Shared Reality and False Polarization in Intergroup Perception

Girts Dimdins

Abstract

False polarization is an overestimation of a gap between the views of two sides that occurs because social perceivers expect their peers to be more susceptible to group norms than themselves. Shared reality is a generic term denoting social verification of information by individuals and groups. This thesis explores the relation between shared social reality and false polarization in situations where two groups take the perspectives of each other. A special emphasis in the thesis is put on intergroup attributions. In Study 1, two experiments examined the differences between evaluative and descriptive aspects of intergroup perception. In Experiment 1, Latvian and Russian students agreed on the descriptive scores of mutual ratings, but they disagreed on the evaluative ones. In Experiment 2, male and female psychology students did not show any disagreement either for evaluative or descriptive scores, but they expected their peers to be in-group biased in evaluation of both groups. In Study 2, supporters and opponents of Latvia's EU membership exhibited a false polarization effect when rating various causal explanations of attitudinal behavior of both groups. Study 3 replicated this result with free response causal explanations. Content analysis of the data showed that simulated explanations of a target behavior given from the opposite group perspectives differed more (in terms of perceiving the actors' awareness of causes of behavior and in terms of using mental state markers) than those given from the partisans' own perspective. In Study 4, two experiments demonstrated that EU supporters and opponents could not differentiate between explanations given by in-group members and out-group members who imagined the same perspective and responded more favorably to explanations given from the explainers' own perspective than from the imagined perspectives. In general, the results demonstrate the robustness of the false polarization effect across various contexts. The findings also illustrate the complexity of the term "shared social reality." Furthermore, in the context of intergroup perception, these findings suggest that the most appropriate operationalization of this concept should be shared stereotypes between groups.

Key words: Shared reality, social reality, false polarization effect, social perception, perception of bias, social categorization, in-group favoritism, intergroup attribution, intergroup perception, causal explanation.

Acknowledgements

The present thesis is based on the following studies:

1. Dimdins, G., & Montgomery, H. (in press). Differentiating in-group favoritism from shared reality in intergroup perception. *Scandinavian Journal of Psychology.*

2. Dimdins, G., Montgomery, H., & Austers, I. (2003). The false polarization effect in explanations of attitudinal behavior. *Current Research in Social Psychology, 8,* 275-301.

3. Dimdins, G., Montgomery, H., & Austers, I. (2003). *False polarization when imagining others' explanations of attitudinal behavior.* Manuscript submitted for publication.

4. Dimdins, G., Montgomery, H., & Austers, I. (in press). Differentiating explanations of attitude-consistent behavior: The role of perspectives and mode of perspective taking. *Scandinavian Journal of Psychology.*

CONTENTS

INTRODUCTION

"In a Grove", a short story by the early 20th century Japanese writer Ryunosuke Akutagawa, develops like a conventional detective story. A samurai has been killed and his wife raped in a cedar grove nearby a country road. The readers learn the details of the crime from first-person reports of a number of witnesses: a woodcutter who found the body; a monk who was the last to see the man alive; the mother of the woman; and the policeman who arrested the suspect. Despite the unusual form of narrative, every detail fits together neatly as the reader and the police commissioner simultaneously solve the crime. That is, everything fits until the reader hears the stories of the three main characters. The suspected murderer boastfully confesses how he killed the man in the hardest swordfight of his life. The tearful woman reveals how she killed her husband out of shame of what happened to her, failing to commit suicide. Finally, speaking through a medium, the dead man reveals that he killed himself because of his wife's infidelity. What happened in the grove was very real; the man died. However, what the police commissioner has in front of him are three separate realities, realities with the same physical proof. Until at least two accounts agree, the commissioner has little if any hope of finding the "real truth".

What will not escape the reader's notice is that each character's story is very much in line with that character's social role. The thief, the samurai, and his wife all behave exactly like representatives of these social groups would be expected to act in traditional Japanese society. Every narrative serves perfectly to save the narrator's face; everybody did the right thing under the circumstances. The commissioner may find it hard to accept, but perhaps none of the characters are lying. There are as many truths as many faces need to be saved. To rephrase this tenet using the terms of social science, there are as many psychological realities as there are social perceivers reflecting on the same physical reality. Each psychological reality is determined by the perceiver's unique perspective, shaped by his or her interests, individual traits, and social category memberships. Unless these realities are compared and verified against each other, each one is perfectly valid on its own. Akutagawa never lets the reader know what *really* happened in the grove.

Divergent accounts of the same event shaped by the perceivers' interests are by no means limited to eccentric literary characters as the classic case study by Hastorf and Cantril (1954) shows. The reports of the same controversial football game in the campus papers of Princeton and Dartmouth sound as different as the stories told by Akutagawa's characters. When the researchers asked undergraduates of both universities to judge the game, both groups differed in the number of infractions they saw each team to commit and in their interpretation of the seriousness of these infractions. These results lead Hastorf and Cantril to conclude that "there is no such 'thing' as a 'game' existing 'out

there' in its own right which people merely 'observe'. The 'game' 'exists' for a person and is experienced by him only in so far as certain happenings have significances in terms of his purpose . . . . The sharing of significances provides the links except for which a 'social' event would not be experienced and would not exist for anyone" (p. 133).

This "sharing of significances" that turns subjective experiences into objective "reality" constitutes the focus of the present thesis. More specifically, this thesis looks at how the interaction between cognitive factors, such as asymmetric perception of bias in self and others, and social factors, such as group norms and social stereotypes, shape the perception and interpretation of a shared social reality between two groups. Two concepts are central to the research reported below. One is "shared reality", a generic term denoting social verification of information by individuals and groups. The other is "false polarization", an overestimation of a gap between the views of two sides involved in a conflict or controversy. Four empirical studies focus on the relation between both concepts. The thesis introduction provides a theoretical background to these studies and discusses their contribution to social psychology.

The thesis introduction consists of three parts. The first part, divided into seven sections, gives an overview of the previous research in the field of intergroup perception that forms a basis for the empirical studies of this thesis. The first section is devoted to the concept of shared social reality, with a special focus on the problem of defining and operationalizing this concept in social psychological research. The second section gives a short overview of the topic of perspective taking. The third section reviews the empirical evidence of asymmetric perception of cognitive biases in self and others and discusses the "naïve realism" theory of social perception. The fourth section gives an overview of the previous studies of the false polarization effect. The fifth section is devoted to the effects of social categorization on the social perception and inference. The sixth section discusses stereotyping as one of the consequences of social categorization. The seventh section provides information about previous empirical research of intergroup attributions and various possible classifications of causal explanations that can be used in attribution studies. The second part of the introduction provides an overview of the four empirical studies upon which this thesis is based. The third part discusses the major findings and contributions of this thesis, the limitations of the reported empirical studies, and possible directions for further research.

Shared Social Reality

Festinger (1950) concluded, "where the dependence upon physical reality is low the dependence upon social reality is correspondingly high. An opinion, a belief, an attitude is 'correct', 'valid', and 'proper' to the extent that it is anchored

2

in a group of people with similar beliefs, opinions, and attitudes" (pp. 272-273). For Festinger (1950, 1954), there exists a duality between physical and social realities; physical reality is given precedence when the conditions allow for easy testing of assumptions about this reality. The social reality forms the basis of opinions, attitudes, and beliefs for which the physical evidence is ambiguous or impossible. Hardin and Higgins (1996) cite plenty of evidence to Festinger's predictions about social verification of opinions, beliefs, and attitudes. However, they assume no distinction between physical and social realities, arguing that any experience, including individual sensations, can be considered reliable and valid only to the extent that it is socially verified. In the absence of such verification, any experience is random and transitory. Indeed, the very criteria of what is considered "objectively" real or "physically" measurable are also determined through social consensus.

An early demonstration of the role of shared social reality in verifying individual experience is offered by Sherif's (1936, chapter 6) autokinetic experiments. In these experiments, participants where asked to judge the magnitude and direction of movement of a point of light in a darkened room. In reality, the point of light is fixed but appears to move erratically. When the participants performed the task alone, they tended to create a subjective point of reference and subjective norm to estimate the "movement"; there was a strong tendency to preserve this norm throughout subsequent trials. When no externally given reference points were available, the first estimation became a standard for further comparisons. As one might have expected, these standards differed between the participants. When the participants first performed the task by themselves and subsequently in groups of two and three, their judgments quickly converged to form a group norm. When the participants started the task in groups, a group norm was established even faster and was maintained by the individuals when they subsequently performed the task alone. The subjective experiences of group members became objective through social verification.

Hardin and Higgins (1996) argue that when an experience is recognized and shared with others in the process of social interaction, it achieves reliability, validity, generality, and predictability. The reliability of a shared experience is demonstrated by its repeated recognition by others. Through this recognition, an individual learns that the experience is reproducible in others and thus is not random or incidental. Social verification also validates an experience, showing that it is not only a subjective perception but corresponds to some objective reality. It is important to stress, however, that the validity of an experience is achieved within a specific social group, and the same socially validated "facts" can be interpreted differently by another group from a different perspective. The human need for social verification of one's experience is demonstrated by Asch's (1958) experiments on conformity. Experimental manipulations showed that the presence of a "true partner" considerably reduced the tendency to conform to the majority's opinion. Even the slightest evidence of social

3

verification significantly boosted people's confidence in the validity and reliability of their own experience and encouraged them to speak up against the majority. More recently, Luus and Wells (1994) demonstrated that individuals' confidence in their own judgements varied significantly depending on social verification of these judgements. Eyewitnesses' confidence in accuracy of identifying a suspect in a photo-lineup largely depended on whether they believed another witness had identified the same suspect.

Hardin and Higgins (1996) define shared reality as an ongoing, dynamic process of social verification through which an experience achieves the phenomenological status of objective reality. This verification can take place through communication when an individual does not have any direct experience with the subject of knowledge. In this case, people simply learn and compare the attitudes, beliefs, and opinions of others. The verification can also happen through active social interaction where individuals have joint experiences and, on the basis of these experiences, form knowledge. Like Festinger, Hardin and Higgins are interested primarily in functions of shared reality for the individual and the consequences of shared reality on interpersonal interaction and communication. Shared reality has important consequences for creating and maintaining the experience of self: the more a given aspect of self has been recognized in the social verification process, the more "real" it becomes, and the more likely it is to be maintained and defended. Shared reality regulates social interaction by allowing individuals to predict expectations and behaviors of others; it regulates communication by allowing individuals to tailor their messages to a particular audience and to evaluate audience feedback as well as to provide appropriate feedback to other individuals.

For the individual social perceiver, any direct or indirect interaction with others provides some information about existence or absence of shared reality. In addition, it allows people to estimate the extent of overlap in opinions, attitudes, and beliefs of self and others. For a social scientist, however, the definition and measurement of shared reality constitutes a methodological challenge. One can think of several ways how the construct of shared reality can be operationalized. As far as the researchers are interested in the consequences of shared reality for an individual-- self-perception, choice of communicative or behavioral strategies, or general beliefs about the world--shared reality can be defined as an individual's subjective experience of sharing his beliefs, opinions, or attitudes with others. For example, Hardin and Higgins (1996, pp. 58-59) defined the self-attributes of participants in their experiments as "shared" if subjects believed at least one other person thought the attribute characteristic of them. From this point of view, either the definition or measurement of shared reality does not pose a real problem.

The situation is different when shared reality has to be defined and measured on an interpersonal level. Those are situations, for example, when presence or absence of shared reality has consequences for the efficiency of

4

communication (e.g., Nickerson, 1999) or group performance (e.g., Moreland, Argote, & Krishnan, 1996). For a communication or cooperation to succeed, it is important that two or more people really share information, beliefs, opinions, attitudes, or common understanding of objects or events; it is not enough that these people subjectively believe in sharing social reality with the relevant other(s). In such a situation, one possibility to measure shared reality is simply by reference to the outcome of the respective social activity for which the necessity of shared reality is assumed. If communication is successful or if a group succeeds on a task that requires shared understanding and knowledge, one may claim that such an outcome confirms the existence of shared reality between or among the group members. Such an approach, however, would be formal and without any exploratory and predictive qualities. Another way to operationalize shared reality is to measure to what extent group members agree in their views of each other and/or in their evaluations of other people or objects. Any "objective" criteria of accuracy, like physical measurements or expert judgements, are of little practical use. A member of a working group may consider oneself hardworking, but all the other group members may consider the same person as having a habit to ingratiate herself with the boss. The same objective measurement--the long hours the person spends at work--may serve to justify both opinions.

What is considered "accurate" depends on the perspective of each social perceiver. The agreement between/among group members can be studied in two ways (Kenny & Albright, 1987). First, one may ask if the other(s) view the person as self does. Second, one may ask if people know how the other(s) see them and/or other people, objects, events. In this sense, one can say that a shared reality exists in a group if at least two individuals agree on their evaluations of each other and/or their evaluations of another person, event, or object. Alternatively, in groups of more than two individuals the mean rating of the same target by all judges can be taken as the measurement of shared reality, and individual ratings can be compared to that mean to estimate shared reality between separate individuals and the group as a whole. Of course, the researcher has to decide what constitutes a high or low agreement especially in cases where only two individuals judge each other.

The social-regulatory functions of shared reality include creation and maintenance of group norms. On the one hand, group members as individuals actively participate in the creation of group standards and norms, which constitute the shared reality in a group. On the other hand, the group as a whole imposes this shared reality on each member. The group norms largely determine what attitudes, beliefs, and opinions each member is supposed to have. In other words, shared reality in a group is created by shared group identity. Research has shown that people's susceptibility to social influence depends on their self-categorization as members of a particular group and their acceptance of the norms of that group (Turner & Oakes, 1989). Thus, information coming from an

in-group source and seen as prototypical of in-group norms is more readily accepted as true than information coming from other sources (Hogg & Turner, 1987; Van Knippenberg & Wilke, 1992).

This pervasive influence of group norms on individual attitudes, beliefs, and behaviors leads to several questions. What happens when two individuals belonging to two different groups (thus representing different shared realities) interact? To what extent do these individuals share a social reality between them? This question can be answered from two perspectives--by looking for shared reality either at a different level of social categorization, or alternatively, at the same level of categorization (i.e., the dimension of comparison that makes individuals' group memberships salient).

Individuals belong to multiple social categories at the same time (Turner, Hogg, Oakes, Reicher, & Wetherell, 1987), and different categories (i.e., group memberships) can become salient depending on the situational context. No matter how antagonistic the group identities of two individuals may be, there is likely to be another and more inclusive social category that would allow these individuals to see themselves as belonging to the same category (e.g., members of the same society, culture, or, most inclusively, humankind). Societal beliefs shared by the same society ensure that any two members of that society are likely to share some social identity, and through that, also some social reality (Bar-Tal, 2000). A militant atheist and an ardently religious person may not share a single common belief regarding meaning or origin of human life. Still, they are likely to share social norms that require handling their controversy either through civilized discussion or by politely ignoring each other without resolving to rudeness or physical violence. Thus, even with their incompatible social identities (atheist or religious) salient, they still share the common social reality of their society or culture. It is a higher-order shared reality in the sense that it is imposed on these individuals (and on the groups that they represent) by a larger, more inclusive social group. This higher-order shared reality is a necessary pre-condition for proper functioning of the society.

Of course, conflicts and controversies between groups are not the only contexts where the higher-order shared reality is displayed. Different groups can share many beliefs as members of the same society. Societal beliefs may include myths, collective memories, symbols, ideologies, self-images, images of other societies, goals, values, or social aspirations (Bar-Tal, 2000). Moreover, two groups can agree on mutual stereotypes as long as these stereotypes are accepted as true on a higher societal level of social verification (Oakes & Reynolds, 1997). However, although competing groups may agree on the descriptive nature of mutual stereotypes, they are likely to see the same traits as positive when those traits refer to the in-group and as negative when the same traits refer to the out-group (Hilton & Von Hippel, 1996; Peabody, 1968).

An interesting question is whether the representatives of opposing groups share some social reality as members of both social categories at the level of

their ideological controversy. Is there a shared reality between two groups in conflict regarding the subject, nature, and contents of this conflict? Tajfel (1981a) emphasizes that social categorization divides the social world. It is impossible to have an in-group without contrasting it to some out-group. Beliefs, attitudes, opinions, and practices that define groups have their special value for the group members because they are seen as superior to those held by other groups. Billig (2003) notes that human thinking is inherently rhetorical in itself, rhetorical in the sense that internal thinking is a form of self-deliberation and self-persuasion. Thus, one's attitude, belief, or opinion is largely a stance directed against possible counter-stances. On an intergroup level, in-group is defined not only by what it *is*, but also by what it *is not*, or by what relevant *out-groups are* in contrast to the in-group. Reicher and Hopkins (1996) found that the self-categories, which representatives of political factions used to depict themselves and their parties, were not fixed but were rhetorically constructed to use against the position and arguments of the other side. To use the stance of the out-group for strengthening one's own arguments, people should have fairly good knowledge of beliefs and attitudes constituting this stance. Moreover, to construct a successful attack on the opponents' position, group members should to some extent understand the reasoning behind out-groups arguments, so their response can be predicted. The same is true about defending one's own stance against a possible attack. McGuire's (1964) classic experiments on inducing resistance to persuasion serve to illustrate this point: previous exposure ("inoculation") to moderate amounts of propaganda that provoked the participants to look for counter-arguments made these participants more resistant to subsequent attempts of persuasion than unprepared participants. The stance of one side is formulated with the stance of the other side in mind; thus both groups can be said to share a common social reality. In a way, on an intergroup level, the function of creating a shared reality is still social verification, but these are not individual experiences that are verified. They are shared beliefs and attitudes of the groups. The method of verification is not through confirmation (i.e., testing if others share the group's belief), but through argumentation (i.e., testing if the group's belief can stand its ground in a contest with an opposing belief).

The definition and measurement of shared reality between groups is similar to that used for individuals. Shared reality can be operationalized as agreement between groups about how these groups see each other and each other's beliefs and attitudes about other groups or topics. One can do this by asking all group members (or representative samples of group members) to provide their own opinions and to predict the responses of another group to the same questions. Similarity between the actual and predicted group means indicates agreement, and thus, existence of a shared reality.

Perspective Taking

As Hardin and Higgins (1996) point out, shared knowledge is constructed in part through a process of reciprocal perspective taking. Mead (1934) distinguished between *I*, the impulsive, non-reflective self, and *me*, the reflective, socially guided self, that is developed through internalizing social norms. Through *me* the individual achieves self-consciousness as a social being, becomes able to communicate with others, and positions and orients oneself in any given social context and in society as a whole. To develop this reflective aspect of self, an individual must learn to see oneself through the eyes of others, or, in other words, to take the perspective of others. The ability of perspective taking is acquired early in life by interacting with other people, which forces children to start seeing themselves from the viewpoint of others. Approximately at the age of 7 children learn to shift from the egocentric perspective of the world to a more self-conscious perspective, becoming aware of the need for verification of their ideas through discursive argumentation (Piaget, 1928/1999). Developmentally, this shift in human thinking coincides with the learning of geometrical representation of perspectives--learning to mentally represent an object from a perspective from which it is not physically perceived (Piaget & Inhelder, 1956/1997). Taking into account other people's perspectives has been shown to be an important factor for efficiency of social communication and regulation of social interaction (Krauss & Fussell, 1996).

Montgomery (1994) points to the similarities between perspectives in perception and thinking. Similar to the way a certain visual perspective results in an observer seeing an object in a certain way (some features in the foreground, some in the background, and some altogether out of the picture), adopting a certain cognitive perspective influences people's judgements and decisions about different objects. It is assumed that the qualities of an object of perception or judgement are independent of the perspective from which the object is viewed. However, depending on the perspective adopted by the perceiver, certain attributes of the object become more salient. (In perceptional terms, they are brought to the foreground). Any object or event can be seen in a positive or negative light depending on the perspective. The same person can be seen as stingy or economical, or the same act can be interpreted as brave or stupidly reckless. According to Montgomery (1994), an object can be seen from an inside or an outside perspective. When an inside perspective is adopted, the object is seen as affiliated with the subject (perceiver), and the subject focuses on the advantages of the object. In contrast, when an outside perspective is adopted, the disadvantages of the object come into the foreground. It is also possible to distinguish between three subject-based perspectives, depending on the identification of the perceiver. The perceiver can identify with the following: (a) the self only (the object evaluation is guided by personal interests of the self); (b) the self and some other entities, for example, an in-group (the

8

object evaluation depends on the group interests); or (c) some other entity or entities such as an out-group (the object of evaluation depends on the perceiver's interpretation of that out-group's interests). It is assumed that the perceiver is able to switch among the different combinations of subject-based and object-based (i.e., inside and outside) perspectives. This assumption has important implications for studies of intergroup perception: by taking the perspective of an out-group, the perceiver should be able to see one's in-group from an outside perspective. That is, the perceiver can see as the out-group members themselves see the perceiver's in-group. Of course, this is only possible if both groups share some social reality; that is, the perceiver must have at least some knowledge about the norms, beliefs, and opinions of the other group.

There is mixed support regarding the effectiveness of perspective taking in interpersonal and intergroup contexts. Some experiments on interpersonal attribution have found that literally changing the physical perspective in an interaction can reverse the actor-observer effect (Regan & Totten, 1975; Storms, 1973). Anderson and Pichert (1978) report that participants recalled different information from a fictitious scenario depending on the imagined perspective from which the recall was made. Frank and Gilovich (1989) asked their participants to recall an earlier conversation from either their own or an observer perspective. The participants made more dispositional attributions to their own behavior when they recalled the conversation from an observer perspective. Some studies have found that when taking an out-group's perspective, group members indeed seem to identify with that group and start seeing their own group from an outside perspective, exhibiting biases in favor of the group whose perspective they are taking (Austers, 2002a; Austers & Montgomery, 2001). Moreover, groups can be fairly accurate in predicting an out-group's opinion from that group's perspective; it is similar to the actual opinion of the group (Kemdal & Montgomery, 2001). However, other research has documented systematic biases in perspective taking such as imputing one's knowledge to others (Fussel & Krauss, 1991; Nickerson, Baddeley, & Freeman, 1987). Mikula, Athenstaedt, Heschgl, and Heimgartner (1998) found systematic differences in the way actors and recipients interpreted negative incidents in interpersonal relationships when taking each other's perspectives. This means that accuracy in perspective taking is by no means granted and may depend on a number of factors. One such factor can be an assumption that other people are biased and thus not able or willing to see things the same way as the social perceiver does.

Perception of Bias in Self and Others

Pronin, Puccio, and Ross (2002) note that "people readily recognize biases in others that they do not recognize in themselves, and as a result, they make overly negative attributions about others whose views and self-interested

motives seem 'conveniently' congruent" (p. 638). Pronin, Gilovich, and Ross (in press) review a number of studies that support this assertion. According to these authors, the evidence showing the asymmetry of perception of cognitive biases in self versus others comes from three types of studies. First, there are studies that demonstrate people's failure to recognize cognitive biases in themselves. For example, research has shown that people are unaware of how their general evaluation of targets influence their ratings of specific target attributes (Nisbett & Wilson, 1977b) even when informed about the "halo effect" and asked to become conscious of the causes of their ratings (Wetzel, Wilson, & Kort, 1981); they are unaware of the effect that outcome knowledge has on their perceptions (Fischhoff, 1975); they display overconfidence in their predictions despite systematic underestimation of time for task completion (Buehler, Griffin, & Ross, 1994); and they underestimate their capacity to generate satisfaction with future outcomes and mistakenly attribute such satisfaction to external agents (Gilbert, Brown, Pinel, & Wilson, 2000).

Results from the second category of studies suggest that people recognize and expect cognitive biases in others. For example, research shows that actors correctly expect observers to make correspondent inferences (Miller, Baer, & Schonberg, 1979; Van Boven, Kamada, & Gilovich, 1999), but may overestimate the extent of observers' correspondence bias (Van Boven et al., 1999); people have intuitive understanding of the false consensus effect and predict it in others (Krueger & Zeiger, 1993); people expect the actor-observer effect to appear in attributions made by others (Austers & Montgomery, 2001; Kemdal & Montgomery, 2001). These findings demonstrate that although the social observers may not notice the presence of cognitive biases in their own judgements, they often have intuitive knowledge of such biases in the general population and are able to use this knowledge in making predictions about the opinions and behaviors of other people.

