0.47	1.23	-1.97	1.36
X yelled at the driver who took the empty parking space. <sup>a</sup>	-0.56	1.16	-1.56	1.18
X decided that everyone at the party was pretty shallow and left early. <sup>a,b</sup>	-0.28	1.19	-1.53	1.36
X prefers to go to a movie alone rather than with a friend. <sup>b</sup>	-0.39	1.52	-2.06	1.86
X did not want to congratulate the winner of the competition. <sup>b</sup>	-1.00	1.65	-2.68	1.34
X didn't go to his/her grandmother's funeral because he/she was too busy with work. <sup>b</sup>	-1.39	2.80	-3.58	1.23
X often doesn't respond when his/her colleagues say hello at work. <sup>b</sup>	-1.66	2.00	-3.19	1.13
X yelled at a little girl for coloring outside the lines. <sup>b</sup>	-1.26	2.41	-3.63	1.04
When asked to donate some money for the victims of the flood, X answered that they shouldn't have lived there in the first place. <sup>b</sup>	-1.91	2.13	-3.67	1.43
X couldn't be bothered to give directions to a stranger. <sup>a</sup>	-0.78	1.46	-2.28	1.16
Neutral				
X enjoys reading a good novel. <sup>a,b</sup>	1.17	1.42	-0.05	1.27
X occasionally likes to go to a good restaurant. <sup>a,b</sup>	0.30	1.29	0.36	1.22
X calls his/her parents every once in a while. <sup>a,b</sup>	-0.07	1.48	0.66	1.78
On most days, X stops to get coffee on his/her way to work. <sup>a,b</sup>	0.49	1.62	-0.11	1.10
X told his parents he/she couldn't come home for the holidays. <sup>a,b</sup>	-0.67	1.62	-1.32	1.78
X likes to go for bike rides in the park. <sup>a,b</sup>	1.68	1.22	0.23	1.27
Sometimes X goes to the gym for a workout. <sup>a,b</sup>	1.67	1.40	0.32	1.09
When X gets home, he/she likes to check his/her e-mail. <sup>a,b</sup>	0.93	1.51	0.75	1.52

<sup>a</sup> Behavior used in Studies 1, 3, 4, and 5.

<sup>b</sup> Behavior used in Study 2.

Received April 5, 2004  
Revision received January 10, 2005  
Accepted May 4, 2005 ■