Finally, the third group includes studies that have compared the perception of cognitive bias in self versus other people. For example, Kruger and Gilovich (1999) report that married couples, debaters, and video game and dart players expected their peers, but not themselves, to be motivationally biased in allocating responsibility for desirable and undesirable joint outcomes. Van Boven, White, Kamada, and Gilovich (2003) found that people expected their peers to make more extreme dispositional inferences than they did themselves for a situationally constrained actor's behavior. Miller and Ratner (1998) found that individuals overestimate the role of self-interest as a motivator of other people's actions. Heath (1999) reports similar results--people expect others to be more motivated than themselves by extrinsic incentives (like job security or pay) and less motivated by intrinsic incentives (like learning new things or skill development). In four studies examining the mechanisms behind the better-than average effect, Epley and Dunning (2000) found that the participants overestimated the likelihood that they would act in generous or

selfless ways, but were quite accurate in predicting selfishly motivated actions in others. Other studies have shown that people see themselves as less likely to display self-serving bias than the average person (Friedrich, 1996); participants in anchoring studies report they have been less influenced by the experimental manipulation than others, but fail to recognize the impact of the anchor on their own estimates (Wilson, Houston, Etling, & Brekke, 1996); and people expect out-group members to be biased in their evaluations whereas seeing their own evaluations as similar to impartial third-party judgements (Vivian & Berkowitz, 1992).

Pronin, Lin, and Ross (2002) report on what appears to be so far the most extensive and systematic study of bias perception in self and others. They asked the participants to rate their own and others' susceptibility to a number of biases. In two surveys, Stanford students judged themselves less susceptible than and average American or an average fellow classmate to eight different biases (self-serving attributions for success versus failure, dissonance reduction after free choice, the positive halo effect, biased assimilation of new information, reactive devaluation of proposals from one's negotiation counterparts, perceptions of hostile media bias toward one's group or cause, the fundamental attribution error in "blaming the victim", and judgments about the "greater good" influenced by self-interest). In the third survey, these results were replicated with travelers at San Francisco international airport comparing themselves with fellow travelers, and with another seven biases added to the design (friend enhancement, trust of strangers, trust of borrowers, generous attribution, downward and upward comparison biases, and the gambler's fallacy). In addition, participants of this survey rated themselves as less susceptible than their peers to low availability biases (but not high availability biases) and "socially undesirable" biases (but not "socially desirable" biases). Interestingly, however, neither student nor traveler participants rated themselves as less prone than their peers to such personal limitations as procrastination, fear of public speaking, and the planning fallacy. Pronin et al. also found that most participants failed to acknowledge the influence of better-than-average effect on their ratings after being made aware of this bias (Study 2) and that the participants were more inclined to detect self-enhancement bias in their dyad partners than in themselves (Study 3). The authors conclude that "knowledge of particular biases in human judgement and inference, and the ability to recognize the impact of those biases on others, neither prevents one from succumbing nor makes one aware of having done so" (p. 378).

It is tempting to explain the asymmetric pattern of bias perception as a manifestation of human need for positive self-evaluation (Tesser, 1988) through social comparison. Given the evidence of the impact of motivation on human judgements (Kunda, 1990) and given the negative connotations associated with the concept of bias, it does not seem unlikely that people overlook biases in themselves and seek them out in others in order to look better in comparison.

Pronin et al. (in press) admit that motivational factors probably play a role in people's claims of less-than-average susceptibility to bias, but they argue that these factors are likely to operate alongside with purely cognitive, non-motivational mechanisms of the human mind to produce the asymmetry effect.

Building on the classic work of Ichheiser (1949), Ross and Ward (1996) have offered an explanation for this asymmetry in people's ability to recognize cognitive biases in themselves and others. They argue that although some aspects of the human egocentrism disappear in the course of experience and maturation, the process is never completed. Despite their ability to recognize that other individuals see the world from different perspectives than themselves, people fail to acknowledge the extent to which the other perspectives can differ from their own point of view. In other words, people make insufficient allowance for the construal differences between themselves and others. Ross and Ward term this worldview "naïve realism". Lay epistemology of naïve realism is based on three tenets:

1. People believe that they see entities and events "as they are in objective reality", and that their attitudes, beliefs, and preferences follow from a direct and unbiased apprehension of this reality.
2. Other social perceivers, who have access to the same information and who have processed this information in a rational manner, should share the naïve realist's reactions, behaviors, and opinions. This belief, for example, results in the false consensus effect (Marks & Miller, 1987; Ross, Green, & House, 1977).
3. The failure of other social perceivers to do so arises from one of three possible sources: (a) the individual or group in question might have been exposed to a different source of information than the naïve realist; (b) the individual or group in question may be unable or unwilling to process the information in a rational and normative way; and (c) the individual or group in question may be biased (either in interpreting the evidence or proceeding from evidence to conclusions) by ideology, self interest, or some other distorting personal influence.

Two mechanisms seem to account for the line of reasoning behind the tenets of naïve realism. First, research has shown that people are not very good at introspection and tend to rely on a priori causal theories rather than true introspection when reporting on their cognitive processes (Nisbett & Wilson, 1977a; Nisbett & Ross, 1980, Chapter 9). This tendency may largely account for the first tenet or the starting point of naïve realism. People simply lack the necessary cognitive ability to detect biases in their thinking. Moreover, people not only fail to find evidence of cognitive bias in their reasoning, but they also do find evidence of their efforts to avoid such a bias. Pronin et al. (in press) refer to this phenomenon as "introspection illusion" because it leads people to mistakenly perceive the results of their own introspection as complete and valid.

When others fail to meet the predictions of the second tenet, the social perceiver is forced to seek explanation for their behavior. The third tenet offers three possible explanations: the first deals with the situational constraints of the person or group in question (i.e., lack of information), and the other two explain the behavior in the terms of that person's or group's dispositional qualities. This is where the second mechanism--correspondence bias (Gilbert & Malone, 1995) or the fundamental attribution error (Ross, 1977)--plays its role in converting the naïve realist into a "naïve cynic". Social perceivers are more inclined to infer stable traits and dispositions from an actor's behavior while overlooking the influence of situational factors (Jones, 1979; Jones & Davis, 1965; Jones & Nisbett, 1972). Repeated observation of behavior may lead the naïve realist to make dispositional inferences about particular people or groups, to make generalizations, and with the time—to develop lay theories about people's tendency to commit various types of cognitive biases and to be self-serving in their inferences. At the same time, as research reported by Pronin et al. (in press) demonstrates, even equipped with such theories of general human behavior, people continue to apply them to others, but continue to rely on introspection when analyzing their own cognitive processes.

Naïve realism has a number of consequences in social behavior. For example, it leads people to assume that others are less susceptible to mass media appeals and propaganda—a much-researched phenomenon also known as the "third person effect" (Cohen, Mutz, Price, & Gunther, 1988; Davison, 1983; Duck & Mullin, 1995). Another media-related phenomenon stemming from the naïve realism is the "hostile media effect" (Vallone, Ross, & Lepper, 1985)—partisans' tendency to see media reports as biased in favor of the opposite side in depictions of controversial issues. Yet another consequence is the "illusion of asymmetric insight" (Pronin, Kruger, Savitsky, & Ross, 2002)—people's tendency to assume that they know their peers better than their peers know them. However, from the point of view of social interaction the most important consequence is the so-called "false polarization effect".

The False Polarization Effect

Pronin, Puccio, and Ross (2002) define the false polarization effect as an overestimation of the construal gap between the modal views of two sides in a controversy or conflict and an underestimation of the amount of common ground that could serve as a basis for reconciliation and constructive action. Robinson, Keltner, Ward, and Ross (1995), and Keltner and Robinson (1997) have carried out systematic studies of the false polarization effect. Robinson et al. (1995) asked pro-choice and pro-life partisans in the abortion rights debate to respond to a number of questions and to indicate the responses that they believed the average in-group and out-group member would give to the same questions. In another study, the same authors asked political liberals and

conservatives a series of questions about a controversial criminal case and the outcomes of the following trial, and they asked them to predict the responses of their political allies and opponents. Both studies showed that although the opposing groups indeed differed in their opinions about the issues in question, the assumed differences were much larger than the actual differences between the groups. Both groups overestimated the extremity of opinions of both their in-group and out-group members. Similarly to other studies reporting asymmetric perception of bias, the participants indicated that their own views and assumptions were generally less shaped by political ideology than by objective or rational pragmatic concerns compared with the views of their peers.

Keltner and Robinson (1997) asked "traditionalist" and "revisionist" university professors to express their opinions related to the so-called Western canon dispute (and to estimate the responses of their peers representing the same and the opposite side in the dispute). They also asked the participants to select a list of books that they would use in teaching an undergraduate course of English literature, and another list of books that would represent the average preferences of the other side. Both groups overestimated the extremity of both sides' opinions (but especially the traditionalist attitudes of the traditionalists). Both groups (but especially traditionalists) also underestimated the overlap of both sides' book selections.

The false polarization effect has been found in several other studies across a number of contexts. Dawes, Singer, and Lemons (1972) asked supporters and opponents of the Vietnam War to write statements that would be endorsed by members of one or the other group. When these statements where presented to the corresponding groups, both supporters and opponents rejected more statements written by their out-group members than by in-group members. The most often given reason for rejection was that these statements were too extreme. Diekman, Eagly, and Kulesa (2002) found that male and female participants overestimated male opposition to pro-female attitude statements. Monin and Norton (2003) found that college student "bathers" and "non-bathers" during a shower ban overestimated the gap between both groups in their concern for the community. In a study by Robinson and Friedman (1995), union and management negotiators overestimated the extremity of the other sides' position in a labor dispute situation. In a study of "adversary's extremity bias" by Rouhana, O'Dwyer, and Morrison Vaso (1997), Israeli and Arab students exaggerated the political views of the other party.

While there is little doubt that the false polarization effect is a consequence of naïve realism and asymmetric bias perception, Pronin, Puccio, and Ross (2002) and Puccio (2003) point at another, more social source of the phenomenon. During contentious discussions, many individuals remain silent. Those who do express their opinions are often the ones with the strongest conviction. Even if the speakers do have some doubts, they hesitate to reveal them. The same is generally true about wider-scale political debates in mass

media. As a result, the social perceivers (partisans involved in the debate and neutral observers alike) are exposed to biased samples of information representing the parties' opinions. The error committed by the social perceivers in such situation can be twofold. First, it is erroneous to assume that the information expressed by the partisans always represent their true opinions. As the classic study of Jones and Harris (1967) demonstrated, even when the situational constraints are highly salient, people tend to make dispositional attributions, assuming that the observable behavior corresponds to the true attitudes (see also Jones, Worchel, Goethals, & Grumet, 1971). People are even more likely to commit such correspondence bias where the situational constraints are not obvious, or where such constraints (e.g., an actor's position as a figurehead for a political faction) can be interpreted as motivational factors contributing to the extremity of the actor's position, or to the lack of ambiguity in an actor's beliefs. Second, it is erroneous to assume that the biased sample of opinions represents an accurate reflection of the average opinion of the larger population of partisans in question. The most extreme opinions in social controversies are often also the most salient, and sometimes, the only ones available. Thus, when the social situation calls for an estimation of the average opinion of a partisan group, people tend to rely on the availability heuristic (Taylor, 1982; Tversky & Kahneman, 1973) in their predictions. Although these errors committed by the social perceiver result directly from biases in social perception, they are preceded or at least facilitated by biases in social presentation.

Puccio (2003) tested the social presentation hypothesis in five studies designed to reduce the false polarization effect. She asked the participants in different experimental settings to estimate the position of partisans who had expressed the arguments supporting their own side, the best arguments for the other side, or arguments for both sides. The results showed that the position of partisans who had expressed arguments favoring the other side (Studies 1 and 3) or both sides (Studies 2, 3, 4, and 5) were judged as significantly less extreme than the position of partisans providing arguments only favoring their own side. Expressing arguments for the other side reduced the false polarization effect and led the participants to perceive the other side as more open-minded and increased the participant's optimism about reaching agreement in a situation of negotiations. The results demonstrate that the presentation of social information indeed influences the false polarization effect. The findings also lead Puccio to conclude that ironically the same correspondent inference bias may increase (when only parties' own positions are communicated) or reduce (when the parties are encouraged to express the opponent's position) the false polarization. In the latter case, the observers tend to overlook the situational constraints put on the actors (experimenter's demand, or, presumably, in real-life situations-- mediator's instructions) and tend to perceive the opponents as more open-minded and less extreme.

Social Categorization and Social Identity

Besides the cognitive biases discussed above, social categorization is perhaps the strongest factor influencing how social perceivers see reality. Keltner and Robinson (1993) found that negotiators unaware of their ideological differences reached agreement faster and rated their partner as more constructive and cooperative than negotiators aware of the differences or negotiators sharing the same ideological views who had been led to believe that their views were different. In three studies of "reactive devaluation", Maoz, Ward, Katz, and Ross (2002) found that Israeli Arabs and Jews derogated proposals attributed to the other side of the conflict even when those proposals were developed by representatives of their own side.

Social category membership adds a strong evaluative element to social judgements. Tajfel (1978) demonstrated that out of their need for self-enhancement, people have a strong motivation to identify with social groups they belong to (in-groups), and to compare these groups favorably to other groups (out-groups). The social identity theory (SIT) is a motivational theory in its basic assumptions. Tajfel and Turner (1979) start from the assertion that individuals strive to maintain or enhance their self-esteem to achieve a positive self-concept. Social groups or categories in any society are associated with positive or negative value connotations. Consequently, social identity (defined as "those aspects of an individual's self-image that derive from the social categories to which he perceives himself as belonging" [p. 40]) can be positive or negative according to the evaluations of those groups, which tend to be socially consensual within or across groups. The evaluation of one's in-group is determined with reference to specific out-groups through social comparisons in terms of value-laden attributes.

The SIT in its condensed form rests on three principles (Tajfel & Turner, 1979). First, individuals strive to achieve and maintain positive social identity. Second, positive social identity is largely based on favorable comparisons between in-group and relevant out-groups. Third, when social identity is unsatisfactory, individuals will strive either to leave their group and join some more positively distinct group, or to make their existing group more positively distinct. The former solution is an individual strategy, designed to change the self-concept of individual, leaving the (low) group status intact. The latter solution can take the form of social creativity (finding a new dimension for intergroup comparison, changing the values associated with comparison dimensions, or finding other, lower status groups for comparison) or direct competition with the out-group. Over two decades of research within the traditions of social identity theory and self-categorization theory have provided plenty of evidence to these principles (Brewer & Brown, 1998; Brown, 2000; Hogg & Abrams, 1999).

This motivation for favorable between-group comparison results in a general tendency to evaluate one's membership group (in-group) more positively than any non-membership group (out-group)–in-group favoritism or in-group bias (Hewstone, Rubin, & Willis, 2002). Of course, because people belong to numerous social categories at the same time, it depends on the particular situational context which social category memberships become most salient and are activated, thus affecting people's judgements about the members of other social categories (Turner et al., 1987). According to Turner et al. (1987), the salience of social categories is determined by the meta-contrast ratio: the ratio of the perceived differences between categories to the perceived differences within categories. The higher the ratio (that is, categories are seen as homogeneous and different from each other), the more salient the social self-categorizations.

Research has identified a number of factors that facilitate in-group favoritism. Such factors include the following: level of identification with (or commitment to) in-group (Branscombe & Wann, 1994; Verkuyten & Nekuee, 1999); the context in which the comparison is made (Kinket & Verkuyten, 1999); group size and status differences (Mullen, Brown, & Smith, 1992; Sachdev & Bourhis, 1987); the perceived threat to the in-group's interests (Ellemers, Spears, & Doosje, 2002; Esses, Jackson, & Armstrong, 1998); the social self-esteem of the group members (Rubin & Hewstone, 1998); and individual difference variables like authoritarianism (Altemeyer, 1998), religious beliefs (Batson & Burris, 1994), and system of values (Biernat, Vescio, & Theno, 1996). An important factor influencing in-group favoritism is the way it is expressed (or the dimension on which it is measured). Studies have shown that in-group favoritism is usually not found when it is framed as out-group derogation rather than in-group enhancement (Brewer, 1999, 2001; Mummendey, 1995; Otten & Mummendey, 2000). However, some studies of trait evaluations have found in-group favoritism on both positive and negative dimensions (Reynolds, Turner, & Haslam, 2000; Rustemli, Mertan, & Cifci, 2000). Reynolds et al. (2000) suggest that when the comparative dimensions are important and relevant to in-group-out-group categorization, in-group favoritism can be expressed to the same extent on both positive and negative dimensions. The relationships between the mentioned factors are complex and result in numerous interactions between them. However, the existing research points to the robustness of the general tendency to see one's own group in a more favorable light than other groups.

Stereotypes

Social categorization is helpful for social interaction because it tells people what to expect from others. It helps one infer motives, predict the reasoning and behavior of others and thus helps them to reduce the uncertainty

of the social environment. The simplified social knowledge about categories is available in the form of stereotypes (Fiske, 1998; Tajfel, 1981b). Hilton and Von Hippel (1996) define stereotypes as "beliefs about the characteristics, attributes, and behaviors of members of certain groups" (p. 240). These beliefs can have two functions and two sources.

First, stereotypes operate like social schemas, allowing to process information about others more quickly and efficiently by simplifying the demands the situation makes on the perceiver (Fiske & Taylor, 1991). As social schemas, stereotypes contain highly abstract and generalized information about members of various social groups. Largely based on and maintained by individual experience, social schemas can reflect (with acceptable accuracy) real characteristics of groups and actual intergroup differences. Schematic processing is necessarily simplistic and may cause the social perceivers to make occasional mistakes. In the long run, however, such schemas have some adaptive value for the social perceiver because most of the time they lead to adequate, if not optimal, inferences about members of a stereotyped group. The need for speed and efficiency, however, does not mean that people use schematic processing automatically and unconditionally (Brewer & Feinstein, 1999; Fiske, Lin, & Neuberg, 1999; Fiske and Neuberg, 1990). Instead, people are "motivated tacticians" who have multiple cognitive strategies available and are able to choose between them based on their goals, motives, and needs (Fiske, 1992; Fiske & Taylor, 1991). This view of the social perceiver suggests that people choose between schematic (or theory-driven, top-down) information processing and data-driven (or bottom-up) information processing depending on their motivation. This means that where the expected costs of making a mistake is low, it can be rational to rely on stereotypes in making social inferences.

The other function (and source) of stereotyping is more social than cognitive in nature. Hilton and Von Hippel (1996) state that stereotypes are more than just beliefs about groups; they are also theories about why certain attributes go together. In this sense, stereotypes can be independent of the real-life group differences. Tajfel (1981b) identified two social functions of stereotypes: (1) creation and maintenance of group ideologies that help to explain or justify a variety of social actions; (2) helping to preserve or cerate positively valued differentiations between one's own and other social groups. This interpretation of stereotypes is similar to Allport's interpretation (1958): "*a stereotype is an exaggerated belief associated with a category. Its function is to justify (rationalize) our conduct in relation to that category*" (p. 187, italics in original). By "our conduct", Allport meant primarily individual social perceivers holding stereotypes of various categories, whereas Tajfel was interested more in the "group ideologies" affecting intergroup behavior and intergroup perception. Tajfel was interested primarily in social stereotypes–beliefs that are shared by large numbers of people within social groups or entities. An important function of social stereotypes is providing explanations of social issues and intergroup

differences. These explanations, in turn, form a basis for a value-laden differentiation between social groups and justification of intergroup behavior. In other words, stereotypes reflect the existing social structure of a given society and the existing rules for social mobility and change. Stereotypes as social schemas are more or less descriptive in nature; stereotypes as shared beliefs are largely prescriptive and normative. Stangor and Schaller (1996) express the same idea; they argue that stereotypes are maintained as norms in a given society, transmitted by language and mass media.

The social nature of shared stereotypes has prompted some researchers to interpret and study them as social representations (Hewstone, Jaspars, & Lalljee, 1982). Moscovici (1981, 1984, 1988) brought the concept of social representations into social psychology in 1970s. Moscovici (1981) defines social representations as "a set of concepts, statements and explanations originating in daily life in the course of inter-individual communications. They are equivalent, in our society, of the myths and belief systems in traditional societies; they might even be said to be the contemporary version of common sense" (p. 181), and elsewhere (1984) as "social reality *sui generis*" (p. 13, italics in original). Social representations range from generalized structures shared by all society members and entire nations to highly specific knowledge structures shared by separate subgroups of society (Moscovici, 1988). An important function of social representations is to conventionalize objects, persons, and events by locating them within a context of familiar, established categories. It is clear from Moscovici's (1981) definition given above that representations not only reflect the way society thinks, but representations are also prescriptive in nature, imposing themselves on the cognitive activity of individuals. According to Moscovici (1984), "representations are prescriptive, that is they impose themselves upon us with an irresistible force" (p. 9). Individuals are often unaware of this influence and remain unaware of the social determination of their thought, preferring to view their reasoning as "common sense". According to Augoustinos and Walker (1995), stereotypes are social representations because they do not simply exist in individuals heads, but they are socially and discursively constructed in the course of everyday communication. Moreover, once established (or "objectified" to use a term from social representation theory) they assume an independent and often prescriptive reality.

Gardner (1994) reviews a number of studies demonstrating the power of shared stereotypes in communication. The cited evidence demonstrates that groups are very efficient in guessing the stereotyped target group from a given small number (as few as two) of stereotypical traits, provided that these stereotypes are shared within the group. Gardner concludes that when the stereotypes are consensual, individuals can understand them regardless of whether or not they subscribe to them, and they are able to communicate through the stereotypes alone. These findings also demonstrate that regardless

of whether social stereotypes reflect real group differences or not, they indeed represent a social reality of its own, shared and communicated by members of the same group. Experimental evidence also shows that stereotypes that are learned from others tend to be more extreme, contain less variability information, and have higher social consensus than stereotypes learned directly from a contact with the stereotyped group (Thomson, Judd, & Park, 1999).

Social categorization also tells the person how to act in specific situations. Group norms largely determine what attitudes a group member is supposed to hold regarding different phenomena, and how the group member is supposed to react to various social stimuli (Abrams, Wetherell, Cochrane, Hogg, & Turner, 2001; Turner, 1991; Turner & Oakes, 1989). Group members endorse shared stereotypes through mutual influence because they reflect shared in-group norms. Self-categorization theory asserts that an important consequence of individuals perceiving themselves in terms of social identity is that they agree with others belonging to the same social category, and they expect these others to agree with them. In-group members are perceived as sources that can consensually validate individual's beliefs. When a person's social identity is salient, in-group opinions perform a regulatory function as a source of normative information against which individual observations are evaluated and validated. As a result, group members tend to adopt the beliefs shared by other group members to form a shared reality. According to Haslam (1997), if there is a disagreement between an individual and other in-group members (a particular belief appears not to be shared), individuals will (a) modify their belief so that it becomes consistent with other group members, (b) attribute the disagreement to differences in the stimulus perception, or (c) re-categorize the former in-group members as an out-group in order to reduce subjective uncertainty that arises from disagreement with people with whom they expect to agree.

The prescriptive nature of social stereotypes has important consequences for the social interaction and intergroup relations. Stereotypes provide an abstract picture of the stereotyped group and instruct the social perceiver how to behave towards the members of that group. In this way, social stereotypes can become self-fulfilling prophecies, both on interpersonal and intergroup level. Self-fulfilling prophecies emerge when people's stereotypes make them (consciously or unconsciously) alter their behavior in such ways that they provoke the expected, stereotype-consistent behavior from the people who are targets of the stereotypical expectancies. Experimental evidence has demonstrated the power of self-fulfilling prophecies in interpersonal communication (Snyder, Tanke, & Berscheid, 1977; Word, Zanna, & Cooper, 1974). One may speculate that the same effect happens on intergroup level with wider societal consequences.

The shared and pervasive nature of social stereotypes has for many decades urged psychologists to discuss the question of stereotype accuracy.

Judd and Park (1993) admit that defining a criterion against which stereotypes could be evaluated as accurate or inaccurate is a "thorny issue" because of the difficulty of objective measurement of a group's location on the relevant attributes. The most frequently used criterion involves self-report evaluations by individuals sampled from the stereotyped group. However, such a method is affected by sample selection biases (in cases of larger populations) and self-report biases (such as social desirability or different interpretations of the attributes in question). Another method might involve collection of objective or behavioral measurements of the stereotyped group. With this method, the problem is finding valid ways to objectively operationalize the attributes in question; for many variables, such operationalization can be virtually impossible. Yet another possible criterion is using judgements of experts who have an extensive knowledge of the stereotyped group. However, such judgements are subject to stereotypes held by the experts themselves. Judd and Park (1993) conclude that stereotype accuracy can be defined only by using a number of criteria and by controlling for the possible biasing factors influencing each criterion. They also stress that the accuracy question concerns the correspondence between individuals' personal beliefs and the true attributes of the group as opposed to the correspondence between knowledge of (shared) cultural stereotypes of a group and the true attributes of the group.

Oakes and Reynolds (1997) emphasize the difficulty of distinguishing the latter two concepts in practice. They argue that social values, which serve as premises for the perceived accuracy of stereotypes, cannot be deemed objectively true or untrue, but rather can be either validated or challenged through processes of social influence and political action. Aggregation of individual-level data does not represent a "true" group-level score because on the group level stereotypes have a different sort of validity that reflects and explains the existing intergroup relationships of the moment. Stereotypes are held about groups as social categories and not groups as aggregations of individuals. Therefore, the "true score" established on one (individual measurement) level cannot serve as criterion for evaluating accuracy on the other level (group perception). Rather, the validity of stereotypes is established at a higher level of social categorization (individual beliefs are validated at the level of group and group beliefs at the level of society). Agreement between two groups on the stereotype of one of them does not necessarily reflect a "kernel of truth" in the stereotype–it may simply reflect the fact that the stereotyped group has accepted the image held by the stereotyping group.

In real life, the cognitive and social aspects of stereotyping are hard to distinguish. Although social stereotypes constitute a reality of their own, they are maintained by individuals, and they manifest themselves through the behavior of individuals. Stereotypes of other groups are not activated and applied unconditionally. First, depending on the situational context, the stereotyped group can be categorized either as out-group, or as a part of a larger,

more inclusive in-group that will influence which stereotypes will be activated (Haslam, Turner, Oakes, McGarty, & Hayes, 1992). Second, individual cognitive processes influence activation of stereotypes. One interesting line of research concerns the automaticity of stereotype activation. Devine (1989) has argued that automaticity develops in social stereotype activation at a young age. However, as people grow older and start to reflect on their own beliefs, they learn to suppress the automatically activated stereotypes and consciously activate more favorable or politically correct beliefs associated with the stereotyped category. This altering of the stereotypic cognitions depends on individual's attitude (prejudice) towards the stereotyped group. Only those low on prejudice (without a negative attitude towards the out-group in question) exert the necessary effort to negate the stereotypes. According to Devine (1989) and Devine and Monteith (1999), stereotypes and personal beliefs are conceptually distinct cognitive structures, each representing a part of individual's entire knowledge base about a particular group. Both concepts may overlap, but they are not identical. All persons belonging to a particular culture should be equally aware of certain cultural stereotypes, but they may differ in their level of prejudice and thus the endorsement of the stereotypes. Hilton and Von Hippel (1996) cite a number of studies confirming Devine's theory. Augoustinos, Ahrens, and Innes (1994) found that participants high in prejudice and low in prejudice could equally list the elements of social stereotypes. However, high-prejudice participants tended to give unqualified pejorative responses, and low-prejudice participants qualified their responses and tended to distance themselves from these responses.

Intergroup Attributions

One area where social categorization and stereotypes have a particularly strong influence is intergroup attribution. Hewstone (1989) defines intergroup attribution as "the ways members of different social groups explain the behavior (as well as the outcomes and consequences of behavior) of members of their own and other social groups" (p. 166). The reason that stereotypes play such a significant role in intergroup attributions is that the social perceiver attributes the behavior of another person not simply to individual characteristics of this person or the situational context but to characteristics associated with the group to which the other person belongs. Moreover, the social perceiver is also a member of a social group, and the norms of that group largely prescribe how the behavior of particular out-groups should be explained. The salience of social categories thus makes intergroup attribution different from interpersonal attribution.

Although the first attempts to apply classical attribution theories to intergroup context appeared in the 1970s, social psychologists interested in intergroup relations had discussed some aspects of group attribution already in

the 1960s. Sherif (1966, pp. 109-110) noted that the assignment of blame in conflict situations is conducted almost entirely from the in-group's point of view. Campbell (1967, p. 825) saw "erroneous causal perception" as one of the undesirable aspects of stereotypes, noting the tendency of white perceivers to attribute racial group differences in education to race rather than environmental background. Tajfel (1969) hypothesized that intergroup attributions would be simpler and more predictable than interpersonal attributions. Following Jones and Davis (1965), Tajfel predicted that the social perceiver's preference for person attributions rather than situation attributions would be extrapolated to group perception. That is, social groups would be personalized, and the explanations would focus on inherent group characteristics rather than transient situational factors.

The earliest demonstration of intergroup attribution was a paper by Taylor and Jaggi (1974) who examined explanations of positive and negative behaviors by in-group and out-group members among Hindu office workers in India. The participants read a short description of a behavior and chose an explanation from a list of possible internal or external causes for the behavior. Taylor and Jaggi found that for socially desirable behaviors the participants made more internal attributions for in-group than out-group actors; for undesirable behaviors, there were fewer internal attributions for in-group than for out-group actors. The results were largely replicated in the United States by Wang and McKillip (reviewed in Pettigrew, 2001). An interesting additional finding in their study was that the more ethnocentric American adult and Chinese college student samples in their study displayed the group serving attributional bias (in opposite directions, as expected), but the less ethnocentric American college students sample did not display such a bias. A decade later, Hewstone and Ward (1985) replicated and extended the results with Malay and Chinese samples in Malaysia and Singapore. They found a clear evidence of ethnocentric attribution, but the effect of in-group favoritism was far stronger than that of out-group derogation. The results of one study also showed that the minority group (Chinese) in fact made out-group favoring attributions, leading the researchers to conclude that ethnocentric attribution is not a universal tendency and can be reversed for low status groups, reflecting the political and cultural background.

Of the intergroup attribution studies published in 1970s, Duncan (1976) offers perhaps the most illustrative evidence of how actor's group membership affects attributions of behavior. Duncan showed white participants a short film where one person shoved the other following an argument. The race of the two actors was varied in a way that a white actor pushed either a white actor or a black actor, or a black actor pushed either a white actor or a black actor. The participants then were asked to categorize the behavior and to rate the extent to which the behavior had internal or external causes. There was a very clear tendency in the results that when the actor delivering the shove was black, his

23

behavior was categorized as more aggressive than when the actor was white; this tendency was especially strong when the actor receiving the shove was white. There was a strong pattern to attribute the shove to dispositional factors when the actor was black and to situational factors when the actor was white. Duncan assumed that the presence of the black actor automatically activated the stereotype that associates black people with violence; once activated, the stereotype biased the observers' interpretation of the events by providing automatic and simple stereotype-consistent explanation.

The evidence of social stereotypes biasing intergroup attribution is not limited to studies of racial or ethnic groups. Deaux and Emswiller (1974) found that on a stereotypically masculine task a male actor's success was more likely to be attributed to ability than a female actor's success on the same task. On a stereotypically female task, no differences were found in attributions for male and female success.

In the light of these findings, in 1979 Pettigrew (2001) suggested a model that allowed making systematic predictions for out-group derogating attributions in intergroup context. Building on Ross's (1977) "fundamental attribution error", Pettigrew tried to explain how intergroup misattributions can help individuals to maintain their negative stereotypes about prejudiced groups. Pettigrew reasoned that if an out-group member were seen performing a negative act consistent with the negative stereotype of that group, the tendency to make internal attributions would be enhanced. When an out-group member was seen performing a single positive act, the perceiver would be motivated to "explain it away", so that the original stereotype needs not to be changed. Depending on the perceiver's interpretation of the situation, the positive behavior can be explained as (a) an exceptional case, (b) actor's special advantage or luck, (c) actor's high motivation and effort, or (d) manipulated situational context. Pettigrew predicted that the "ultimate attribution error" would be sharper for prejudiced individuals, more likely to occur when the group memberships are salient, and it would be stronger when the groups in question have histories of intense conflict and possess negative stereotypes of each other. Furthermore, the ultimate attribution error would be more likely to occur when racial and ethnic differences covary with national and socioeconomic differences.

In his review of 19 studies of intergroup attribution conducted over two decades, Hewstone (1990) reports limited support to Pettigrew's predictions. Hewstone differentiates between categorization effects and outcome effects in intergroup attribution. Categorization effects compare the attributions made for in-group and out-group members separately for each type of behavior. Outcome effects compare the attributions made for different types of behavior (e.g., positive/negative; success/failure) separately for in-group and out-group actors. Hewstone concludes that there is a tendency to attribute negative out-group behavior to personal causes; however, he notes that this tendency more often

takes the form of a categorization effect (more internal attribution for a negative act by out-group than an in-group member) rather than an outcome effect for out-group actors (e.g., more internal attributions for negative than for positive acts). There is also a tendency for out-group failure to be attributed more to lack of ability than is for in-group failure. Regarding explanation of positive out-group behavior, the predictions have not been systematically tested because none of the reviewed studies examined all four possible explanations suggested by Pettigrew. There is some evidence that out-group success is explained away in terms of luck, high effort, or ease of task. Sometimes less internal attributions have been found for positive out-group behavior than positive in-group behavior. The research evidence supports the prediction that group-serving attributions may be stronger when perceivers are aware of their own and the actor's group memberships. Group serving attributions indeed vary across situations and may be stronger when the groups have histories of conflict and possess negative stereotypes of each other; however, the bias can disappear or even be reversed for members of low-status groups. Although the tendency for group-serving attributions may be stronger for prejudiced individuals, the effect has not always been bound. More recent studies have provided further support for the general tendency for group-serving attributions (Austers, 2002a; Hunter, Stringer, & Watson, 1991; Jackson, Sullivan, & Hodge, 1993), but also demonstrated the context-dependence of these effects (Islam & Hewstone, 1993).

Despite the long tradition of intergroup attribution research and the clear evidence of how social categorization and stereotyping affects group attributions, there have been relatively few studies examining the effects of group perspective taking on intergroup attributions. Consequently, not much evidence is accumulated about the false polarization effect in intergroup attributions, although the otherwise robust nature of the false polarization suggests that it is likely to appear in causal explanations as well. Kemdal and Montgomery (2001) found that animal experimenters and animal rights activists could take each other's perspectives, resulting in a reversed actor-observer effect. Only the animal rights supporters showed some false polarization when guessing the ratings of internal causes for activists' behavior from the researchers' perspective. Austers (2002a) asked ethnic Latvian and Russian schoolteachers to rate positive and negative behaviors from their own and out-group perspective on a number of attributional dimensions. The results indicated some false polarization effect in predicting out-group's responses but not consistently. Using a free response format, Austers and Montgomery (2001) asked their participants to explain an out-group behavior from their own and three imagined perspectives: an out-group, a neutral observer, and a "reflected out-group perspective". That is, they were asked to explain how an out-group member would explain a behavior from the participant's in-group perspective. From a reflected out-group perspective, participants used more personality-

related explanations than from a self perspective, and the tendency was reversed for situational explanations indicating that the participants had expected an ideological bias to influence the explanations by an out-group member. However, this only indicates an expected false polarization on behalf of an out-group rather than an actual bias. It should be noted, however, that in none of these studies direct statistical tests were performed to check for the false polarization effect, and the observations reported above are based on general patterns of means. In addition, in all of the reported studies, the participants were asked to take an out-group perspective but not an in-group perspective. Robinson et al. (1995) tested some aspects of intergroup attribution when they asked their participants to rate the basis of their own political judgments and that of their in-group and out-group members. The respondents indicated that they personally had been less influenced by ideology or political orientation than either their peers or opponents. These findings suggest a presence of the false polarization effect in the context of intergroup attributions.

The results of interpersonal or intergroup attribution studies largely depend on the methodology used for categorizing the attributions. The oldest and most widely used classification of attributions, based on Heider's (1958) model of interpersonal attributions, distinguishes between internal (dispositional) and external (situational) attributions. This classification has been criticized on both theoretical and practical methodological grounds (Miller, Smith, & Uleman, 1981; Ross, 1977). Islam and Hewstone (1993) proposed an alternative four-dimensional model, measuring attributions along the continuums of causal locus, stability, controllability (from Weiner, 1986), and globality (from Abramson, Seligman, & Teasdale, 1978). Kruglanski (1975) has stressed the necessity to differentiate between endogenous and exogenous attributions, where an endogenously attributed action is an end in itself and an exogenously attributed action is one that mediates a further goal. Kruglanski (1979) also distinguishes between causal explanations (accounting for what caused an act) and teleological explanations (accounting for why the act was accomplished), and between actions (actor's intentional behavior) and occurrences (behavior independent of the actor's will). Similarly, Buss (1979) suggested categorizing attributions into causes and reasons. Locke and Pennington (1982) further elaborated this division, interpreting reasons as one type of internal causes of behavior, which can further be divided into situational reasons and psychological reasons.

More recently, Malle and colleagues (Malle, 1999, 2001; Malle & Knobe, 1997; Malle, Knobe, O'Laughlin, Pearce, & Nelson, 2000) have attempted to integrate the previous efforts at categorizing attributions into a new framework. The main idea behind their model of folk explanations of behavior (FEB) is that rather than concentrating on the explicit contents of causal attributions, people's behavior explanations must instead be considered in terms of their conceptual structure, the network of concepts and assumptions on which these explanations

are based. Central to this structure is the folk concept of intentionality (Malle, 2001; Malle & Knobe, 1997). According to the classification suggested by Malle (1999), explanations that interpret an agent's behavior as unintentional are considered cause explanations. Explanations that interpret the behavior as intentional are divided into three major groups: (a) causal history of reasons (factors from agent's personal history that cause the intentional behavior without the agent being aware of them); (b) reasons (factors that the agent considered when forming the intention to act); and (c) enabling factors (factors that clarify how it was possible that the agent completed the intended action). Reasons are further divided into three types: desires, beliefs, and valuings. A distinction is made between reasons marked with a mental state description and those unmarked. In addition, the coding scheme developed by Malle (2000) uses a number of secondary codes to reflect the contents of each explanation. The FEB coding scheme has been used to examine the differences between how individuals explain individual and group behavior (O'Laughlin & Malle, 2002), but it does not seem to have been applied in studies of intergroup attributions.

Aims of the Thesis

The general aim of the thesis is to explore how the false polarization effect influences perception and interpretation of the shared social reality when two groups take the perspective of each other. This thesis examines what can or should be understood with respect to shared social reality in intergroup perception. All four studies address this issue in various ways to find common patterns of agreement and disagreement in intergroup perception. Study 1 and Study 4 are designed to test the effects of shared reality more directly. In addition, this thesis tests the robustness of the false polarization effect in various previously unexplored contexts, focusing on intergroup attributions. Study 1 looks at the false polarization effect in intergroup ratings of personality traits. Study 2 examines the false polarization in intergroup ratings on a number of attributional dimensions. Study 3 explores the influence of the false polarization effect on causal explanations given in a free response format. Finally, Study 4 checks the possible influences of the false polarization effect on perception of causal explanations previously provided from own perspective and imagined group perspectives.

The false polarization effect is a complex phenomenon because it is hypothesized to be a result of three interacting factors: asymmetric perception of bias, social categorization, and biased presentation of information. Of the three mentioned factors, the effects of social categorization are examined more deeply. Social categorization through stereotypes associated with each social category tells the social perceiver when to expect group-serving effects in groups' opinions and in which direction these effects are supposed to occur. One question addressed in this thesis is to what extent the false polarization effect

coincides with social categorization effects. That is, does the false polarization effect happen only when the group norms clearly state which responses are group serving and which are not? This question is addressed most directly in Study 1, which looks separately at categorization effect and false polarization effect in descriptive and evaluative aspects of intergroup ratings. Study 2 provides some additional information on this topic by using a number of attributional dimensions where categorization effects cannot be predicted in advance. In all the reported studies, to elicit the categorization effects influencing group judgements more efficiently, groups were asked to evaluate two targets (in-group target and out-group target).

In addition, this thesis tests a promising method for coding intergroup attributions–the Folk Explanations of Behavior coding scheme that has been developed and mostly used in interpersonal contexts. The Folk Explanations of Behavior coding scheme is applied in Study 3. Study 2 provides some additional information about how some of the concepts underlying the coding categories are perceived in intergroup settings.

OVERVIEW OF THE EMPIRICAL STUDIES

Study 1. Differentiating In-group Favoritism from Shared Reality in Intergroup Perception

This study explores the effects of two factors influencing mutual ratings of social groups: in-group favoritism and shared reality. In-group favoritism is a consequence of salient social categorization. In line with the social identity theory (Tajfel & Turner, 1979) and self-categorization theory (Turner et al., 1987), in-group favoritism should influence intergroup ratings in a way that allows each group to see itself more favorably than the other group. In other words, in-group favoritism should influence the evaluative aspects of intergroup ratings. Shared reality, on the other hand, can be defined as the shared stereotypes of both groups; that is, the characteristics of both groups that they agree about and that presumably have been verified on a higher (i.e., societal) level of social categorization (Oakes & Reynolds, 1997). Because of in-group favoritism, the shared stereotypes are likely to be expressed differently by both groups. The same descriptive attribute (e.g., tendency to save money) can be seen either as a positive trait (thrifty) or a negative one (stingy), depending on whether seen from a group's inside perspective (i.e., by the group members themselves) or an outside perspective (i.e., by members of the out-group) (Montgomery, 1994). Peabody (1968, 1985) separates both aspects of intergroup ratings. This study applied Peabody's method in two intergroup settings. We predicted that each group would show in-group favoritism on the evaluative aspects of ratings, but both groups would agree on the descriptive ratings of each other. We also predicted that the groups would show false

polarization on the evaluative dimension, but they would not overestimate the group differences on the descriptive dimension.

Method

In Experiment 1, 86 Latvian and 77 Russian university students rated two targets–a Latvian living in Latvia and a Russian living in Latvia–on 13 pairs of scales designed to differentiate between the evaluative and descriptive aspects of each rating. For each scale pair, we calculated an evaluative score as the mean of ratings on both scales and a descriptive score as the difference of ratings on both scales divided by two. We then conducted a 2 (judge ethnicity: Latvian vs. Russian) by 2 (target ethnicity: Latvian vs. Russian) by 13 (scale pair) mixed ANOVA with repeated measures on the last two factors separately for the evaluative and descriptive scores.

In Experiment 2, 44 female and 33 male psychology students rated two targets–a male psychology student at Stockholm University and a female psychology student at Stockholm University–on 8 pairs of scales. Each participant rated both targets from his or her own perspective as well as from the perspectives of a male psychology student and a female psychology student. Similar to Experiment 1, we calculated evaluative and descriptive scores, which were then separately entered into a 2 (gender of judge) by 2 (gender of target) by 3 (perspective: own vs. male vs. female) by 8 (scale pair) mixed ANOVA with repeated measures on the last two factors.

Major Findings

The analysis of evaluative scores in Experiment 1 revealed a tendency that each target was evaluated more positively by in-group judges than by out-group judges. Russian judges showed clear in-group favoritism, evaluating the Russian target significantly higher than the Latvian target. Latvian judges did not see both targets as significantly different in evaluative terms. The analysis of descriptive scores showed that both groups of judges tended to agree on the descriptive attributes of each target. Furthermore, the judges saw both targets as different from each other.

In Experiment 2, the analysis of evaluative scores showed that neither of the two groups of judges displayed significant in-group favoritism in evaluation of both targets. However, when taking the group perspectives, the participants gave the highest evaluation to each target from the target's in-group perspective, expecting the ratings by their peers to be positively biased towards in-group. Similar to the results in Experiment 1, both groups of judges tended to agree on the descriptive attributes of each target. Although there were some differences in ratings from own and the imagined group perspectives, these differences were not systematical and did not reflect any false polarization.

Discussion

The results demonstrated that intergroup ratings reflect agreement on the contents of shared group stereotypes despite the bias introduced by in-group

favoritism (Experiment 1). This agreement is largely independent of whether these ratings are made from one's own perspective or imagined group perspectives (Experiment 2). The results of Experiment 2 produced a moderate false polarization effect for the evaluative scores, although the participants themselves did not show any in-group favoritism. The results support the hypothesis that the false polarization effect is more likely to occur on evaluative dimensions where the knowledge of group norms provide guidelines about how group members are supposed to judge different targets. However, it would be interesting to replicate the results of Experiment 2 in a more competitive intergroup context where the social categorization is more salient.

Study 2. The False Polarization Effect in Explanations of Attitudinal Behavior.

This study tests whether the false polarization effect is present in explanations of causes of attitudinal behavior when these explanations are given from in-group and out-group perspectives. The false polarization effect has been previously found in studies of political beliefs and attitudes (Pronin, Puccio, & Ross, 2002). However, apart from some aspects of studies conducted by Robinson et al. (1995), this phenomenon has not been systematically researched in the area of intergroup attributions. Our study used a number of different attributional dimensions reported in literature on previous research. Given the robust nature of the false polarization effect, we predicted that it should appear in intergroup attributions as well. That is, when imagining how specific causes for a given target behavior are rated from the target's in-group and out-group perspectives, participants would give significantly higher (or lower) ratings than the respective in-group and out-group members themselves. Theoretically, the mean differences should occur in the group-serving direction. However, from the previous reports of intergroup attribution studies, it was not possible to formulate precise hypotheses about what should be seen as the group serving direction for all variables employed. Therefore, our general prediction was that the mean differences from opposing group perspectives should be in opposite directions for the same target behavior; that is, what is seen as favorable by a target's in-group would be seen as less favorable by the out-group and vice versa.

Method

Ninety-two supporters and 49 opponents of Latvia's EU membership rated two opposite attitudinal behaviors (actively supporting EU membership vs. actively protesting against EU membership) on 17 rating scales constructed to cover various aspects of intergroup attributions. Attitudinal behavior is general, non-specific behavior that expresses the actors' attitude towards a certain issue. In other words, actors behaving in line with their attitudes. Each participant rated each target behavior from three perspectives: one's own

perspective and the imagined group perspectives of EU supporters and EU opponents. Of the 17 rating scales, 2 scales measured participants' evaluation of the target behavior, and 15 scales represented various attributional dimensions. Besides the variables of causal locus, controllability, stability, and globality (Islam & Hewstone, 1993), we included a number of explanatory dimensions taken from the FEB model (Malle, 1999, 2001). In addition, the participants also indicated their agreement/disagreement with seven items measuring their support to EU membership. The participants rated the same items also from both group perspectives.

Major Findings

For two attributional dimensions–causal locus (actor vs. situation) and globality of causes–there were no significant mean differences between judges or target behaviors and no false polarization effect. The false polarization effect appeared in estimations of stability of causes of behavior and in ratings of the actors' control over the causes of their behavior. From the target's imagined in-group perspective, the participants rated the causes as more stable, and the actors more in control of the causes of their behavior, than the target's in-group did from their own perspective. Interestingly, most of the time the ratings from the judges' own perspective did not differ significantly between both groups. The false polarization effect was not equally strong for all target behaviors and from all imagined group perspectives, but it always occurred in the expected direction.

After a principal component analysis, two indexes–rationality index and subjectivity index–were constructed from the items derived from the FEB scheme. The pattern of means for the rationality index showed that the participants saw rationality as a favorable cause of behavior and showed the false polarization effect when rating the rationality of causes of both target behaviors. The participants generally saw subjectivity as an unfavorable factor in determining causes of behavior. Regarding the means of subjectivity index, the participants showed false polarization when rating both target behaviors from the targets' in-group perspectives but not from out-group perspectives. False polarization of smaller magnitude was also found in the ratings of actors' acceptance of other people's opinions, which was generally seen as an unfavorable cause of behavior.

Discussion

The study confirmed that the false polarization effect is a robust phenomenon, which can affect not only prediction of group attitudes but also causal explanations of attitudinal behavior. Interestingly, we did not find any group-serving effects in ratings of globality of causes and the person-situation attributions, nor did we find any false polarization effects on these scores. One explanation for this finding is that the study did not compare positive and negative behaviors by both groups. Although all four dimensions have been

31

shown to yield group-serving biases for comparisons of positive and negative in-group and out-group behaviors (cf., Islam & Hewstone, 1993), the same effects may be less pronounced when explaining non-specific attitudinal behavior. Another possible explanation is that globality of causes and the person-situation distinction per se have neither positive nor negative connotations (whereas for example stability, control, and rationality as concepts have positive connotations). It is likely that the evaluative connotation of a rating dimension triggers the false polarization effect in estimating in-group and out-group ratings along with the general social categorization effects (i.e., in-group favoritism).

Study 3. False Polarization when Imagining Others' Explanations of Attitudinal Behavior

This study had two aims. First, we wanted to examine whether the false polarization effect would appear in free-response explanations of attitudinal behavior. Second, we wanted to explore the possibilities of using the BEB coding scheme in studies of intergroup attributions. We made two predictions regarding the false polarization effect. We hypothesized that one dimension that could elicit group-serving biases and false polarization could be the ratio between the number of reason explanations and causal history of reason (CHR) explanations used by the participants. Reason explanations imply actors' awareness of the causes of their behavior; CHR explanations imply a lack of such awareness. Awareness, according to Locke & Pennington (1982), in the context of attributions largely equals rationality, which, in turn, is seen as a favorable causal factor (Kenworthy & Miller, 2002). Our prediction was that participants should provide more reason (relative to CHR) explanations when taking the (imagined) in-group perspective of the target than the target's in-group members have provided from their own perspective. And vice versa, simulated explanations given from a target's out-group perspective should contain more CHR (relative to reason) explanations compared to the explanations given by that target's out-group members. Another prediction concerned the use of marked and unmarked reason explanations. In line with Malle's (1999) reasoning that mental state markers are likely to be used by observers to distance themselves from the agent's subjective reasons, we hypothesized that one should expect more marked (relative to unmarked) reason explanations when a target's out-group members explain the target's behavior than if the same behavior is explained by the target's in-group members. Consequently, our prediction for the false polarization effect was that participants should use more mental state markers when taking the (imagined) out-group perspective of the target than the target's out-group have used from their own perspective; the opposite should be true regarding the target's in-group perspective.

Method

Sixty-one supporters and 23 opponents of Latvia's EU membership provided free-response explanations to two attitudinal behaviors—people actively supporting and people actively opposing EU membership. Each participant explained each behavior from one's own perspective and from the imagined perspective of either EU supporters or EU opponents (the group perspective was varied between participants). The participants also rated their support to EU membership on seven items, and they answered the same items from the same imagined group perspective from which they had explained the behavior.

Two independent judges then carried out a content analysis of the responses, using the F.EX Coding Scheme for Folk Explanations of Behavior (Malle, 2000). Explanation types (cause explanations, causal history of reason [CHR] explanations, reason explanations [marked and unmarked], and enabling factor explanations) were used as the primary codes. The contents of each explanation were coded using a number of secondary codes. There was a high agreement between the judges on the primary codes (88.3% of the units), but unsatisfactory agreement on the secondary codes (65.7%). Thus, only the primary codes were used in the data analysis. As predicted, the numbers of reason and CHR explanations and the numbers of marked and unmarked reason explanations were negatively correlated for each target behavior. This allowed us to construct a reason-CHR index for each participant by subtracting the number of CHR explanations from the number of reason explanations and dividing the difference by the total number of explanations given. A marked-unmarked index was constructed in an identical manner.

Major Findings

The predicted pattern of false polarization appeared in the means of reason-CHR index. Generally, each target's in-group members showed the strongest false polarization (in terms of using more CHR explanations relative to reason explanations) when explaining the target behavior from its out-group perspective. For the marked-unmarked index, there was some evidence of false polarization, but not all mean differences were significant and in the expected direction.

Discussion

The results of Study 3 are interesting from two aspects. First, it was shown that the explainers tended to use more reason explanations for attitude-congruent (in-group) behavior and more causal history of reason explanations for attitude-incongruent (out-group) behavior. The ratio between both types of explanations indeed reflects group-serving biases in causal attribution. In addition, the use of mental state markers is related to the explainers' attitude towards the target behavior (with more markers used to distance oneself from the actors' motives, which are unacceptable to the explainer). These results

show that the FEB coding scheme can be successfully applied to the context of intergroup attributions and its categories used for calculating indexes that reflect in-group favoritism in attributions. Second, the results demonstrate that the false polarization effect occurs both in numerical ratings and in choice of expressions when predicting a group's position. The explanations given from group perspectives were more simplistic than explanations provided from the explainers' own perspective. The study demonstrated the robust nature of the false polarization effect, although it could not answer if this effect happens because of asymmetric perception of bias or biased presentation of social information (or a combination of both factors).

Study 4. Differentiating Explanations of Attitude-consistent Behavior: The Role of Perspectives and Mode of Perspective Taking.

This study was designed to test whether participants could differentiate between causal explanations previously provided by their in-group and out-group members from their own perspective and the imagined perspectives of both groups. The study had two aims. First, we wanted to test if the participants would be able to see a difference between explanations provided from the same perspective when this perspective is taken as an in-group perspective and when it is taken as an out-group perspective. If the shared reality hypothesis is true, and the groups formulate their arguments with the opposite arguments in mind, they should be able to "emulate" out-group opinion effectively when giving explanations from an out-group perspective. Consequently, the explanations provided from the same perspective by members of both groups should be of similar quality, and perceivers unaware of the origin of the explanations should see no differences between them. Second, we wanted to test whether the participants would be able to see any differences between explanations given when a given group perspective is taken as explainer's own perspective and an imagined in-group or out-group perspective. If the "naïve realism" hypothesis is true, and the participants formulate the group opinions with an assumption that their peers are more biased than they are, the explanations provided from the imagined group perspectives should be more simplified and exaggerated. Consequently, perceivers unaware of the origin of explanations should see some difference in quality between both types of explanations regardless of the group perspective.

Method

In Experiment 1, 78 supporters and 57 opponents of Latvia's EU membership expressed their agreement or disagreement with a number of causal explanations selected from the data of Study 3. In Experiment 2, 60 supporters and 28 opponents of EU membership expressed their agreement with the same explanations, rated each explanation on a number of scales, and attempted to

guess whether the explanation had been initially provided by a supporter or opponent of the EU membership.

We selected 72 explanations from Study 3, 36 for each target behavior (actively supporting vs. actively protesting Latvia's EU membership). Half of the explanations were provided by EU supporters, and the other half by the opponents from the three possible perspectives: one's own perspective, imagined supporter perspective, and imagined opponent perspective. In Experiment 1, the explanations were randomly assigned to three different versions of the questionnaire. The participants rated their agreement or disagreement with 12 explanations for each target behavior. In Experiment 2, six different versions of the questionnaire were created, and participants rated six explanations for each target behavior. In both experiments, the explanations came from the six possible combinations of perspective (supporter vs. opponent) and mode of perspective taking (i.e., how that perspective had been taken: as own perspective, as an imagined in-group perspective, or as an imagined out-group perspective). In Experiment 2, in addition to rating their agreement with each explanation, the participants also rated how positive/negative the explanation was towards the target group, to what extent the explanation depicted the target behavior as rational or irrational, how typical the explanation was for the Latvian society, and how substantiated the explanation was. After that, the participants were informed that the explanations they had just rated had been provided by supporters and opponents of EU membership from their own perspective and imagined group perspectives. The participants were then asked to guess if the explanation had been initially provided by an EU supporter or an opponent.

Major Findings

In line with the predictions, in both experiments participants agreed more with explanations provided from their in-group perspective than from an out-group perspective. The agreement was independent of whether their in-group members or out-group members had taken that perspective. In addition, the participants tended to agree more with the explanations provided by explainers expressing own opinion rather than the explanations given from the imagined group perspectives. Furthermore, in Experiment 2 own explanations were rated as more positive towards the target group, better substantiated, and depicting the target behavior as more rational than the simulated in-group or out-group explanations. Interestingly, own explanations were rated also as more typical than the group explanations, although for this variable the mean differences were smaller. In line with the initial predictions, the participants were able to "guess" only the perspective from which the explanation had been provided, but they could not "guess" the explainer attitude when this attitude was not congruent with the perspective.

In general, group members can recognize the perspective from which an explanation has been provided irrespective of whether the perspective has been taken by an in-group or out-group member. However, they are not able to distinguish effectively between perspectives that are congruent (in-group perspective) or incongruent (out-group perspective) with the explainer attitude. There may be no perceivable basis for distinguishing between these two types of perspectives as long as they reflect the common knowledge about how both sides *should* explain both target behaviors. In general, the results of Study 4 support the assumption that group perspective taking is at least in part based on shared social reality, i.e., on being well-informed of out-groups' opinions and arguments. Interestingly, the participants perceived explanations provided from own perspective more favorably than the simulated explanations. These results support the hypothesis that the false polarization effect results from overestimation of bias in other people. One may speculate that the participants' expectation of other people's bias in Study 3 resulted in producing lower quality (perhaps simplified, more extreme, or less reasonable) explanations from the imagined group perspectives, which, in turn, resulted in lower agreement with these explanations in Study 4.

GENERAL DISCUSSION

Main Findings and Contributions of this Research

This thesis aimed to gain a better understanding of what shared reality means in intergroup perception. The results of the reported studies illustrate the complexity of the concept of shared reality and point to a variety of possible interpretations of the term. The results show that even within groups there can be different ways to define and measure shared reality. One interpretation of shared reality within a group is a purely sociological one. That is, shared reality can be defined as the similarity between the actual opinions of the group members. Although in all the reported studies there was certain variability in the group opinions, for many (although not all) variables the opinions of the members of the same group were more similar than the opinions of members of different groups. This is not the sort of shared reality that Festinger (1950, 1954) or Hardin and Higgins (1996) had in mind; in this sense, people who have never heard about each other can share a social reality. Thus, the reality is shared in sociological sense, but it is not shared psychologically. However, such an interpretation and measurement of shared reality is of some practical and theoretical importance. Of practical importance, this offers a baseline estimate of what the actual opinions are against which the group members' beliefs about the opinions can be measured. Of theoretical importance, it reflects the internalized group norms among individuals belonging to the same group.

Admittedly, the fact that group members have similar opinions regarding some topic or object of judgement does not say anything about the causal relationships between social category membership and this similarity. The group members may reason in a certain way because of their group membership (a likely direction of causality in Study 1), or they may be in the same group precisely because they share the same opinions (as in Studies 2-4). In any case, the mean opinion of the group is the actual manifestation of the group norm as far as it has any influence on the group members, and thus it serves as a baseline estimate for comparison when studying the beliefs *about* the group norm.

Group members' beliefs about what other members of their group think constitute an alternative way to operationalize shared reality within a group. Such a definition of shared reality–as group members' beliefs about the beliefs of other group members–more closely corresponds to the "classical" interpretation of shared reality as a process of social verification. What is interesting, however, is that the studies reported in this thesis as well as the previous work on the false polarization effect (Pronin, Puccio, & Ross, 2002) have documented a discrepancy between the group members' beliefs about the opinions within a group and the actual mean opinion in the group. The psychologically shared social reality can actually be (and often is) different from the social reality in sociological sense. The group members agree in their view about the average opinion of their group, although most of them personally do not share this opinion. This situation is very similar to the one used to define pluralistic ignorance: "the case in which virtually every member of a group privately rejects a belief, opinion, or practice, yet believes that virtually every other member privately accepts it" (Prentice & Miller, 1996, p. 161). Of course, in the case of false polarization, group members do not reject the prevalent opinion; they just believe they endorse it to a lesser extent than others. Otherwise, the two phenomena have much in common and, as Miller, Monin, and Prentice (2000) observe, may share a common source: the fact that partisans' public behavior reflect the sharply defined norms and positions of their social identity rather than the less extreme positions of their private attitudes. The false polarization effect reflects group members' knowledge of group norms and their estimated support to these norms in the group. That they significantly overestimate this support may appear "ignorant" from a sociological perspective. Yet, as long as the knowledge of group norms and the perceived acceptance of these norms influence public behavior to a larger extent than the group members' private attitudes (Miller et al., 2000), it is feasible to consider them a more valid measurement of the shared reality than the mean measurement of the actual private opinions. Moreover, research has shown that the perception of support to the norm may cause these private opinions to gravitate towards that norm (Prentice & Miller, 1993), so the "ignorant misperception" of today may turn out to be a "true estimate" of tomorrow.

Such a definition of shared reality works well also in an intergroup context. As the reported results show, the actual disagreement between groups on various issues may vary from none to moderate, but on many occasions the groups almost perfectly agree on how much they should disagree on these issues. The groups demonstrate a very good knowledge of each other's norms, reflected in stereotypical images about the opinions of these groups. The reported studies strongly support interpretation of shared reality as shared stereotypes between groups, i.e., the agreement between groups on their mutual stereotypes. In both experiments of Study 1, groups of judges agreed on the descriptive attributes of target groups; in Experiment 2, they also agreed that each group should see its attributes as more favorable than the other group. In Study 2, both groups agreed on how each of them should agree or disagree to various causal explanations of in-group or out-group attitudinal behavior. In Study 3, both groups produced highly similar explanations when taking the perspective of the same group, and as Study 4 demonstrated, the explanations provided from the same perspective were of the same quality even those given by holders of an opposite attitude. Results reported by Austers (2002b) also support the interpretation of shared reality as shared stereotypes between groups. He found that Latvian and Russian students agreed on the value stereotypes of Latvians (majority group) and contemporary people. Both stereotypes were significantly different from individual self-ratings, reported by both groups, which in turn were very similar to each other. However, both groups disagreed on the stereotype of Russians. Russian students saw Russian targets as more similar to themselves, whereas Latvian students saw Russians as more similar to contemporary people in general. This result can be explained by the status differences between the groups.

All these findings point to the fact that regardless of how strongly the stereotypes are actually endorsed by individual group members, they have a reality of their own, and this is the reality that the group members acknowledge. These stereotypes are true to the extent that the groups believe and agree that they are true (Oakes & Reynolds, 1997). In this sense, the groups were very effective in taking each other's perspective just as one would have predicted from Montgomery's (1994) perspective theory. But the research raised another question: What can be understood with perspective taking in intergroup context? Do group members really try to imagine themselves in the other group's position and try to emulate their way of reasoning, or do they simply recite readily available stereotypes about how the other group is supposed to reason? The reported research leaves this question without a definite answer. Results of Study 3 and Study 4 suggest that group members may rely on readily available stereotypes rather than exerting the cognitive effort of imagining themselves in the position of the other group, at least in the cases where perspectives of large, diffuse groups are taken. Explanations provided from group perspectives were simplified and exaggerated; moreover, explanations

provided from the perspective of the same group were remarkably similar regardless of whether the perspective was taken by in-group members or out-group members. This, in turn, could mean that the polarized ratings and simplified arguments produced from the imagined group perspectives do not really reflect assumption of bias in others but rather represent a condensed version of the group norm. The results of Study 4 cast a doubt on such an explanation and suggest that the assumption of bias in others at least to some extent influences the quality of simulated arguments provided from group perspectives. If the imagined explanations reflected the group norms, the participants of Study 4 should have perceived these explanations as favorably as those provided from own perspective (or even more favorably), but the results reflected an opposite pattern. However, to answer these questions properly, further research is needed to focus on how exactly individuals come up with predictions of group opinion.

In general, the present research shows that, at least descriptively, it is more appropriate to define shared social reality (both within and between groups) as shared beliefs about beliefs of the group members rather than defining it simply as shared beliefs. Admittedly, such a definition does not offer a very rational portrait of the social perceiver. Both the present studies and previous research on the false polarization effect, pluralistic ignorance, and the third person effect show that such beliefs about the beliefs are often exaggerated in comparison with the actual beliefs or even opposite to the actual beliefs. From a normative point of view, it is tempting to label this sort of agreement a shared "pseudo-reality", something that only seems real and could be proven wrong at a closer look. However, real-life examples illustrate the power of such "pseudo-real" beliefs to become self-fulfilling prophecies. Representatives of conflicting parties feel pressured by the opinions of their groups when negotiating (Rubin, Pruitt, & Kim, 1994) and may fail to reach an agreement simply because they underestimate the extent to which the "street opinion" favors compromise and an end to the conflict. Investors try to predict how others will react to a piece of news and act on these predictions, which in turn affects stock prices (Wärneryd, 2001).

The admission that the shared social reality upon which individuals and groups act is vulnerable to psychological biases has implications for both theoretical reasoning and practical research. From a theoretical point of view, it underscores the necessity to define very precisely what one understands by shared reality if such a term is used to explain social psychological phenomena, as well as the necessity to specify the mechanisms of shared reality that influence these phenomena. Alternatively, it is worth considering to use the term "shared reality" only to refer to an individuals' perception of the extent to which their views are shared by others (see Hardin & Higgins, 1996), and to find other terms to describe the actual similarity between the views of different individuals and groups.

From a practical point of view, the most interesting and relevant question for research is how to reduce the gap between the "perceived" and the "actual" reality, that is, to reduce the bias in prediction of other people's opinions. Puccio (2003) presents an example of such research by demonstrating how simple manipulations in presenting information about the parties' views can reduce the false polarization effect. It would be interesting to replicate these results in a real-world situation. Another challenge in this area could be laboratory research on bias recognition in self. One may speculate that making the social perceivers aware of their own susceptibility to cognitive biases could reduce the perceived gap between one's own opinion and that of the group. Previous research has shown that introspection is unlikely to bring about such awareness. Perspective taking could serve as an alternative method of self-analysis. Given the social perceivers' competence in recognizing cognitive biases in others, seeing oneself from an outside perspective, i.e., imagining how another critical person would see the social perceiver, might improve the recognition of cognitive biases in self.

An interesting result of the reported studies was that the group-serving effects of social categorization did not happen on all the measured variables; in addition, the false polarization effect did not influence all the variables. In Study 2, we did not find any categorization effects or any false polarization for such attributional variables as globality of causes and causal locus of behavior. On some other variables, we did not find actual group differences in explanations of the same target behavior, but nevertheless the participants expected such differences in responses of their peers. One possible explanation why the same pattern did not occur for the globality and locus variables is that these dimensions do not have either positive or negative connotations. Thus, they did not provide clear cues for the participants which answer should be seen as favoring one or the other group. In other words, there was no information about the group norms regarding these two variables. Consequently, there was no prescribed direction into which a systematic overestimation of group positions might occur. Such an explanation is supported by the fact that we did not find any systematic overestimation of group position in prediction of descriptive scores in Experiment 2 of Study 1, but we found an in-group serving overestimation for the evaluative scores. Awareness of in-group favoritism might have prompted the participants to expect that each target would be seen favorably by its in-group members, but there was no such normative knowledge available regarding the descriptive aspects of ratings. Of course, this hypothesis is a post-hoc explanation and should be tested in a properly designed study. If this assumption is true, it may have important implications for the studies of the false polarization effect. Some variables could be more polarizing than others because the social norms predict that groups should strongly disagree on certain topics. These social categorization effects could be independent of other factors influencing false polarization such as asymmetric bias perception or biased

presentation of information. Thus, their influence should be distinguished from the other factors and measured separately if one wants to gain full understanding of the mechanisms behind the false polarization effect.

All four reported studies demonstrate the robustness of the false polarization effect across various contexts and measurements. Although moderate in size, the effect was present when the respondents imagined how their peers assign positive and negative traits to both gender groups in Experiment 2 of Study 1. The effect was even stronger in intergroup attributions as shown in Study 2 and Study 3. The results of Study 3 are especially illustrative because they show that the effect occurs not only in "placement" of the average group opinion on a given rating dimension but also in emulating the group members' way of thinking. The results of Study 4 showed that the simulated explanations were perceived less favorably by the group members than the explanations given from participants' own perspective. These differences in perception point to possibilities of further research that could provide additional information about the cognitive mechanisms behind the false polarization effect.

Study 3 confirms that the FEB coding scheme (Malle, 1999, 2001) can be successfully applied to studying intergroup attributions. More specifically, both dimensions of comparison used in the study–the number of reason vs. causal history of reason explanations and the use of mental state markers for reason explanations–reflect intergroup bias when explaining in-group and out-group behavior. Perhaps there are other variables in the coding scheme that can be used in a similar manner. In Study 3, we did not use the contents of the explanations in data analysis because of a relatively low initial agreement between judges who conducted the content analysis. However, a repeated content analysis with experienced/better trained judges might reveal new dimensions in the data that could be used for studying attributional biases in intergroup contexts and that could provide additional information about how the false polarization effect is manifested in free responses.

In a wider theoretical context, the results of Study 2 and Study 3 confirmed that awareness of the causes of one's behavior is seen as a positive, desirable causal factor. These results are similar to the findings of Pronin, Lin, and Ross (2002) who found that the social desirability of various biases largely overlapped with the cognitive availability of these biases and that the participants did not show asymmetry in perception of cognitively available biases. (Examples of cognitively available and socially desirable biases include friend enhancement, trust of strangers, trust of borrowers, and upward and downward comparison.) Our results are also similar to those of Kenworthy and Miller (2002), who found that rationality was seen as a desirable cause of group attitudes more often attributed to in-group members than out-group members. It would be interesting to study how the perception of awareness of causes of behavior is related to the correspondence bias/fundamental attribution error. It is

reasonable to expect some connection between both constructs. An actor who is reacting to a situation could arguably be seen as more aware of the causes of his behavior than an actor who behaves in line with his dispositional qualities (e.g., personality traits). Further studies in this direction could shed more light on the cognitive mechanisms responsible for correspondence bias.

Shortcomings and Limitations

One limitation of the reported studies is that all of them were largely exploratory. Although predictions were formulated and tested for each study, none of them were designed with a precise, refutable hypothesis in mind. This was largely because of limitations of the previous knowledge basis. Despite the relatively wide use of the shared reality concept in the previous literature, attempts to operationalize it have been rare. The present studies have illustrated how difficult it is to define shared reality as a measurable construct and raised a question of how useful it is to use this term in social psychological studies despite of its convenience in post hoc explanations of various phenomena. Regarding the proper definition of shared social reality, the research reported in this thesis has admittedly raised more questions than it has provided answers. Rather, it points to necessity of a discussion about the term among social psychologists that would conventionalize its use in a more precise way and allow formulating refutable hypotheses about the nature and effects of shared reality both within and between groups.

The situation is more optimistic regarding various aspects of intergroup attribution research addressed in this thesis. Studies 2 and 3 were based on prudent and general predictions because they addressed previously unexplored aspects of intergroup attribution. Study 2 applied a number of known attributional dimensions in a new context, causal explanations of attitudinal behavior. Most of the previous research had been conducted within the "ultimate attribution error" paradigm (Pettigrew, 2001) with its division between positive and negative outcomes, and it could not serve to formulate refutable hypotheses for Study 2. Study 3 used new variables for measurement of bias in intergroup attributions based on the FEB coding scheme that had not been used in intergroup perception before. Now, however, it is possible to build further studies on these findings by formulating precise, refutable hypotheses.

Another limitation of the reported studies is the use of a repeated-measures design. In all four studies, all participants rated both targets. In Study 1 (Experiment 2) and Study 2, all participants rated the targets from own perspective and both imagined group perspectives. The use of a repeated-measures design could have had some effect on the research results because it made the research hypothesis more transparent to the participants. Moreover, one may speculate that a repeated-measures design triggers a different cognitive mechanism in estimating the group opinion by asking the participants to

juxtapose their own opinion to the group (in contrast to asking only about a group opinion in a between-subjects design). It would be interesting to replicate the reported results with a between-subjects design. Such a replication would improve the validity of the findings and provide more information about the cognitive mechanisms involved in predicting group opinions. We conducted a preliminary study to explore the effects of data-gathering design on the false polarization effect (Dimdins & Montgomery, 2003). The study included 150 first year psychology students at Stockholm University. The study took place two weeks before the scheduled national referendum on Sweden's membership in the European Monetary Union (EMU). The participants rated their attitude towards EMU and predicted the consequences of EMU membership on Swedish economy, welfare, and democracy. In the between-subjects condition, the participants answered the questions either from their own perspective or from one of the group perspectives (i.e., imagining how either EMU supporters or opponents in their class answered the same questions). In the within-subjects condition, each participant answered the questions from all three perspectives. The data analysis showed that across all the variables the participants in the within-subjects condition showed significantly higher false polarization effect than did the participants in the between-subjects condition. However, the pattern of results varied between separate variables and between both group perspectives. Therefore, additional research is needed to document the nature and magnitude of the differences between both types of data-gathering designs.

The third limitation of the reported studies is the definition of groups whose opinions the participants were asked to predict. Only in one of the reported surveys (Experiment 2 of Study 1), the question about the imagined group perspectives was defined precisely about the group whose data we used for estimation of the group members' own opinions. In Studies 2 and 3, the group perspectives were defined more broadly, and thus the respondents' opinions were technically calculated from a sample that was not representative of the groups about whose perspectives the respondents were asked to take. However, in Study 2 we used two different formulations for the imagined group perspectives and did not find any systematic differences in results between them. It should be noted that both formulations—EU supporters/opponents in general vs. EU supporters/opponents among University of Latvia students—defined the group as more inclusive than the actual respondent group (who were students of one faculty only). Generally, it is very unlikely that this incongruity might have accounted for the false polarization effect in the results. In a number of other studies, researchers have taken special care of the formulation of group-perspective questions and still found the false polarization effect (Robinson et al., 1995; Puccio, 2003). Moreover, we found some false polarization in Experiment 2 of Study 1 where the sample was representative of the group whose perspective the respondents were asked to take. We also found the false polarization effect in the EMU study among psychology students at

Stockholm University, where the respondents were asked to predict the responses of their classmates who filled in the same questionnaire (Dimdins & Montgomery, 2003). At the same time, it makes intuitive sense to assume that as the number of people in a group decreases, and the group members' familiarity with each other increases, the accessibility of information about the group members' actual opinions should increase as well. Consequently, the group members' reliance on stereotypes in predicting others' opinions should decrease, which should in turn reduce or eliminate the false polarization effect. This assumption emphasizes the importance of precise formulation of the group perspective to ensure validity of the study results, and it also points to some possibilities for future studies to be discussed below.

Directions for Further Research

Mapping the Mechanisms of Group Perspective Taking

Most of the reported research on group perspective taking has focused on the results of this process. One prospective direction for future studies would include asking the social perceivers to report on the process of perspective taking; for example, asking them to respond using the think-aloud protocols. Two questions in this regard would be of special interest. First, recording think-aloud protocols would help to identify the source of information that the social perceivers use to define group position. Second, this method would help to uncover the mechanism that accounts for the divergence between a respondent's own opinion and the imagined group opinion. For example, it would help to answer the question whether the false polarization effect is indeed a result of polarizing the imagined group opinion or of moderating or elaborating one's own opinion in relation to a general group norm. In addition, the two ratings could be unrelated and result from different cognitive processes such as relying on one's experience to formulate one's own opinion and relying on general stereotypes in case of the group opinion.

Future research on false polarization should stress the difference between the perceptions of what opinion an average group member *does* have and what opinion a group member *should* have in accordance with the group norms. Mapping the mechanisms of how social perceivers explain others' susceptibility to group norms could help to understand the role of group norms and social categorization effects in false polarization. In addition, making this difference salient could reduce the false polarization effect.

Cognitive Representations of Groups

Related to the previous research area is the question of how groups are represented in the social perceiver's mind when the imagined group opinion is formulated. Mullen, Rozell, and Johnson (1996) observe that instructions to process information about a target group with prototype (Cantor & Mischel, 1979) representations are likely to exaggerate such consequences of intergroup

perception as in-group bias and relative homogeneity. Alternatively, instructions to process information about a target group with exemplar representations (Smith & Zarate, 1992) are likely to mitigate some of these consequences. The hypothesis would be that if the group opinion is based on prototype representations, the false polarization should be stronger than if it is based on exemplar representations. On the other hand, a highly polarized imagined group opinion could be based on few extreme and salient real cases (exemplars), confirming the biased social presentation hypothesis.

Influence of Group Size and Previous Contact

Related to the problem of group representation is the question of group size and the social perceiver's experience in individual contacts with members of the group. A number of previous studies have suggested that both of these factors might influence the extent of false polarization. For example, decision making research has shown that manipulating the group size and emotional significance of hypothetical group members can eliminate the framing effects in the life-death decision dilemma (Wang, 1996). Epley and Dunning (2000) found that the participants felt superior (in the sense of overestimating the likelihood of generous and selfless behavior) only in relation to a random peer but not in relation to an individualized peer. Kemdal and Montgomery (2000) found that animal rights activists and animal experimenters could take each other's perspectives without showing any false polarization effect. One explanation for this result is that both groups had been in contact for a long time and they knew and understood their opponent's opinions. These results seem to suggest that the false polarization is more likely to occur in relation to large, diffuse groups rather than to small, well-defined groups whose individual opinions are known to the social observer.

Perception of Simulated Group Opinions

One interesting direction for further research would be to elaborate on the results in Study 4. The answer to the question of why the simulated explanations were on average perceived less favorably than own explanations might provide some variables for systematic evaluation of consequences of the false polarization effect. Such variables, in turn, could help validate eventual results acquired from the think-aloud protocols described above.

Concluding Remarks

The present thesis illustrates several problem areas within social psychology. One is the difference between interpersonal perception and interaction and individual-group and intergroup perception and interaction. Because people define themselves in terms of social category memberships and non-memberships, there is a strong motivation for the social perceivers to stay at the level of group norms and group ideologies in their social judgements. The false polarization effect seems to be a direct consequence of such ideological,

normative judgements. It offers a measurement of the discrepancy between the intergroup social reality that is constructed at the level of group norms, ideologies, and stereotypes, and the reality in sociological sense as reflected in people's actual opinions. The mechanisms of interpersonal perception/interaction are simply insufficient for explaining intergroup perception/interaction.

The false polarization effect also illustrates another problem area–the complex interaction between cognitive and social factors in creating a social reality between groups. Groups as entities are as real for the social perceiver as are other individuals. Their influence on the social perceiver can be even stronger than that of another individual. Unlike individuals, groups cannot provide the social perceiver with unequivocal feedback about how correct his or her assumptions about the group beliefs and traits are. Any feedback that an individual receives from a group is likely to be much more susceptible to multiple interpretations than feedback received from an individual. Because the assumption of bias in others, individual motivational factors, and biased social presentation, the individual's exaggerated beliefs about group opinion may never be corrected. Neither cognitive nor social factors alone can account for such phenomena as the false polarization effect, pluralistic ignorance, or the third person effect. It is the interaction of both types of factors that makes individual judgement primitive and amplifies individual bias into social action, like herd behavior.

The third problem highlighted in the present thesis is more practical than theoretical. Although the reported studies have demonstrated the robustness of the false polarization effect, the actual process of reasoning that leads to these robust results is still not clear. Of course, the results are interesting on their own. However, these results are open to diverse theoretical speculations regarding the process that led to them. Finding out how the group opinions are represented in people's minds and how exactly people come up with predictions of group opinions could help to identify those parts in the process of individual reasoning where the discrepancy between one's own and the predicted group opinion appears. This, in turn, could point to possible mechanisms for reducing this discrepancy. Such mechanisms, especially if applicable outside the laboratory, might help to reduce intergroup misunderstanding and conflict in various social contexts.

REFERENCES

Abrams, D., Wetherell, M., Cochrane, S., Hogg, M. A., & Turner, J. C. (2001). Knowing what to think by knowing who you are: Self-categorization and the nature of norm formation, conformity and group polarization. In M. A. Hogg & D. Abrams (Eds.), *Intergroup relations: Essential readings* (pp. 270-288). Philadelphia, PA: Psychology Press.

Abramson, L. Y., Seligmann, M. E. P., & Teasdale, J. D. (1978). Learned helplessness in humans: Critique and reformulation. *Journal of Abnormal Psychology, 87,* 49-74.

Allport, G. W. (1958). *The nature of prejudice.* Garden City, NY: Doubleday.

Altemeyer, B. (1998). The other "authoritarian personality". In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 30, pp. 47-92). San Diego, CA: Academic Press.

Anderson, R. C., & Pichert, J. W. (1978). Recall of previously unrecallable information following a shift of perspective. *Journal of Verbal Learning and Verbal Behavior, 17,* 1-12.

Asch, S. E. (1958). Effects of group pressure upon the modification and distortion of judgments. In E. E. Maccoby, T. M. Newcomb, & E. L. Hartley (Eds.), *Readings in Social Psychology* (3rd ed., pp. 174-183). New York: Henry Holt & Company.

Augoustinos, M., Ahrens, C., & Innes, J. M. (1994). Stereotypes and prejudice: The Australian experience. *British Journal of Social Psychology, 33,* 125-141.

Augoustinos, M., & Walker, I. (1995). *Social cognition: An integrated introduction.* London: Sage.

Austers, I. (2002a). Ability to take an out-group's perspective in explaining positive and negative behaviors. *Scandinavian Journal of Psychology, 43,* 407-412.

Austers, I. (2002b). Attribution of value stereotypes as a consequence of group membership: Latvian and Russian students living in Latvia compared. *International Journal of Intercultural Relations, 26,* 273-285.

Austers, I., & Montgomery, H. (2001). Explaining an out-group's political behavior from four different perspectives. *Journal of Baltic Psychology, 2,* 23-33.

Bar-Tal, D. (2000). *Shared beliefs in a society: Social psychological analysis.* Thousand Oaks, CA: Sage.

Batson, C. D., & Burris, C. T. (1994). Personal religion: depressant or stimulant of prejudice and discrimination? In M. P. Zanna & J. M. Olson (Eds.), *The psychology of prejudice: The Ontario Symposium, Volume 7* (pp. 149-170). Hillsdale, NJ: Lawrence Erlbaum.

Biernat, M., Vescio, T. K., & Theno, S. A. (1996). Violating American values: A "value congruence" approach to understanding outgroup attitudes. *Journal of Experimental Social Psychology, 32*, 387-410.

Billig, M. (2003). Political Rhetoric. In D. O. Sears, L. Huddy, & R. Jervis (Eds.), *Oxford handbook of political psychology* (pp. 222-250). New York: Oxford University Press.

Branscombe, N. R., Wann, D. L. (1994). Collective self-esteem consequences of out-group derogation when a valued social identity is on trial. *European Journal of Social Psychology, 24*, 641-657.

Brewer, M. B. (1999). The psychology of prejudice: ingroup love or outgroup hate? *Journal of Social Issues, 55*, 429-444.

Brewer, M. B. (2001). Ingroup identification and intergroup conflict: When does ingroup love become outgroup hate? In R. D. Ashmore, L. Jussim, & D. Wilder (Eds.), *Social identity, intergroup conflict, and conflict reduction* (pp. 17-41). New York: Oxford University Press.

Brewer, M. B., & Brown, R. J. (1998). Intergroup relations. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *The handbook of social psychology* (4th ed., Vol. 2, pp. 554-594). New York: McGraw-Hill.

Brewer, M. B., & Feinstein, A. S. H. (1999). Dual processes in the cognitive representation of persons and social categories. In S. Chaiken, & Y. Trope (Eds.), *Dual-process theories in social psychology* (pp. 255-270). New York: Guilford.

Brown, R. (2000). Social identity theory: past achievements, current problems and future challenges. *European Journal of Social Psychology, 30*, 745-778.

Buehler, R., Griffin, D., & Ross, M. (1994). Exploring the "planning fallacy": Why people underestimate their task completion times. *Journal of Personality and Social Psychology, 67*, 366-381.

Buss, A. R. (1979). On the relationship between causes and reasons. *Journal of Personality and Social Psychology, 37*, 1458-1461.

Campbell, D. T. (1967). Stereotypes and the perception of group differences. *American Psychologist, 22*, 817-829.

Cantor, N., & Mischel, W. (1979). Prototypes in person perception. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 12, 3-52). New York: Academic Press.

Cohen, J., Mutz, D., Price, V., & Gunther, A. (1988). Perceived impact on defamation: An experiment on third-person effects. *Public Opinion Quarterly, 52*, 161-173.

Davison, W. P. (1983). The third-person effect in communication. *Public Opinion Quarterly, 47*, 1-15.

Dawes, R. M., Singer, D., & Lemons, F. (1972). The contrast effect and its implications for intergroup communication and the indirect assessment of attitude. *Journal of Personality & Social Psychology, 21*, 281-295.

Deaux, K, & Emswiller, T. (1974). Explanations of successful performance on sex-linked tasks: What is skill for the male is luck for the female. *Journal of Personality and Social Psychology, 29,* 80-85.

Devine, P. G. (1989). Stereotypes and prejudice: Their automatic and controlled components. *Journal of Personality and Social Psychology, 56,* 5-18.

Devine, P. G., & Monteith, M. J. (1999). Automaticity and control in stereotyping. In S. Chaiken, & Y. Trope (Eds.), *Dual-process theories in social psychology* (pp. 339-360). New York: Guilford.

Diekman, A. B., Eagly, A. H., & Kulesa, P. (2002). Accuracy and bias in stereotypes about the social and political attitudes of women and men. *Journal of Experimental Social Psychology, 38,* 268-282.

Dimdins, G., & Montgomery, H. (2003). *The false polarization effect in between-subjects and within-subjects designs.* Unpublished data, Stockholm University.

Duck, J. M., & Mullin, B. (1995). The perceived impact of the mass media: Reconsidering the third person effect. *European Journal of Social Psychology, 25,* 77-93.

Duncan, B. L. (1976). Differential social perception and attribution of intergroup violence: Testing the lower limits of stereotyping of blacks. *Journal of Personality and Social Psychology, 34,* 590-598.

Ellemers, N., Spears, R., & Doosje, B. (2002). Self and social identity. *Annual review of Psychology, 53,* 161-186.

Epley, N., & Dunning, D. (2000). Feeling "holier than thou": Are self-serving assessments produced by errors in self- or social prediction? *Journal of Personality & Social Psychology, 79,* 861-875.

Esses, V. M., Jackson, L. M., & Armstrong, T. L. (1998). Intergroup competition and attitudes toward immigrants and immigration: An instrumental model of group conflict. *Journal of Social Issues, 54,* 699-724.

Festinger, L. (1950). Informal social communication. *Psychological Review, 57,* 271-282.

Festinger, L. (1954). A theory of social comparison processes. *Human Relations, 7,* 117-140.

Fischhoff, B. (1975). Hindsight ≠ foresight: The effect of outcome knowledge on judgment under uncertainty. *Journal of Experimental Psychology: Human Perception & Performance, 1,* 288-299.

Fiske, S. T. (1992). Thinking is for doing: Portraits of social cognition from daguerreotype to laserphoto. *Journal of Personality and Social Psychology, 63,* 877-889.

Fiske, S. T. (1998). Stereotyping, prejudice and discrimination. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *The handbook of social psychology* (4th ed., Vol. 2, pp. 357-411). New York: McGraw-Hill.

Fiske, S. T., Lin, M., & Neuberg, S. L. (1999). The continuum model: Ten years later. In S. Chaiken, & Y. Trope (Eds.), *Dual-process theories in social psychology* (pp. 231-254). New York: Guilford.

Fiske, S. T., & Neuberg, S. L. (1990). A continuum of impression formation, from category-based to individuating processes: Influences of information and motivation on attention and interpretation. In M. P. Zanna (Ed.), *Advances in Experimental Social Psychology* (Vol. 23, pp. 1-74). San Diego, CA: Academic Press.

Fiske, S. T., & Taylor, S. E. (1991). *Social cognition* (2nd ed.). New York: McGraw Hill.

Frank, M. G., & Gilovich, T. (1989). Effect of memory perspective on retrospective causal attributions. *Journal of Personality and Social Psychology, 57,* 399-403.

Friedrich, J. (1996). On seeing oneself as less self-serving than others: The ultimate self-serving bias? *Teaching of Psychology, 23,* 107-109.

Fussell, S. R., & Krauss, R. M. (1991). Accuracy and bias in estimates of others' knowledge. *European Journal of Social Psychology, 21,* 445-454.

Gardner, R. C. (1994). Stereotypes as consensual beliefs. In M. P. Zanna & J. M. Olson (Eds.), *The psychology of prejudice: The Ontario Symposium. Volume 7* (pp. 1-31). Hillsdale, NJ: Lawrence Erlbaum.

Gilbert, D. T., Brown, R. P., Pinel, E. C., & Wilson, T. D. (2000). The illusion of external agency. *Journal of Personality and Social Psychology, 79,* 690-700.

Gilbert, D. T., & Malone, P. S. (1995). The correspondence bias. *Psychological Bulletin, 117,* 21-38.

Hardin, C.D., & Higgins, E. T. (1996). Shared reality: How social verification makes the subjective objective. In R. M. Sorrentino, & E. T. Higgins (Eds.), *Handbook of motivation and cognition.* (Vol. 3, pp. 28-84). New York: The Guilford.

Haslam, S. A. (1997). Stereotyping and social influence: Foundations of stereotype consensus. In R. Spears, P. J. Oakes, N. Ellemers, & S. A. Haslam (Eds.), *The social psychology of stereotyping and group life* (pp. 119-143). Oxford: Blackwell.

Haslam, S. A., Turner, J. C., Oakes, P. J., McGarty, C., & Hayes, B. K. (1992). Context-dependent variation in social stereotyping 1: The effects of intergroup relations as mediated by social change and frame of reference. *European Journal of Social Psychology, 22,* 3-20.

Hastorf, A. H., & Cantril, H. (1954). They saw a game: A case study. *Journal of Abnormal and Social Psychology, 49,* 129-134.

Heath, C. (1999). On the social psychology of agency relationships: Lay theories of motivation overemphasize extrinsic incentives. *Organizational Behavior and Human Decision Processes, 78,* 25-62.

Heider, F. (1958). *The psychology of interpersonal relations.* New York: Wiley.

Hewstone, M. (1989). *Causal attribution: From cognitive processes to collective beliefs.* Oxford: Basil Blackwell.

Hewstone, M. (1990). The "ultimate attribution error"? A review of the literature on intergroup causal attribution. *European Journal of Social Psychology, 20,* 311-335.

Hewstone, M., Jaspars, J., & Lalljee, M. (1982). Social representations, social attribution and social identity: The intergroup images of 'public' and 'comprehensive' schoolboys. *European Journal of Social Psychology, 12,* 241-269.

Hewstone, M., Rubin, M., & Willis, H. (2002). Intergroup bias. *Annual Review of Psychology, 53,* 575-604.

Hewstone, M., & Ward, C. (1985). Ethnocentrism and causal attribution in Southeast Asia. *Journal of Personality and Social Psychology, 48,* 614-623.

Hilton, J. L., & Von Hippel, W. (1996). Stereotypes. *Annual Review of Psychology, 47,* 237-71.

Hogg, M. A., & Abrams, D. (1999). Social identity and social cognition: Historical background and current trends. In D. Abrams & M. A. Hogg (Eds.), *Social identity and social cognition* (pp. 1-25). Oxford: Blackwell.

Hogg, M. A., & Turner, J. C. (1987). Social identity and conformity: A theory of referent information influence. In W. Doise, & S. Moscovici (Eds.), *Current issues in European social psychology* (Vol. 2, pp. 139-182). Cambridge: Cambridge University Press.

Hunter, J. A., Stringer, M., & Watson, R. P. (1991). Intergroup violence and intergroup attribution. *British Journal of Social Psychology, 30,* 261-266.

Ichheiser, G. (1949). Misunderstandings in human relations: A study in false social perception. *American Journal of Sociology, 55 (Suppl.).*

Islam, M. R., & Hewstone, M. (1993). Intergroup attributions and affective consequences in majority and minority groups. *Journal of Personality and Social Psychology, 64,* 936-950.

Jackson, L. A., Sullivan, L. A., & Hodge, C. N. (1993). Stereotype effects of attributions, predictions, and evaluations: No two social judgements are quite alike. *Journal of Personality and Social Psychology, 65,* 69-84.

Jones, E. E. (1979). The rocky road from acts to dispositions. *American Psychologist, 34,* 107-117.

Jones, E. E., & Davis, K. E. (1965). From acts to dispositions: The attribution process in person perception. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 2, pp. 219-266). New York: Academic Press.

Jones, E. E., & Harris, V. A. (1967). The attribution of attitudes. *Journal of Experimental Social Psychology, 3,* 1-24.

Jones, E. E., & Nisbett, R. E. (1972). The actor and the observer: Divergent perceptions of the causes of behavior. In E. E. Jones, D. E. Kanouse, H.

H. Kelley, R. E. Nisbett, S. Valins, & B. Weiner (Eds.), *Attribution: Perceiving the causes of behavior* (pp. 79-94). Morristown, NJ: General Learning.

Jones, E. E., Worchel, S., Goethals, G. R., & Grumet, J. F. (1971). Prior expectancy and behavioral extremity as determinants of attitude attribution. *Journal of Experimental Social Psychology, 7,* 59-80.

Judd, C. M., & Park, B. (1993). Definition and assessment of accuracy in social stereotypes. *Psychological Review, 100,* 109-128.

Keltner, D., & Robinson, R. J. (1993). Imagined ideological differences in conflict escalation and resolution. *International Journal of Conflict Management, 4,* 249-262.

Keltner, D., & Robinson, R. J. (1997). Defending the status quo: Power and bias in social conflict. *Personality and Social Psychology Bulletin, 23,* 1066-1077.

Kemdal, A. B., & Montgomery, H. (2000). *Seeing it from the other side: Perspectives and attitudes towards animal experimentation.* Unpublished manuscript, Stockholm University.

Kemdal, A. B., & Montgomery, H. (2001). Explaining own and others' behavior in a controversial issue: Animal experimentation. *Journal of Social Psychology, 141,* 693-713.

Kenny, D. A., & Albright, L. (1987). Accuracy in interpersonal perception: A social relations analysis. *Psychological Bulletin, 102,* 390-402.

Kenworthy, J. B., & Miller, N. (2002). Attributional biases about the origins of attitudes: Externality, emotionality, and rationality. *Journal of Personality and Social Psychology. 82,* 693-707.

Kinket, B., & Verkuiten, M. (1999). Intergroup evaluations and social context: A multilevel approach. *European Journal of Social Psychology, 29,* 219-237.

Krauss, R. M., & Fussell, S. R. (1996). Social psychological models of interpersonal communication. In E. T. Higgins, & A. W. Kruglanski (Eds.), *Social psychology: Handbook of basic principles* (pp. 655-701). New York: Guilford.

Krueger, J., & Zeiger, J. S. (1993). Social categorization and the truly false consensus effect. *Journal of Personality and Social Psychology, 65,* 670-680.

Kruger, J., & Gilovich, T. (1999). "Naïve cynicism" in everyday theories of responsibility assessment: On biased assumptions of bias. *Journal of Personality and Social Psychology, 76,* 743-753.

Kruglanski, A. W. (1975). The endogenous-exogenous partition in attribution theory. *Psychological Review, 82,* 387-406.

Kruglanski, A. W. (1979). Causal explanation, teleological explanation: On radical particularism in attribution theory. *Journal of Personality and Social Psychology, 37,* 1447-1457.

Kunda, Z. (1990). The case for motivated reasoning. *Psychological Bulletin, 108*, 480-498.

Locke, D., & Pennington, D. (1982). Reasons and other causes: Their role in attribution process. *Journal of Personality and Social Psychology, 42*, 212-223.

Luus, C. A. E., & Wells, G. L. (1994). The malleability of eyewitness confidence: Co-witness and perseverance effects. *Journal of Applied Psychology, 79*, 714-723.

Malle, B. F. (1999). How people explain behavior: A new theoretical framework. *Personality and Social Psychology Review, 3*, 23-48.

Malle, B. F. (2000). *F. Ex: A coding scheme for folk explanations of behavior (Version 4.0)* [On-line]. Retrieved December 17, 2002, from: http://darkwing.uoregon.edu/~interact/fex.html

Malle, B. F. (2001). Folk explanations of intentional action. In B. F. Malle, L. J. Moses, & D. A. Baldwin (Eds.), *Intentions and intentionality: Foundations of social cognition* (pp. 265-286). Cambridge, MA: MIT Press.

Malle, B. F., & Knobe, J. (1997). The folk concept of intentionality. *Journal of Experimental Social Psychology, 33*, 101-121.

Malle, B. F., Knobe, J., O'Laughlin, M. J., Pearce, G. E., & Nelson, S. E. (2000). Conceptual structure and social functions of behavior explanations: Beyond person-situation attributions. *Journal of Personality and Social Psychology, 79*, 309-326.

Maoz, I., Ward, A., Katz, M., & Ross, L. (2002). Reactive devaluation of an "Israeli" vs. "Palestinian" peace proposal. *Journal of Conflict Resolution, 46*, 515-546.

Marks, G., & Miller, N. (1987). Ten years of research on the false-consensus effect: An empirical and theoretical view. *Psychological Bulletin, 102*, 72-90.

McGuire, W. J. (1964). Inducing resistance to persuasion: Some contemporary approaches. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 1, pp. 191-229). New York: Academic Press.

Mead, G. H. (1934). *Mind, self and society: From the standpoint of a social behaviorist* (C. W. Morris, Ed.). Chicago: University of Chicago Press

Mikula, G., Athenstaedt, U., Heschgl, S., & Heimgartner, A. (1998). Does it only depend on the point of view? Perspective-related differences in justice evaluations of negative incidents in personal relationships. *European Journal of Social Psychology, 28*, 931-962.

Miller, A. G., Baer, R., & Schonberg, P. (1979). The bias phenomenon in attitude attribution: Actor and observer perspectives. *Journal of Personality and Social Psychology, 37*, 1421-1431.

Miller, D. T., Monin, B., & Prentice, D. A. (2000). Pluralistic ignorance and inconsistency between private attitudes and public behaviors. In D. J.

Terry, & M. A. Hogg (Eds.), *Attitudes, behavior, and social context: The role of norms and group membership* (pp. 95-113). Mahwah, NJ: Erlbaum.

Miller, D. T., & Ratner, R. K. (1998). The disparity between the actual and assumed power of self-interest. *Journal of Personality and Social Psychology, 74*, 53-62.

Miller, F. D., Smith, E. R., & Uleman, J. (1981). Measurement and interpretation of situational and dispositional attributions. *Journal of Experimental Social Psychology, 17*, 80-95.

Monin, B., & Norton, M. I. (2003). Perceptions of a fluid consensus: Uniqueness bias, false consensus, false polarization, and pluralistic ignorance in a water conservation crisis. *Personality and Social Psychology Bulletin, 29*, 559-567.

Montgomery, H. (1994). Towards a perspective theory of decision making and judgement. *Acta Psychologica, 87*, 155-178.

Moreland, R. L., Argote, L., & Krishnan, R. (1996). Socially shared cognition at work: Transactive memory and group performance. In J. L. Nye, & A. M. Brower (Eds.), *What's social about social cognition? Research on socially shared cognition in small groups* (pp. 57-84). Thousand Oaks, CA: Sage.

Moscovici, S. (1981). On social representations. In J. P. Forgas (Ed.), *Social cognition: Perspectives on everyday understanding* (pp. 181-209). London: Academic Press.

Moscovici, S. (1984). The phenomenon of social representations. In R. M. Farr, & S Moscovici (Eds.), *Social representations* (pp. 3-69). Cambridge: Cambridge University Press.

Moscovici, S. (1988). Notes towards a description of social representations. *European Journal of Social Psychology, 18*, 211-250.

Mullen, B., Brown, R., & Smith, C. (1992). In-group bias as a function of salience, relevance, and status: An integration. *European Journal of Social Psychology, 22*, 103-122.

Mullen, B., Rozell, D., & Johnson, C. (1996). The phenomenology of being in a group: Complexity approaches to operationalizing cognitive representation. In J. L. Nye, & A. M. Brower (Eds.), *What's social about social cognition? Research on socially shared cognition in small groups* (pp. 205-229). Thousand Oaks, CA: Sage.

Mummendey, A. (1995). Positive distinctiveness and social discrimination: an old couple living in divorce. *European Journal of Social Psychology, 25*, 657-670.

Nickerson, R. S. (1999). How we know--and sometimes misjudge--what others know: Imputing one's own knowledge to others. *Psychological Bulletin, 125*, 737-759.

Nickerson, R. S., Baddeley, A., & Freeman, B. (1987). Are people's estimates of what other people know influenced by what they themselves know? *Acta Psychologica, 64,* 245-259.

Nisbett, R., & Ross, L. (1980). *Human inference: Strategies and shortcomings of social judgement.* Englewood Cliffs, NJ: Prentice-Hall.

Nisbett, R. E., & Wilson, T. D. (1977a). Telling more than we can know: Verbal reports on mental processes. *Psychological Review, 8,* 231-259.

Nisbett, R. E., & Wilson, T. D. (1977b). The halo effect: Evidence for unconscious alteration of judgments. *Journal of Personality & Social Psychology, 35,* 250-256.

Oakes, P. J., & Reynolds, K. J. (1997). Asking the accuracy question: Is measurement the answer? In R. Spears, P. J. Oakes, N. Ellemers, & S. A. Haslam (Eds.), *The social psychology of stereotyping and group life* (pp. 51-71). Oxford: Blackwell.

O'Laughlin, M. J., & Malle, B. F. (2002). How people explain actions performed by groups and individuals. *Journal of Personality and Social Psychology, 82,* 33-48.

Otten, S., & Mummendey, A. (2000). Valence-dependent probability of ingroup favoritism between minimal groups: an integrative view on the positive-negative asymmetry in social discrimination. In D. Capozza, & R. Brown (Eds.), *Social identity processes: Trends in theory and research* (pp. 96-116). London: Sage.

Peabody, D. (1968). Group judgements in the Philippines: Evaluative and descriptive aspects. *Journal of Personality and Social Psychology, 10,* 290-300.

Peabody, D. (1985). *National characteristics.* Cambridge: Cambridge University Press.

Pettigrew, T. F. (2001). The ultimate attribution error: Extending Allport's cognitive analysis of prejudice. In M. A. Hogg & D. Abrams (Eds.), *Intergroup relations: Essential readings* (pp. 162-173). Philadelphia, PA: Psychology Press.

Piaget, J. (1999). *Judgement and reasoning in the child* (M. Warden, Trans.). London: Routledge. (Original work published 1928)

Piaget, J., & Inhelder, B. (1997). *The child's conception of space* (F. J. Langdon, & J. L. Lunzer, Trans.). London: Routledge. (Original work published 1956)

Prentice, D. A., & Miller, D. T. (1993). Pluralistic ignorance and alcohol use on campus: Some consequences of misperceiving the social norm. *Journal of Personality and Social Psychology, 64,* 243-256.

Prentice, D. A., & Miller, D. T. (1996). Pluralistic ignorance and the perpetuation of social norms by unwitting actors. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 28, pp. 161-209). San Diego, CA: Academic Press.

Pronin, E., Gilovich, T., & Ross, L. (in press). Objectivity in the eye of the beholder: Divergent perceptions of bias in self versus others. *Psychological Review.*

Pronin, E., Kruger, J., Savitsky, K., & Ross, L. (2001). You don't know me, but I know you: The illusion of asymmetric insight. *Journal of Personality & Social Psychology, 81,* 639-656.

Pronin, E., Lin, D. Y., & Ross, L. (2002). The bias blind spot: Perceptions of bias in self versus others. *Personality and Social Psychology Bulletin, 28,* 369-381.

Pronin, E., Puccio, C., and Ross, L. (2002). Understanding misunderstanding: Social psychological perspectives. In T. Gilovich, D. Griffin, & D. Kahneman (Eds.), *Heuristics and biases: The psychology of intuitive judgment* (pp. 636-665). Cambridge: Cambridge University Press.

Puccio, C. (2003). *The search for common ground: Overcoming false polarization and unwarranted pessimism about ideological differences.* Unpublished doctoral dissertation, Stanford University.

Regan, D. T., & Totten, J. (1975). Empathy and attribution: Turning observers into actors. *Journal of Personality and Social Psychology, 32,* 850-856.

Reicher, S., & Hopkins, N. (1996). Self-category constructions in political rhetoric: An analysis of Thatcher's and Kinnock's speeches concerning the British Miners strike (1984-85). *European Journal of Social Psychology, 26,* 353-371.

Reynolds, K. J., Turner, J. C., & Haslam, S. A. (2000). When are we better than them and they worse than us? A closer look at social discrimination in positive and negative domains. *Journal of Personality and Social Psychology, 78,* 64-80.

Robinson, R. J., & Friedman, R. A. (1995). Mistrust and misconstrual in union-management relationships: causal accounts in adversarial contexts. *International Journal of Conflict Management, 3,* 312-327.

Robinson, R. J., Keltner, D., Ward, A., & Ross, L. (1995). Actual versus assumed differences in construal: "Naïve realism" in intergroup perception and conflict. *Journal of Personality and Social Psychology, 68,* 404-417.

Ross, L. (1977). The intuitive psychologist and his shortcomings. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 10, pp. 173-220). New York: Academic Press.

Ross, L., Greene, D., & House, P. (1977). The false consensus effect: An egocentric bias in social perception and attribution processes. *Journal of Experimental Social Psychology, 13,* 279-301.

Ross, L., & Ward, A. (1996). Naive realism in everyday life: Implications for social conflict and misunderstanding. In T. Brown, E. Reed, & E. Turiel (Eds.), *Values and knowledge* (pp. 103-135). Hillsdale, NJ: Erlbaum.

Rouhana, N. N., O'Dwyer, A., & Morrison Vaso, S. K. (1997). Cognitive biases and political party affiliation in intergroup conflict. *Journal of Applied Social Psychology, 27,* 37-57.

Rubin, J. Z., Pruitt, D. G., & Kim, S. H. (1994). *Social conflict: Escalation, stalemate and settlement* (2nd ed.). New York: McGraw Hill.

Rubin, M., & Hewstone, M. (1998). Social identity theory's self-esteem hypothesis: A review and some suggestions for clarification. *Personality and Social Psychology Review, 2,* 40-62.

Rustemli, A., Mertan, B., & Ciftci, O. (2000). In-group favoritism among native and immigrant Turkish Cypriots: Trait evaluations of in-group and out-group targets. *Journal of Social Psychology, 140,* 26-34.

Sachdev, I., & Bourhis, R. Y. (1987). Power and status differentials in minority and majority intergroup relations. *European Journal of Social Psychology, 21,* 1-24.

Sherif, M. (1936). *The psychology of social norms.* New York: Harper & Brothers.

Sherif, M. (1966). *Group conflict and co-operation: Their social psychology.* London: Routledge.

Smith, E. R., & Zarate, M. A. (1992). Exemplar-based model of social judgement. *Psychological Review, 99,* 3-21.

Snyder, M., Tanke, E. D., & 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.

Stangor, C., & Schaller, M. (1996). Stereotypes as individual and collective representations. In C. N. Macrae, C. Stangor, & M. Hewstone (Eds.), *Stereotypes and stereotyping* (pp. 3-37). New York: Guilford.

Storms, M. D. (1973). Videotape and the attribution process: Reversing actors' and observers' points of view. *Journal of Personality and Social Psychology, 27,* 165-175.

Tajfel, H. (1969). Cognitive aspects of prejudice. *Journal of Social Issues, 25*(4), 79-97.

Tajfel, H. (1978). Social categorization, social identity and social comparison. In H. Tajfel (Ed.), *Differentiation between social groups: Studies in the social psychology of intergroup relations* (pp. 61-76). London: Academic Press.

Tajfel, H. (1981a). *Human groups and social categories.* Cambridge: Cambridge University Press.

Tajfel, H. (1981b). Social stereotypes and social groups. In: J.C. Turner, & H. Giles (Eds.), *Intergroup Behavior* (pp. 144-167). Oxford: Blackwell.

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

Taylor, D. M., & Jaggi, V. (1974). Ethnocentrism and causal attribution in a South Indian context. *Journal of Cross-Cultural Psychology, 5,* 162-171.

Taylor, S. E. (1982). The availability bias in social perception and interaction. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgement under uncertainty: Heuristics and biases* (pp. 190-200). Cambridge: Cambridge University Press.

Tesser, A. (1988). Toward a self-evaluation maintenance model of social behavior. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 21, pp. 181-227). San Diego, CA: Academic Press.

Thomson, M. S., Judd, C. M., & Park, B. (1999). The consequences of communicating social stereotypes. *Journal of Experimental Social Psychology, 36,* 567-599.

Turner, J. C. (1991). *Social influence.* Milton Keynes: Open University Press.

Turner, J. C., Hogg, M. A., Oakes, P. J., Reicher, S. D., & Wetherell, M. S. (1987). *Rediscovering the social group: A self-categorization theory.* Oxford: Blackwell.

Turner, J. C., & Oakes, P. J. (1989). Self-categorization theory and social influence. In P. B. Paulus (Ed.), *Psychology of group influence* (2nd ed., pp. 233-275). Hillsdale, NJ: Lawrence Erlbaum.

Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology, 5,* 207-232.

Vallone, R. P., Ross, L., & Lepper, M. R. (1985). The hostile media phenomenon: Biased perception and perceptions of media bias in coverage of the Beirut massacre. *Journal of Personality and Social Psychology, 49,* 577-585.

Van Boven, L., Kamada, A., & Gilovich, T. (1999). The perceiver as perceived: Everyday intuitions about the correspondence bias. *Journal of Personality and Social Psychology, 77,* 1188-1199.

Van Boven, L., White, K., Kamada, A., & Gilovich, T. (2003). Intuitions about situational correction in self and others. *Journal of Personality and Social Psychology, 85,* 249-258.

Van Knippenberg, D., & Wilke, H. (1992). Prototypicality of arguments and conformity to ingroup norms. *European Journal of Social Psychology, 22,* 141-155.

Verkuiten, M., & Nekuee, S. (1999). Ingroup bias: the effect of self-stereotyping, identification and group threat. *European Journal of Social Psychology, 29,* 411-418.

Vivian, J. E., & Berkowitz, N. H. (1992). Anticipated bias from an outgroup: An attributional analysis. *European Journal of Social Psychology, 22,* 415-424.

Wang, X. T. (1996). Domain-specific rationality in human choices: Violations of utility axioms and social contexts. *Cognition, 60,* 31-63.

Weiner, B. (1986). *An attributional theory of motivation and emotion.* New York: Springer Verlag.

Wetzel, C. G., Wilson, T. D., & Kort, J. (1981). The halo effect revisited: Forewarned is not forearmed. *Journal of Experimental Social Psychology, 17,* 427-439.

Wilson, T. D., Houston, C. E., Etling, K. M., & Brekke, N. (1996). A new look at anchoring effects: Basic anchoring and its antecedents. *Journal of Experimental Psychology: General, 125,* 387-402.

Word, C. O., Zanna, M. P., & Cooper, J. (1974). The nonverbal mediation of self-fulfilling prophecies in interracial interaction. *Journal of Experimental Social Psychology, 10,* 109-120.

Wärneryd, K.-E. (2001). *Stock-market psychology: How people value and trade stocks.* Cheltenham: Edward Elgar.

STUDY 1

Differentiating In-Group Favoritism from Shared Reality in Intergroup Perception

Girts Dimdins and Henry Montgomery
Department of Psychology, Stockholm University, Sweden

Abstract
Two basic factors influence mutual ratings of social groups: in-group favoritism (related to the evaluative aspects of a rating) and the perception of shared reality (related to the descriptive aspects). In two studies, we examine the usefulness of Peabody's (1968) method of separating evaluative and descriptive aspects of rating in intergroup judgments. In Study 1, Latvian and Russian students made different evaluations of both groups, but the same groups agreed on the descriptive ratings. In Study 2, male and female psychology students rated each other from own, in-group, and out-group perspectives. The participants did not show any in-group favoritism in their own ratings, but they expected their fellow students to be in-group biased. The participants agreed on the descriptive ratings of both groups. The results demonstrate that shared reality influences intergroup ratings, despite differences in evaluations.

Key words: Shared reality, in-group favoritism, evaluative ratings, descriptive ratings, perspective taking

According to Tajfel (1981), an important function of social stereotypes is that it creates and preserves positively valued differentiation between one's own and other social groups. This implies that it is essential in an intergroup context to view positively one's in-group (any social category that individuals belong to that is salient at the given moment) more than any out-group (social categories that individuals do not identify with at that same moment). Empirical studies support this assumption (Campbell, 1967; Esses, Haddock, & Zanna, 1994; Esses & Zanna, 1995; Peabody, 1968; Rustemli, Mertan, & Ciftci, 2000; Saenger & Flowerman, 1954; Scaillet & Leyens, 2000; Smedley & Bayton, 1978).

At the same time, judging a social group is difficult, especially if based only on evaluative terms. Most often groups that are stereotyped by the social perceiver live in the same social space as the perceiver. Hardin and Higgins (1996) note that "social interaction is predicated upon and regulated by the establishment of shared reality" (p.30). Because interacting groups share the same social reality, they are likely to possess some evaluation-free, descriptive information about each other. A group's auto-stereotypes and the stereotypes held about the same group by other groups are often similar (McAndrew et al., 2000; Peabody, 1985). However, this is not the case when one group has little knowledge about another group—that is, when the group being judged and the judges themselves do not share the same reality. In studies during the Cold War and shortly after the end of the cold war, the stereotypes of Russians expressed by representatives of other nations differed significantly from the stereotypes Russians expressed about themselves, their auto-stereotypes (Peabody & Shmelyov, 1996; Stephan et al., 1993). The distinction between evaluative and descriptive aspects of judgment has also been emphasized in the research on stereotype accuracy (Judd & Park, 1993), decision-making (Montgomery, 1994), and the factor structure of personality traits (Peabody, 1984; Peabody & Goldberg, 1989; Saucier, 1994).

Peabody (1968) describes a method for differentiating between both aspects of group ratings. He argues that single-trait ratings typically confound the evaluative and descriptive aspects of judgment. A similar problem is present when a target is judged on a single scale, offering a contrast between two traits. For example, if a typical member of a group is rated as quite thrifty on a bipolar scale of *extravagant-thrifty*, it is not clear whether this rating reflects the rater's belief that a typical target group member is prone to control their expenditures, or whether it reflects the rater's bias for that particular group, resulting in an adjective with a positive connotation (*thrifty*) instead of an adjective with a negative connotation (*extravagant*). Peabody suggests that this question could be answered by rating the target on an additional bipolar scale, maintaining the same descriptive contrast but reversing the evaluative connotations of the adjectives: *stingy-generous*. A rating closer to *stingy* relative to *generous* would indicate that the overall rating is more descriptive, whereas the opposite would

mean that evaluative aspects play a major role in the rater's judgments. When both scales are arranged in a way that the higher rating corresponds to the positive adjective, the evaluative aspect of the rating will contribute to the score on both scales, but the descriptive aspect will do so only on one scale, decreasing the score on the other scale.

Peabody (1968) demonstrates that in intergroup ratings (two groups rating each other) the two aspects could be estimated by conducting a 2 (judge) by 2 (scale) ANOVA on the ratings of each target, with repeated measures on the last factor. The main effect of judge corresponds to the evaluative disagreement between groups about the evaluation of a target (the target's in-group judges giving higher scale averages than the target's out-group judges). The main effect of scale corresponds to the agreement between groups about the descriptive characteristics of the target (the difference between the scales being in the same direction for both groups of judges).

Peabody (1970, 1985) suggests that the two aspects of group ratings can be expressed directly as indexes. The rating on each of the two scales (a and b) can be seen as a combination of an evaluative (E) and a descriptive (D) component: $a = E + D$, and $b = E - D$. By solving for the two components, it is possible to calculate their value from the ratings on both scales: $E = (a + b)/2$; $D = (a - b)/2$. We refer to these indexes as evaluative and descriptive scores.

Although this method has been neglected in stereotype and intergroup perception studies, we believe it offers a simple and elegant possibility to differentiate between both aspects of group ratings. We intend to investigate the usefulness of Peabody's approach for studying group perception of shared reality. We will also propose possible improvements on Peabody's original method as well as discuss its limitations and the possibilities of further methodological refinements.

One aspect of Peabody's method that has been largely unexplored is a direct comparison of targets. In previous studies of intergroup perception (Peabody, 1968, 1985), judgments about each target were analyzed separately. In the studies reported below, we include both targets as within-subjects variables in the same analysis. This approach yields more precise information about how each group perceives the differences (or similarities) between the targets.

Another drawback of Peabody's method is that one faces a methodological problem when attempting to generalize the results across a number of scale pairs. Whereas the evaluative scores can be averaged across any number of scales, it is not possible to construct one index reflecting the descriptive agreement between groups of judges because the descriptive effects on different scale pairs can occur in different directions.

To estimate both effects across all scale pairs, we suggest conducting a 3-way (Target by Judge by Scale Pair) ANOVA separately on the descriptive scores and evaluative scores as the dependent variables. The analysis of the

3

evaluative scores serves to estimate the effects of in-group favoritism. The descriptive scores reflect the shared reality of intergroup ratings.

In-group favoritism should result in a tendency that the mean evaluative score of an in-group target is higher than the same mean of an out-group target. This tendency should hold true for both groups of judges and should result in a significant Judge by Target interaction (if there are no evaluative differences among the scale pairs) or a Judge by Target by Scale Pair interaction (if there are such differences).

Shared reality should be reflected in a tendency that (on most or all scale pairs) the mean descriptive score of the same target does not differ significantly between both groups of judges. This tendency should hold true for both targets, and should result in a strong Target by Scale pair interaction. At the same time, Target by Scale Pair by Judge interaction should be weak or insignificant because it reflects the extent of disagreement between judges on the descriptive ratings of each target.

It is important to note that the size of effects is contingent on how the descriptive scores are calculated for each scale pair. Which rating is subtracted from the other one will affect the sign of the descriptive score. Thus the direction of subtraction can influence all the ANOVA effects calculated across scale pairs. Although these differences in calculations create certain indeterminacy in the analysis of results, they cannot introduce or eliminate any effects related to the existence or absence of shared reality because the mean differences between targets and between judges for each separate scale pair remain the same, only their signs change.

We illustrate the use of Peabody's method in two studies of intergroup perception. In Study 1, the design is similar to Peabody's (1968) original research. In Study 2, we apply the same method in a more complex design of intergroup perception. Judges rate the targets not only from their own perspective but also from imagined in-group and out-group perspectives.

Study 1

In this study, Latvian and Russian students rated targets representing both ethnic groups. The relations between Latvians (the titular group) and Russians (the largest ethnic minority group) in Latvia have a long and complicated history. On the one hand, there exists a certain tension between both groups on a political and an everyday level. On the other hand, the groups have shared the same social universe for decades and thus should have a fairly good knowledge of each other. One would expect a fairly strong in-group favoritism to be present in the judgments of both groups of one another. However, there should also be evidence of some "shared reality" in these judgments; that is, both groups should agree to some extent on how they describe each other.

Method

Translation and pre-testing of the scales. Fourteen scale pairs taken from Peabody (1985) and eight scale pairs taken from Montgomery and Skaldeman (2001) were first translated to Latvian independently by two professional translators, who afterwards discussed the differences in translations. Two other translators then provided independent back-translations to English. All four met to identify possible translation problems and finalize the Latvian version of the scales. The same procedure was used to translate the scales into Russian. Four bilingual judges (one professional translator, one journalist, and two psychologists) independently compared the Latvian and Russian versions of the questionnaire and noted the possible discrepancies between both versions. Their suggestions were used to revise both versions to make them comparable.

The questionnaire was then pre-tested by asking 16 people (8 Russians and 8 Latvians) to answer the questions in the presence of an interviewer and to express any comments regarding their contents. Next, the interviewer explained the logic of the evaluative and descriptive contrasts to the respondents and asked for comments about how well the adjectives of each scale pair represented these contrasts. The comments were used to further improve the wording of the scales. Finally, all four judges checked and approved the final versions of the questionnaire in Latvian and Russian.

Nine pairs of scales were excluded from the study at various stages of the translation/pre-testing procedure. Although the correspondence to the English original was important, the main criteria were the relationship between the trait adjectives in each pair and the comparability between the Latvian and Russian versions of each scale pair.

Sample. The study included 86 Latvian and 77 Russian respondents (76% women). All participants were social sciences and technical science students from five different Latvian universities (most of them were in their second or third year of studies). The average age of the respondents was 19.5 years. Age and gender proportions were similar in the Latvian and Russian respondent groups.

Questionnaire. The questionnaire was taken from the middle section of a larger survey on ethnic identity and values. The participants were asked to rate the likelihood of presence of the given traits in 2 ethnic groups, *Latvians living in Latvia* and *Russians living in Latvia*, on 13 pairs of scales. Each scale ranged from 1 to 9. Both extremes (1 and 9) indicate a very large possibility of the presence of the given trait, and 5 indicates an equal possibility of both or neither of the traits. The adjectives at both ends of each scale were opposites, representing a contrast for both evaluative and descriptive aspects (e.g., *thrifty-extravagant*). Each scale had a pair that reversed the evaluative aspects but was similar regarding the descriptive aspects (e.g., *stingy-generous*). In the sequence of scales, the evaluative favorable end of the scale altered from the right to the

left (being assigned the anchors 1 or 9). The order of the scales in the questionnaire was arranged to separate those of related contents.

Results and Discussion

Before the calculations, the data were recoded so that 9 corresponded to the positive adjective on all scales. Appendix A gives all means and standard deviations. First, an evaluative score was calculated for each scale pair as the mean of ratings on both scales. Figure 1 gives the means for the evaluative score. A 2 (judge) by 2 (target) by 13 (scale pair) mixed ANOVA with repeated measures on the last two factors was carried out on the evaluative scores as dependent variables.
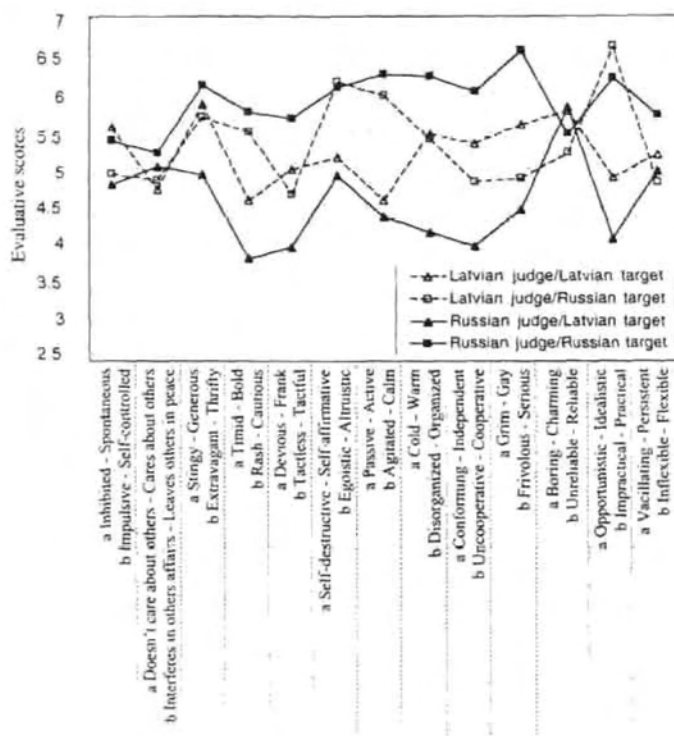
Figure 1. Mean evaluative scores for Latvian and Russian judges. Higher scores indicate more positive evaluation.

The ANOVA yielded a Target by Judge interaction, $F(1, 128) = 32.43$, $p < .01$, and a Target by Judge by Scale Pair interaction, $F(12, 128) = 7.63$, $p < .01$. Figure 1 shows that the Target by Judge interaction reflects an overall tendency that each target was evaluated more positively by in-group judges and less positively by out-group judges. As the three-way interaction shows, there

6

are differences between scale pairs regarding this pattern. However, whereas the Russian judges displayed a clear in-group favoritism evaluating Russian target significantly higher than the Latvian target (estimated marginal means 5.91 [SE = .10] and 4.57 [SE = .11] respectively), the Latvian judges did not see both targets as significantly different (M_{Lat} = 5.22 [SE = .10] and M_{Rus} = 5.36 [SE = .09]). This pattern results in a highly significant main effect of target: $F(1, 128)$ = 49.06, $p < .01$. A significant Target by Scale Pair interaction, $F(12, 128)$ = 16.24, $p < .01$, indicates that this pattern was not present or equally strong on all scale pairs. There was also a main effect of scale pair, $F(12, 128) = 13.17$, $p <$ .01, showing that for some scale pairs the evaluative contrast was stronger than for others. A significant Judge by Scale Pair interaction, $F(12, 128) = 3.52$, p < .01, shows that for some scale pairs the evaluative contrast was not equally strong for both groups of judges. Perhaps this result reflects some subtle differences between the Latvian and Russian versions of the questionnaire. In some cases the evaluative contrast in one language was stronger than in the other language.

The descriptive score for each scale pair was calculated as the difference of ratings on both scales divided by two. To reduce indeterminacy in the calculations, we chose to follow a common rationale in calculating the difference scores for all scale pairs. For each scale pair, we subtracted the rating on the scale where the positive adjective represented unassertiveness or impulse control (e.g., *cautious* in *rash-cautious*) from the rating on the scale where the positive adjective represented assertiveness or impulse expression (e.g., *bold* in *timid-bold*). A larger descriptive score thus corresponds to assertiveness/impulse expression, and a lower score to unassertiveness/ impulse control. The terms *assertiveness/ unassertiveness* and *impulse expression/ impulse control* are taken from Peabody and Goldberg's (1989) structure of personality traits, where they denote two separate factors. It makes intuitive sense, though, to use both dimensions alongside, because together they represent a contrast between more dynamic traits on the one side, and more passive traits on the other. It is important to note that this arrangement of scales represents only one of many possible ways of calculating the difference scores. However, such a calculation makes the size and direction of descriptive scores interpretable.

A 2 (judge) by 2 (target) by 13 (scale pair) mixed ANOVA with repeated measures on the last two factors was carried out on the descriptive scores as dependent variables. The analysis yielded two strong main effects: target, $F(1, 132) = 130.72$, $p < .01$, and scale pair, $F(12, 132) = 15.93$, $p < .01$. There were also strong Target by Scale Pair, $F(12, 132) = 39.28$, $p < .01$, and Target by Judge, $F(1, 132) = 10.89$, $p < .01$, interactions. There were also two weaker interactions, Target by Scale Pair by Judge, $F(12, 128) = 5.87$, $p < .01$, and Scale Pair by Judge, F(12, 128) = 2.81, p < .05. Figure 2 shows that both targets were seen as descriptively different by both groups of judges; at the same time,

7

there was a very good (although not perfect as reflected in the interactions involving judge) agreement between judges on the descriptive attributes of each target. The strong main effect of target reflects the tendency that Russians were rated as more assertive/prone to impulse expression than Latvians by both groups of judges.

The Target by Judge interaction reflects a tendency that Latvian judges saw both targets as more descriptively different than Russian judges did. This effect occurs because Russian judges displayed stronger in-group favoritism in their ratings as could be seen from the analysis of the evaluative scores.
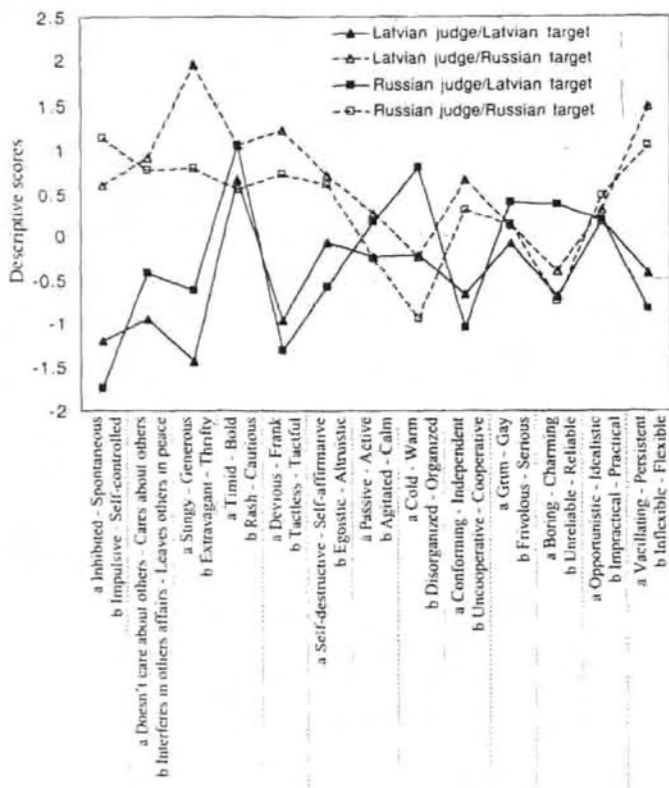
Figure 2. Mean descriptive scores for Latvian and Russian judges. Higher scores indicate that a target is rated as assertive/prone to impulse expression. When calculating the descriptive scores, rating *b* was subtracted from rating *a*.

Taken together, the results of Study 1 support our predictions that both groups agree on the descriptive aspects and disagree on the evaluative aspects of mutual ratings. One finding stands out from the predicted pattern. Although stronger in-group favoritism by Russian participants (as ethnic minority group) could be expected (Brewer & Gardner, 1996; Hewstone, Rubin, & Willis, 2002;

Mullen, Brown, & Smith, 1992), the lack of in-group favoritism demonstrated by the Latvian participants was somewhat of a surprise. One explanation might be the influence of social desirability on the responses of Latvian judges. As Fiske (2000) points out, modern racism theories hypothesize a self-esteem-based motive not to appear racist, and a similar hypothesis can also be made about the motivation to hide one's ethnocentrism. Maass, Castelli and Arcuri (2000) quote a number of studies that show a tendency for majority group members to avoid actions that might be seen as discriminating against minority group members.

This result also illustrates the usefulness of direct comparisons between targets in the analysis. If we had followed Peabody's (1968) original approach and ran the analysis separately on each target, we might have (erroneously) concluded that both groups displayed similar levels of in-group favoritism because each target received higher evaluations from its in-group judges than from out-group judges.

Study 2

Related to shared reality is the concept of perspective. Decision-making studies demonstrate that the same object attributes can be seen as positive or negative depending on the perspective taken by the perceiver (Montgomery, 1994). In addition, the members of opposing groups can successfully take the perspective of their opponent (Kemdal & Montgomery, 2001). This can reduce in-group bias and stereotyping of an out-group (Galinsky & Moskowitz, 2000).

At the same time, taking a group's perspective may lead to an exaggeration of the group's position (Robinson, Keltner, Ward, & Ross, 1995) and overestimation of differences between various groups (Diekman, Eagly, & Kulesa, 2002; Keltner & Robinson, 1997; Robinson & Friedman, 1995; Rouhana, O'Dwyer, & Morrison Vaso, 1997). Overestimation of a group score on a certain attribute in comparison with the actual group score on that attribute is known as the *false polarization effect* (Pronin, Puccio, & Ross, 2002).

Study 2 was designed to examine the effects of perspective taking on the intergroup perception. In addition to rating both targets from their own perspective, both groups were asked to make their ratings from the perspective of their in-group and out-group. As in Study 1, we predicted that both groups of judges would show in-group favoritism, disagreeing on the evaluative aspects of ratings of each target, but would agree on the descriptive aspects. In line with the perspective theory (Montgomery, 1994), we expected the false polarization effect to occur primarily in evaluations of targets in positive-negative terms but not in the descriptive ratings of the targets. We predicted that the judges would indicate stronger in-group favoritism when rating the targets from the imagined group perspectives than their own perspective. We also expected no perspective effects in the analysis of the descriptive scores.

9

Method

Selection of scales. Because the participants would be asked to make all the ratings six times (2 targets X 3 perspectives), it was necessary to reduce the number of scales used in the questionnaire. Thus only eight scale pairs already available in Swedish (taken from Montgomery and Skaldeman, 2001) were used. The Swedish scales were initially developed according to a similar pre-testing procedure as the scales in Study 1. The only difference was that out of the eight scale pairs three were not taken from Peabody but developed by the researchers (directly in Swedish). Montgomery and Skaldeman ran a multidimensional scaling on their data (with 206 participants). For each scale pair, they found clear descriptive and evaluative contrasts in line with Peabody's model.

Sample. Seventy-one students (44 women) at the Department of Psychology, Stockholm University, participated in the study. All participants were first year students and received points for course requirements for taking part in the study.

Questionnaire. We asked the participants to rate two targets: (a) *a male psychology student at Stockholm University* and (b) *a female psychology student at Stockholm University* on 8 pairs of scales. The participants were asked to indicate the usual presence of the given traits in each target. The scale ranged from 1 to 9, standing for *extremely*, with the scale midpoint 5 for *neither*. Each participant rated both targets from his or her own perspective, as well as from the perspective of (a) male psychology student at Stockholm University and (b) female psychology student at Stockholm University.

Results and Discussion

Appendix B gives the means and standard deviations for all scales. A 2 (judge) by 2 (target) by 3 (perspective) by 8 (scale pair) mixed ANOVA with repeated measures on the last three factors was carried out on the evaluative scores of all 8 scale pairs (calculated as the average rating on both scales). The analysis yielded a significant main effect of target, $F(1, 57) = 4.40, p < .05$, main effect of scale pair, $F(7, 57) = 4.18, p < .01$, and a Target by Scale Pair interaction, $F(7, 57) = 13.48, p < .01$. On average, the female target received higher evaluation than the male target. The estimated marginal means and standard errors were 5.73 (0.09) and 5.62 (0.08). As can be seen in Figure 3, the difference between targets varied depending on the scale pair, resulting in the Target by Scale Pair interaction.

In contrast to Study 1, the participants did not show significant in-group bias in evaluation of both targets. There are two possible explanations of this. First, although male and female psychology students represent distinct social categories, a salient common category (psychology student) may contribute to seeing both targets as equally positive. Second, there is no history of conflict or competition between groups that reduces the need for self-enhancement. In

10

other words, the social categorization in the present context was probably not strong enough to induce in-group favoritism.
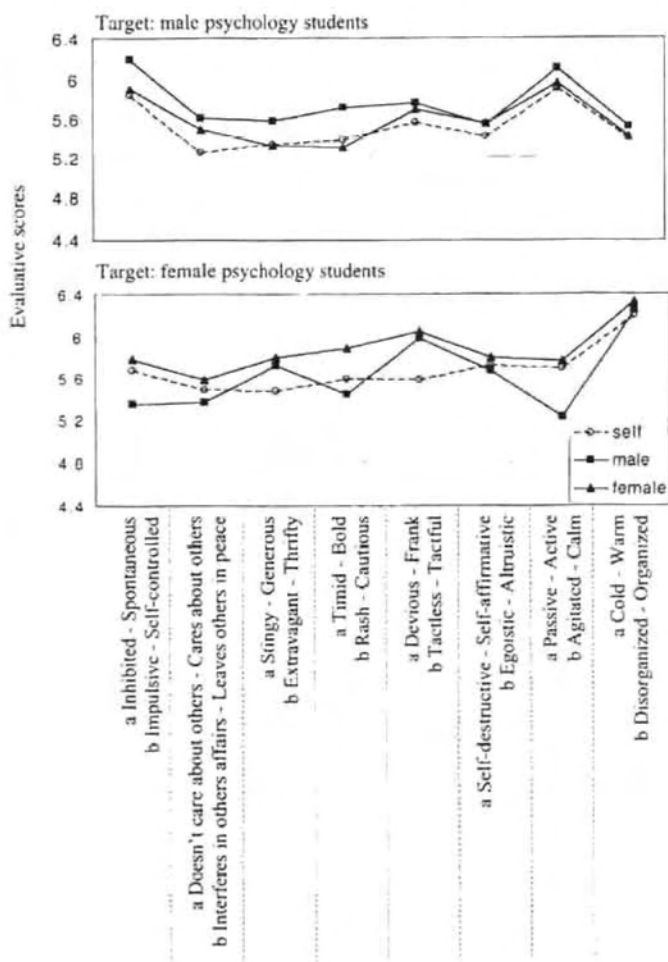
Figure 3. Mean evaluative scores for male and female targets from own, male, and female perspectives across both groups of judges. Higher scores indicate more positive evaluation.

The analysis also yielded a significant Target by Perspective interaction, $F(2, 57) = 4.20$, $p < .05$. Figure 3 shows a clear pattern: the participants gave the highest evaluation to each target from the target's in-group perspective. That is, they judged the male target most positively from a male perspective and female target most positively from a female perspective. Although participants did not display significant in-group favoritism, the pattern of means indicates that they expected the ratings of fellow participants to be biased towards in-group.

Target: male psychology students

1.5

—•— self
1
—◦— male
—△— female
0.5

0

-0.5

-1

Descriptive scores

Target: female psychology students

1.5

1

0.5

0

-0.5

-1

a Inhibited - Spontaneous
b Impulsive - Self-controlled
a Doesn't care about others - Cares about others
b Interferes in others affairs - Leaves others in peace
a Stingy - Generous
b Extravagant - Thrifty
a Timid - Bold
b Rash - Cautious
a Devious - Frank
b Tactless - Tactful
a Self-destructive - Self-affirmative
b Egoistic - Altruistic
a Passive - Active
b Agitated - Calm
a Cold - Warm
b Disorganized - Organized

Figure 4. Mean descriptive scores for male and female targets from own, male, and female perspectives across both groups of judges. Higher scores indicate that a target is rated as assertive/prone to impulse expression. When calculating the descriptive scores, rating *b* was subtracted from rating *a*.

A 2 (judge) by 2 (target) by 3 (perspective) by 8 (scale pair) mixed ANOVA with repeated measures on the last three factors was carried out on the descriptive scores of all 8 scale pairs, calculated as the difference between ratings on both scales divided by two. As in Study 1, a higher descriptive score corresponds to assertiveness/ impulse expression. The analysis yielded a significant main effect of scale pair, $F(7, 57) = 15.69$, $p < .01$, and a significant Target by Scale Pair interaction, $F(7, 57) = 19.41$, $p < .01$. There was also a weak but significant Perspective by Scale Pair, $F(14, 57) = 2.73$, $p < .05$. interaction. Figure 4 shows that although both targets were seen as fairly

different and the means occasionally differed for the three perspectives, the overall pattern across all scale pairs shows a good agreement among the judges about the descriptive attributes of each target. There seem to be no systematic differences between the ratings provided from own perspective and the imagined group perspectives.

General Discussion

The present results demonstrate the usefulness of Peabody's (1968, 1985) model of separating evaluative and descriptive components in mutual ratings of social groups. In both studies, evaluative and descriptive scores could be used as convenient and efficient tools for estimating the effects of in-group favoritism and shared reality. In a broader theoretical context, the results of Study 1 showed that intergroup ratings reflect agreement on the contents of the shared social reality despite the bias introduced by in-group favoritism. Study 2 demonstrated that this agreement is largely independent of whether these ratings are made from one's own perspective or from imagined group perspectives.

The results of Study 2 point to a promising direction for further research. We did not find any systematic disagreement between judges on the evaluations of targets (i.e., no in-group favoritism). However, when rating the targets from the imagined group perspectives, both groups of judges expected in-group favoritism to influence the scores. The perspective effects were moderate in size, but the expected bias was clearly in the predicted direction. It would be interesting to repeat this study in a more competitive social context as in Study 1. Some perspective effects were found also in the analysis of descriptive scores, but the mean differences were not systematic and cannot be attributed to the false polarization effect. However, given the robust nature of this effect (Pronin et al., 2002), we may find that it affects both the evaluative and descriptive aspects of ratings made from imagined group perspectives.

The results supported our suggested improvements to Peabody's method. We believe that directly comparing the targets in the same analysis is particularly useful both in studies of in-group favoritism and the contents of shared reality. Our other suggestion – inclusion of all scale pairs in the same analysis – also proved useful. Although the data analysis is based on complicated three-way and four-way ANOVAs, the results are easy to grasp because the number of effects corresponding to the in-group favoritism and shared reality is limited in each analysis. In addition, our approach allows conducting two analyses across all scale pairs instead of conducting separate analyses for each pair.

In both studies, we found differences between scale pairs both for evaluative and descriptive scores. One possible way of further refining Peabody's method would be to look at (or control for) the influence of social context and relevance of rating criteria on the intergroup judgments. It has been shown that dimensions of comparison between groups and the social context (in

which the comparison is made) can significantly influence the patterns of displaying in-group favoritism and social discrimination (Mummendey, 1995; Reynolds, Turner, & Haslam, 2000) as well as out-group stereotyping (Haslam, Turner, Oakes, McGarty, & Hayes, 1992). Taking these factors into account might help to understand the differences in results between the scale pairs and develop more reliable sets of scales for use in specific social contexts. Reducing the differences between the scale pairs would allow constructing reliable indexes for evaluative and descriptive components across all pairs.

We believe that the method outlined above can be applied to studies of stereotype accuracy (cf. Judd & Park, 1993; Oakes & Reynolds, 1997) and consensus (cf. Haslam, 1997). These are issues of great social relevance that continue to attract the interest of psychologists. We hope that our research will encourage other researchers to re-visit the work of Peabody and to further contribute to the development of his method and its integration into modern social psychology.

References

Brewer, M. B., & Gardner, W. (1996). Who is this "we"? Levels of collective identity and self representations. *Journal of Personality and Social Psychology, 71,* 83-93.

Campbell, D. T. (1967). Stereotypes and the perception of group differences. *American Psychologist, 22,* 817-829.

Diekman, A. B., Eagly, A. H., & Kulesa, P. (2002). Accuracy and bias in stereotypes about the social and political attitudes of women and men. *Journal of Experimental Social Psychology, 38,* 268-282.

Esses, V. M., Haddock, G., & Zanna, M. P. (1994). The role of mood in the expression of intergroup stereotypes. In M. P. Zanna & J. M. Olson (Eds.), *The psychology of prejudice: The Ontario Symposium* (Vol. 7, pp. 77-101). Hillsdale, NJ: Lawrence Earlbaum Associates

Esses, V. M., & Zanna, M. P. (1995). Mood and the expression of ethnic stereotypes. *Journal of Personality and Social Psychology, 69,* 1052-1068.

Fiske, S. T. (2000). Stereotyping, prejudice, and discrimination at the seam between the centuries: Evolution, culture, mind and brain. *European Journal of Social Psychology, 30,* 299-322.

Galinsky, A. D., & Moskowitz, G. B. (2000). Perspective-taking: Decreasing stereotype expression, stereotype accessibility, and in-group favoritism. *Journal of Personality and Social Psychology, 78,* 708-724.

Hardin, C. D., & Higgins, E. T. (1996). Shared reality: How social verification makes the subjective objective. In R. M. Sorrentino & E. T. Higgins

(Eds.), *Handbook of motivation and cognition: Interpersonal context* (Vol. 3, pp. 28-84). New York: Guilford.

Haslam, S. A. (1997). Stereotyping and social influence: Foundations of stereotype consensus. In R. Spears, P. J. Oakes, N. Ellemers, & S. A. Haslam (Eds.), *The social psychology of stereotyping and group life* (pp. 119-143). Oxford: Blackwell.

Haslam, S. A., Turner, J. C., Oakes, P. J., McGarty, C., & Hayes, B. K. (1992). Context-dependent variation in social stereotyping 1: The effects of intergroup relations as mediated by social change and frame of reference. *European Journal of Social Psychology, 22,* 3-20.

Hewstone, M., Rubin, M., & Willis, H. (2002). Intergroup bias. *Annual Review of Psychology, 53,* 575-604.

Judd, C. M., & Park, B. (1993). Definition and assessment of accuracy in social stereotypes. *Psychological Review, 100,* 109-128.

Keltner, D., & Robinson, R. J. (1997). Defending the status quo: Power and bias in social conflict. *Personality and Social Psychology Bulletin, 23,* 1066-1077.

Kemdal, A. B., & Montgomery, H. (2001). Explaining own and others' behavior in a controversial issue: Animal experimentation. *Journal of Social Psychology, 141,* 693-713.

Maass, A., Castelli, L., & Arcuri, L. (2000). Measuring prejudice: Implicit versus explicit techniques. In D. Capozza, & R. Brown (Eds.), *Social identity processes: Trends in theory and research* (pp. 96-116). London: Sage.

McAndrew, F. T., Akande, A., Bridgstock, R., Mealey, L., Gordon, S. C., Scheib, J. E., Akande-Adetoun, B. E., Odewale, F., Morakinyo, A., Nyahete, P., & Mubvakure, G. (2000). A multicultural study of stereotyping in English-speaking countries. *Journal of Social Psychology, 140,* 487-502.

Montgomery, H. (1994). Towards a perspective theory of decision making and judgement. *Acta Psychologica, 87,* 155-178.

Montgomery, H., & Skaldeman, P. (2001). *Shared realities and private worlds in more and less happy couples.* Unpublished manuscript. Department of Psychology, Stockholm University.

Mullen, B., Brown, R., & Smith, C. (1992). In-group bias as a function of salience, relevance, and status: An integration. *European Journal of Social Psychology, 22,* 103-122.

Mummendey, A. (1995). Positive distinctiveness and social discrimination: An old couple living in divorce. *European Journal of Social Psychology, 25,* 657-670.

Oakes, P.J., & Reynolds, K.J. (1997). Asking the accuracy question: Is measurement the answer? In R. Spears, P. J. Oakes, N. Ellemers, & S. A.

Haslam (Eds.), *The social psychology of stereotyping and group life* (pp. 51-71). Oxford: Blackwell.

Peabody, D. (1968). Group judgements in the Philippines: Evaluative and descriptive aspects. *Journal of Personality and Social Psychology, 10,* 290-300.

Peabody, D. (1970). Evaluative and descriptive aspects in personality perception: A reappraisal. *Journal of Personality and Social Psychology, 16,* 639-646.

Peabody, D. (1984). Personality dimensions through trait inferences. *Journal of Personality and Social Psychology, 46,* 384-403.

Peabody, D. (1985). *National characteristics.* Cambridge: Cambridge University Press.

Peabody, D., & Goldberg, L. R. (1989). Some determinants of factor structures from personality-trait descriptors. *Journal of Personality and Social Psychology, 57,* 552-567.

Peabody, D., & Shmelyov, A. G. (1996). Psychological characteristics of Russians. *European Journal of Social Psychology, 26,* 507-512.

Pronin, E., Puccio, C., & Ross, L. (2002). Understanding misunderstanding: Social psychological perspectives. In T. Gilovich, D. Griffin, & D. Kahnemen (Eds.), *Heuristics and biases: The psychology of intuitive judgement* (pp. 636-665). Cambridge: Cambridge University Press.

Reynolds, K. J., Turner, J. C., & Haslam, S. A. (2000). When are we better than them and they worse than us? A closer look at social discrimination in positive and negative domains. *Journal of Personality and Social Psychology, 78,* 64-80

Robinson, R. J., & Friedman, R. A. (1995). Mistrust and misconstrual in union-management relationships: causal accounts in adversarial contexts. *International Journal of Conflict Management, 3,* 312-327.

Robinson, R. J., Keltner, D., Ward, A., & Ross, L. (1995). Actual versus assumed differences in construal: "Naïve realism" in intergroup perception and conflict. *Journal of Personality and Social Psychology, 68,* 404-417.

Rouhana, N. N., O'Dwyer, A., & Morrison Vaso, S. K. (1997). Cognitive biases and political party affiliation in intergroup conflict. *Journal of Applied Social Psychology, 27,* 37-57.

Rustemli, A., Mertan, B., & Ciftci, O. (2000). In-group favouritism among native and immigrant Turkish Cypriots: Trait evaluations of in-group and out-group targets. *Journal of Social Psychology, 140,* 26-34.

Saenger, G., & Flowerman, S. (1954). Stereotypes and prejudicial attitudes. *Human Relations, 7,* 217-238.

Saucier, G. (1994). Separating description and evaluation in the structure of personality attributes. *Journal of Personality and Social Psychology, 66,* 141-154.

Scaillet, N., & Leyens, J.-P. (2000). From incorrect deductive reasoning to ingroup favouritism. In D. Capozza, & R. Brown (Eds.), *Social identity processes: Trends in theory and research* (pp. 49-61). London: Sage.

Smedley, J. W., & Bayton, J. A. (1978). Evaluative race-class stereotypes by race and perceived class of subjects. *Journal of Personality and Social Psychology, 36,* 530-535.

Stephan, W. G., Ageyev, V., Stephan, C. W., Abalakina, M., Stefanenko, T., & Coates-Shrider, L. (1993). Measuring stereotypes: A comparison of methods using Russian and American samples. *Social Psychology Quarterly, 56,* 54-64.

Tajfel, H. (1981). Social stereotypes and social groups. In: J.C. Turner, & H. Giles (Eds.), *Intergroup Behavior* (pp. 144-167). Oxford: Blackwell.

Appendix A. Mean ratings and standard deviations on all scales for Latvian and Russian judges

| Scale | Target: Latvian | | | |
| | Latvian judges | | Russian judges | |
	Mean	SD	Mean	SD
a Inhibited (-) – Spontaneous (+)	3.27	1.64	4.67	2.69
b Impulsive (-) – Self-controlled (+)	6.04	1.85	5.60	2.70
a Doesn't care about others (-) – Cares about others (+)	4.95	1.91	3.56	1.67
b Interferes in others affairs (-) – Leaves others in peace (+)	5.01	2.49	4.55	2.39
a Stingy (-) – Generous (+)	4.23	1.91	3.04	1.84
b Extravagant (-) – Thrifty (+)	6.76	1.67	6.42	1.96
a Timid (-) – Bold (+)	4.23	2.02	3.66	2.05
b Rash (-) – Cautious (+)	5.90	1.95	6.17	2.35
a Devious (-) – Frank (+)	4.79	2.04	3.03	1.77
b Tactless (-) – Tactful (+)	5.74	1.80	5.17	2.41
a Self-destructive (-) – Self affirmative (+)	5.26	1.48	4.97	2.14
b Egoistic (-) – Altruistic (+)	3.95	1.57	2.87	1.55
a Passive (-) – Active (+)	4.82	1.86	4.05	2.13
b Agitated (-) – Calm (+)	5.60	1.99	5.70	2.41
a Cold (-) – Warm (+)	5.65	1.99	3.40	2.01
b Disorganized (-) – Organized (+)	5.25	1.89	4.67	2.07
a Conforming (-) – Independent (+)	4.39	2.12	4.54	2.40
b Uncooperative (-) – Cooperative (+)	4.94	1.73	4.12	2.03
a Grim (-) – Gay (+)	5.61	5.71	4.40	2.28
b Frivolous (-) – Serious (+)	6.72	1.59	5.36	2.27
a Boring (-) – Charming (+)	5.45	1.90	4.79	2.10
b Unreliable (-) – Reliable (+)	5.70	1.61	4.09	2.43
a Opportunistic (-) – Idealistic (+)	5.04	1.57	6.01	2.14
b Impractical (-) – Practical (+)	6.51	1.99	5.41	2.34
a Vacillating (-) – Persistent (+)	5.14	2.11	4.32	2.04
b Inflexible (-) – Flexible (+)	4.75	1.80	3.99	1.83

| Scale | Target: Russian | | | |
| | Latvian judges | | Russian judges | |
	Mean	*SD*	*Mean*	*SD*
a Inhibited (-) – Spontaneous (+)	6.65	1.74	5.83	2.26
b Impulsive (-) – Self-controlled (+)	2.96	1.67	4.46	2.26
a Doesn't care about others (-) – Cares about others (+)	5.29	5.58	6.21	2.03
b Interferes in others affairs (-) – Leaves others in peace (+)	4.04	2.27	4.97	1.88
a Stingy (-) – Generous (+)	5.61	1.83	6.58	1.95
b Extravagant (-) – Thrifty (+)	4.40	1.80	4.30	2.35
a Timid (-) – Bold (+)	7.34	1.67	6.78	1.78
b Rash (-) – Cautious (+)	5.05	1.99	5.36	2.12
a Devious (-) – Frank (+)	5.37	2.13	6.25	1.93
b Tactless (-) – Tactful (+)	4.24	1.73	5.59	1.85
a Self-destructive (-) – Self affirmative (+)	6.59	1.58	6.20	1.95
b Egoistic (-) – Altruistic (+)	4.43	1.77	5.17	2.11
a Passive (-) – Active (+)	6.42	1.95	6.91	1.99
b Agitated (-) – Calm (+)	3.45	1.55	4.71	1.94
a Cold (-) – Warm (+)	5.68	2.03	7.12	1.74
b Disorganized (-) – Organized (+)	5.20	2.07	5.38	2.10
a Conforming (-) – Independent (+)	6.24	1.95	5.86	2.09
b Uncooperative (-) – Cooperative (+)	5.67	1.97	6.64	2.00
a Grim (-) – Gay (+)	6.50	1.76	6.86	1.89
b Frivolous (-) – Serious (+)	4.82	1.77	5.39	2.16
a Boring (-) – Charming (+)	5.08	1.73	6.54	1.76
b Unreliable (-) – Reliable (+)	4.77	1.88	6.26	1.80
a Opportunistic (-) – Idealistic (+)	4.90	1.60	4.82	2.11
b Impractical (-) – Practical (+)	5.63	1.79	6.30	1.90
a Vacillating (-) – Persistent (+)	6.78	1.91	6.74	1.87
b Inflexible (-) – Flexible (+)	6.23	2.02	5.64	1.98

Note. For all scales 1 stands for the negative end of the continuum and 9 for the positive end. Adjacent scales are from the same pair. Rating *b* was subtracted from rating *a* when calculating the descriptive scores.

Appendix B. Mean ratings and standard deviations on all scales for male and female judges from own, male and female perspectives

| Scale | Target: "Male psychology student" | | | | | |
| | Male judges | | | Female judges | | |
	Own	Male	Female	Own	Male	Female
a Inhibited –	5.59	5.58	5.64	5.48	6.29	5.63
Spontaneous	(1.31)	(1.33)	(1.38)	(1.23)	(1.45)	(1.42)
b Impulsive –	6.30	6.54	6.19	6.07	6.62	6.26
Self-controlled	(1.30)	(1.07)	(1.13)	(1.61)	(1.61)	(1.85)
a Doesn't care about	4.70	4.96	5.19	5.35	6.55	5.93
others – Cares about o.	(1.59)	(1.56)	(1.42)	(1.88)	(1.58)	(1.78)
b Interferes in o. affairs	5.41	5.58	5.50	5.47	5.57	5.58
– Leaves others in peace	(1.87)	(1.42)	(1.48)	(1.45)	(1.60)	(1.37)
a Stingy – Generous	5.41	5.62	5.58	5.26	6.17	5.55
	(1.01)	(1.02)	(1.03)	(1.18)	(1.45)	(1.27)
b Extravagant – Thrifty	5.56	5.35	5.15	5.28	5.49	5.30
	(1.31)	(1.26)	(1.19)	(1.24)	(1.58)	(1.44)
a Timid – Bold	5.48	5.85	5.54	5.33	6.17	5.51
	(1.19)	(1.32)	(1.53)	(1.29)	(1.58)	(1.52)
b Rash – Cautious	5.37	5.23	4.54	5.37	5.62	5.67
	(1.21)	(1.18)	(1.36)	(1.38)	(1.34)	(1.60)
a Devious – Frank	5.37	5.69	5.54	5.49	5.60	6.19
	(1.62)	(1.72)	(1.53)	(1.62)	(1.81)	(1.64)
b Tactless – Tactful	5.67	5.62	5.27	6.12	6.38	6.33
	(1.44)	(1.33)	(1.34)	(1.37)	(1.50)	(1.59)
a Self-destructive –	6.44	6.27	6.31	5.91	5.95	6.16
Self affirmative	(1.28)	(0.87)	(1.05)	(1.54)	(1.56)	(1.48)
b Egoistic – Altruistic	4.63	4.35	4.73	4.93	5.64	5.35
	(1.25)	(1.20)	(1.08)	(1.22)	(1.69)	(1.53)
a Passive – Active	5.82	5.77	5.89	5.98	6.36	5.61
	(1.36)	(1.42)	(1.51)	(1.46)	(1.83)	(1.72)
b Agitated – Calm	5.89	5.96	6.24	6.26	6.36	6.16
	(1.70)	(1.51)	(1.39)	(1.65)	(1.62)	(1.75)
a Cold – Warm	5.48	5.50	5.62	5.74	6.07	6.14
	(1.22)	(1.14)	(1.39)	(1.63)	(1.50)	(1.66)
b Disorganized –	5.29	5.00	4.80	4.72	5.60	5.05
Organized	(1.20)	(1.50)	(1.92)	(1.47)	(1.88)	(1.75)

| Scale | Target: "Female psychology student" | | | | | |
| | Male judges | | | Female judges | | |
	Own	Male	Female	Own	Male	Female
a Inhibited –	4.56	4.85	5.27	5.56	4.83	5.61
Spontaneous	(1.40)	(1.46)	(1.64)	(1.30)	(1.41)	(1.45)
b Impulsive –	6.15	5.93	5.96	6.72	6.02	6.19
Self-controlled	(1.63)	(1.66)	(1.89)	(1.33)	(2.09)	(1.82)
a Doesn't care about	5.85	6.15	5.89	6.26	6.69	6.81
others – Cares about o.	(1.56)	(1.54)	(2.01)	(2.00)	(1.80)	(1.76)
b Interferes in o. affairs	4.74	4.44	4.62	5.05	4.41	4.77
– Leaves others in peace	(1.23)	(1.48)	(1.88)	(1.70)	(1.77)	(1.46)
a Stingy – Generous	5.41	5.82	5.73	5.79	6.00	5.88
	(1.22)	(0.92)	(1.37)	(1.37)	(1.41)	(1.35)
b Extravagant – Thrifty	5.44	5.56	5.69	5.21	5.45	5.67
	(1.16)	(1.31)	(1.49)	(1.39)	(1.78)	(1.44)
a Timid – Bold	5.00	4.78	5.46	5.42	4.81	5.14
	(1.41)	(1.50)	(1.53)	(1.18)	(1.07)	(1.13)
b Rash – Cautious	5.96	6.15	6.19	5.67	6.10	6.35
	(1.16)	(1.46)	(1.23)	(1.49)	(1.76)	(1.45)
a Devious – Frank	4.85	4.74	5.23	5.63	5.64	5.91
	(1.61)	(1.61)	(1.99)	(1.73)	(1.64)	(1.80)
b Tactless – Tactful	5.93	6.59	6.04	6.07	6.91	6.79
	(1.66)	(1.34)	(1.51)	(1.98)	(1.34)	(1.52)
a Self-destructive –	6.26	5.85	6.19	6.05	6.10	6.09
Self affirmative	(1.43)	(1.32)	(1.47)	(1.69)	(1.62)	(1.65)
b Egoistic – Altruistic	4.96	5.30	5.31	5.65	5.76	5.70
	(1.32)	(1.27)	(1.69)	(1.46)	(1.46)	(1.26)
a Passive – Active	5.96	5.52	6.08	5.91	5.02	5.93
	(1.72)	(1.89)	(1.38)	(1.43)	(1.75)	(1.75)
b Agitated – Calm	5.04	4.82	5.19	5.91	5.52	5.54
	(1.51)	(1.42)	(1.74)	(1.78)	(1.95)	(1.93)
a Cold – Warm	5.82	5.82	5.96	6.33	6.43	6.61
	(1.47)	(1.50)	(1.56)	(1.39)	(1.58)	(1.61)
b Disorganized –	6.37	6.44	6.46	6.12	6.29	6.33
Organized	(1.57)	(1.31)	(1.45)	(1.79)	(2.02)	(1.86)

Note. Standard deviations are indicated in parentheses. For all scales, 1 stands for the negative end of the continuum and 9 for the positive end. Adjacent scales are from the same pair. Rating *b* was subtracted from rating *a* when calculating the descriptive scores

STUDY 2

THE FALSE POLARIZATION EFFECT IN EXPLANATIONS OF ATTITUDINAL BEHAVIOR

Girts Dimdins
Stockholm University

Henry Montgomery
Stockholm University

Ivars Austers
University of Latvia

ABSTRACT
Supporters and opponents of Latvia's EU membership rated attitudinal behavior of EU supporters and opponents on a number of causal explanation scales from their own perspective and simulated perspectives of both groups. From a target's in-group perspective, both groups rated the causes of behavior as more stable, controllable and rational, and less subjective and less influenced by others than the respective group did from their own perspective. Although the results were not consistent for all rating scales and all perspectives, the study demonstrates that a false polarization effect occurs in explanations of attitudinal behavior. Directions for future research are discussed.

Key words: Intergroup attribution, group attribution, false polarization effect, perspective taking

THEORETICAL BACKGROUND
Lee Ross (1977) notes that "Individuals must, for the most part, share a common understanding of the social actions and outcomes that affect them, for without such consensus, social interaction would be chaotic, unpredictable, and beyond the control of the participants" (p. 179). For the past several decades, psychologists and other researchers have focused on the construal of social reality (see Hewstone, 1989 for a review). A significant portion of these studies deals with causal attribution in an intergroup context (Hewstone, 1990; Deschamps, 1983). These studies consistently have found that individuals prefer in-group-serving attributions relative to out-group-serving attributions (Hewstone, 1990).

The definition of "group-serving" has varied from study to study, depending on which attributes are used. The oldest classification of intergroup attributions,

based on Heider's (1958) model of interpersonal attributions, distinguishes between internal (dispositional) and external (situational) attributions. The basic hypothesis is that more internal attributions would be made for in-group members' socially desirable behavior and more external attributions for socially undesirable behavior. The opposite should hold true for out-group members' behaviors. Several classic studies using this classification (Taylor & Jaggi, 1974; Duncan, 1976; Rosenberg & Wolfsfeld, 1977; Hewstone & Ward, 1985) have provided mixed support for this hypothesis. The classification itself has been criticized for several methodological drawbacks (see Miller, Smith, & Uleman, 1981).

An improved model was proposed by Islam and Hewstone (1993), who suggest measuring attributions along the continuums of causal locus, stability, controllability (Weiner, 1986), and globality (Abramson, Seligman, & Teasdale, 1978). The hypothesis for this model is that positive in-group behavior and negative out-group behavior would be attributed to internal, stable, controllable (by the actor), and global causes; negative in-group and positive out-group behavior should be attributed to more external, unstable, uncontrollable, and specific causes. A number of successful attribution studies have used this classification (Islam & Hewstone, 1993; Wilder, Simon & Faith, 1996; Lee & Robinson, 2000; Austers, 2002). Again, the support for the hypothesis has been mixed.

Malle (1999, 2001) suggests another alternative to the internal/external classification. Central to the model of folk explanations of behavior (FEB) is the folk concept of intentionality (Malle & Knobe, 1997). Within the model, explanations that interpret an agent's behavior as unintentional are considered *cause* explanations. Explanations that interpret an agent's behavior as intentional are divided into three major groups. *Causal history of reasons* contains factors from an agent's personal history that cause the intentional behavior without the agent being aware of them. *Reasons* are factors that the agent considered when forming the intention to act, and *enabling factors* are factors that clarify how it was possible that the agent completed the intended action. Reasons are further divided into *desires*, *beliefs*, and *valuing*. Whereas the FEB model has been used mostly in explanations of individual behavior (Malle, Knobe, O'Laughlin, Pearce, & Nelson, 2000), it can also be successfully applied in analyzing the explanations of group behavior (O'Laughlin & Malle, 2002).

A number of studies of group attributions have been carried out that ask the participants to take the perspective of their in-group or out-group members. Kemdal and Montgomery (2001) found that animal experimenters and animal rights activists could take each other's perspective, resulting in a reversed actor-observer effect. Austers (2002) and Austers and Montgomery (2001) found

similar results. Robinson, Keltner, Ward, and Ross (1995) tested some aspects of group attribution when they asked their participants to rate the basis of their own political judgments and that of their in-group and out-group members. The respondents indicated that they personally had been less influenced by ideology or political orientation than either their peers or opponents. These findings suggest a presence of the false polarization effect in the context of intergroup attributions.

The false polarization effect (see Pronin, Puccio, & Ross [2002] for a review of this phenomenon) is better known from the studies of group attitudes. Generally, the effect can be defined as an overestimation of the expected group score on a certain attribute in comparison with the actual group score on that attribute, which has been found across a number of different contexts (Robinson, Keltner, Ward, & Ross, 1995; Keltner & Robinson, 1997; Robinson & Friedman, 1995; Rouhana, O'Dwyer, & Morrison Vaso, 1997; Diekman, Eagly, & Kulesa, 2002). These findings show that the false polarization effect is a very robust phenomenon, influencing various types of ratings. There is no reason why the same effect should not appear in attribution studies, but very few published papers address this issue. Austers (2002) asked ethnic Latvian and Russian schoolteachers to rate positive and negative behaviors from their own and out-group perspective on a number of attributional dimensions. The results indicated some false polarization effect in predicting out-group's responses, but not consistently. To our knowledge, the paper of Robinson et al. (1995) provides the most convincing evidence of the false polarization effect in the context of group attributions. However, in their study the effect is reported on a small number of context-specific dimensions. We designed this study to explore the presence of the false polarization effect on more general attributional dimensions, including "traditional" dimensions familiar from the previous research.

This study tests whether the false polarization effect would be present in explanations of causes of attitudinal behavior when these explanations are given from in-group and out-group perspectives. By attitudinal behavior we understand general, non-specific behavior that expresses the actor's attitude towards a certain issue, in other words--acting in line with one's attitudes. We chose attitudinal behavior instead of more specific group behaviors because it allowed us to construct simple, context-independent and unambiguous stimuli for our study.

In our study we tried to use all the attributional dimensions reported in earlier literature. In addition to the four variables of causal locus, controllability, stability, and globality, we included a number of explanatory dimensions taken from Malle's FEB model. The items constructed on the basis of the FEB model

3

represented factors differentiating among the major groups of explanations within the model, as well as various types of explanations within each group. The differentiating factors were intentionality (distinguishing cause explanations from all other explanations) and awareness (distinguishing causal history of reasons from reason explanations). The specific explanations included in the questionnaire were various types of causal history of reasons and reason explanations. (In an earlier unpublished study, where we asked participants to give free-response explanations of attitudinal behavior, more than 95% of the explanations fell into these two categories.)

We included two items measuring the evaluation of the target behavior in positive-negative terms, to control whether the group serving bias occurs in the expected direction (i.e., each target behavior is seen as more positive from its in-group perspective than out-group perspective). Finally, we also included an item asking the participants to estimate the distribution of people sharing their opinions and those with opposite opinions in the society. This was another control question because stronger in-group favoritism can be expected among the members of minority groups (Hewstone, Rubin, & Willis, 2002; Mullen, Brown, & Smith, 1992).

We predicted that when imagining how specific causes for a given target behavior are rated from the target's in-group and out-group perspectives, participants would give significantly higher (or lower) ratings than the respective in-group and out-group members themselves. Theoretically, the mean differences should occur in the group-serving direction. However, because we were using general descriptions of two groups' attitudinal behavior rather than descriptions of positive and negative behavior, it was difficult to make specific predictions about the direction of mean differences. Although one would expect that each target behavior should be seen more favorably from the target's in-group perspective than the out-group perspective, it does not necessarily mean that in-group behavior is seen as explicitly positive, and out-group behavior as negative. In other words, from the previously published group attribution studies we could not predict with certainty how in-group favoritism should manifest itself on each of the attributional variables. Our general prediction was that the mean differences from opposing perspectives should be in opposite directions for the same behavior. For example, if the causes for supporter behavior are rated as highly stable from the supporter perspective (implying that stability is seen as a favorable attribute) they should be rated as relatively unstable from the opponent perspective, and vice versa.

After a pilot study of several topics eliciting political attitudes among Latvian students, we decided that Latvia's membership in the European Union is a controversial political issue worthy of study. The topic provides an excellent

4

context for examining perspective taking in explanations of group behavior. Over the last two years, the debate on the issue has grown with the opinions of both supporters and opponents highly salient and well represented in mass media. Although traditionally EU membership supporters have been in a slight majority in Latvia, the opinion polls predict a close race in the referendum planned for autumn 2003.

METHOD

Sample. One hundred forty-one students (110 women) at the University of Latvia participated in the study. The participants were undergraduate students of education. The mean age of the participants was 19 years. Majority of respondents (82%) were ethnic Latvians.

Questionnaire. First, we asked the participants to indicate their age, gender, ethnicity, and general attitude towards joining the European Union (pro vs. against). We also asked whether they believed (a) that EU supporters were in the considerable majority in Latvia, (b) that there was approximately equal number of opponents and supporters, or (c) that the opponents were in the majority.

In the second part of the questionnaire, we provided descriptions of two opposite behaviors and asked participants to rate various explanations of these behaviors from their own perspective as well as from the perspective of both EU supporters and EU opponents. The opposite behaviors were described with the following statements: (a) *"There are many people in Latvia who actively support Latvia entering the European Union"*; (b) *"There are many people in Latvia who actively protest against Latvia entering the European Union"*. Each description was written on top of a separate sheet.

We asked the participants to rate the causes of both target behaviors from their own perspective, EU opponent, and EU supporter perspective by answering a number of questions. All but the last two questions offered various explanations for the behavior, and they were constructed to cover both "traditional" attribution dimensions and the FEB coding scheme. The survey items are shown in Appendix A. The participants rated to what extent in their opinion each explanation accounted for the target behavior. The last two questions asked the participants to evaluate the target behavior in positive-negative terms. All ratings were made on a five-point Likert scale.

Correspondingly, for group perspectives, half of the participants were asked the following: *"How would people who support (oppose) Latvia entering the European Union answer the question . . .".* For the other half, the following question was asked: *"How would University of Latvia students who support*

5

(oppose) Latvia entering the European Union answer the question . . . ". The two different wordings were used to control for the possible effects of how the group perspectives were defined.

The order of items and target behaviors was counterbalanced across the questionnaires. The participants first rated each target behavior from their own perspective. Then they rated each target from both group perspectives in random order. To reiterate, each participant made ratings of two target behaviors from three perspectives, answering the same set of questions six times.

In the third part of the questionnaire, we asked the participants to indicate their agreement/disagreement with seven items measuring their support to EU membership (see Appendix A). Afterwards, the participants rated the same items from both group perspectives.

RESULTS

First, we checked for the effects of questionnaire type (wording for the group perspective: supporters/opponents in general vs. LU students) and the effects of perceived balance (supporters in majority vs. opponents in majority vs. both groups equal). We included each of these variables as a between-subjects factor in a 4-way ANOVA with participant attitude (supporter vs. opponent) as another between-subjects variable and target behavior (supporter vs. opponent) and perspective (self vs. supporter vs. opponent) as within-subjects variables. We ran the ANOVA on all variables in our study both on raw data and the indexes reported below. We found some significant effects involving both variables, but the directions of mean differences varied from analysis to analysis. Because the effects of both variables did not seem to be systematic, we do not report them.

We based our data analysis on planned pairwise comparisons between the rating means from own perspective and the corresponding means from both group perspectives. We used a one-tailed t-test where we had clear predictions about the directions of mean differences (i.e., attitude ratings) and a 2-tailed t-test for the other comparisons (i.e., causal explanation ratings). Such a direct test of the false polarization effect increased the power of the analysis in comparison with a full ANOVA. For each comparison, we computed effect size (Cohen's d) to provide an estimate of the magnitude of mean differences. According to conventional estimates suggested by Cohen (1988), an effect size of about 0.2 can be seen as small, an effect size of about 0.5 as medium, and an effect size of 0.8 or higher as large.

Ninety-two participants (65%) identified themselves as supporters of Latvia's EU membership, and forty-nine (35%) identified themselves as opponents. The top row of Table 1 shows the means of strength of support to EU membership from own and both imagined group perspectives. The index was calculated as the average of the seven items in the third part of the questionnaire. The index values can vary from 1 to 5, higher score corresponding to stronger support to Latvia's EU membership. The false polarization effect is clearly visible in the results. Both groups overestimated the support displayed by EU supporters and the opposition displayed by EU opponents.

Table 1

Mean attitude ratings from own perspective and the imagined group perspectives

	Ratings from own perspective		Ratings from the supporter perspective		Ratings from the opponent perspective	
	By supporters ($n = 92$)	By opponents ($n = 49$)	By supporters	By opponents	By supporters	By opponents
Support to EU member-ship	3.88 (0.55)	2.32 (0.56)	4.30*** (0.61)	4.29*** (0.65)	1.91** (0.85)	1.94** (0.74)
Effect size			0.72	0.68	0.56	0.57
Evaluation of target behavior:						
Supporter behavior	3.53 (0.67)	2.77 (0.76)	4.13*** (0.73)	3.96*** (0.87)	1.96*** (0.82)	2.36** (0.91)
Effect size			0.85	0.55	1.02	0.54
Opponent behavior	2.63 (0.81)	3.52 (0.82)	1.99*** (0.86)	2.13*** (0.81)	3.85* (1.06)	3.80 (0.97)
Effect size			0.76	0.62	0.34	0.31

Note. Standard deviations are given in parentheses. Asterisks indicate that rating from the imagined group perspective is significantly different (1-tailed t-test) from the corresponding rating from own perspective. Means in columns 3 and 4 are compared to means in column 1; means in columns 5 and 6 are compared to means in column 2. The effect size (Cohen's d) for each comparison is given below the respective mean in columns 3-6. * $p < .05$. ** $p < .01$. *** $p < .001$.

We calculated an index measuring the evaluation of target behavior as the average of two items: liking/disliking the behavior, and the belief that the behavior benefits the interests of the Latvian society. Cronbach's Alphas for

both items were calculated separately for each perspective and target behavior (because each of target/perspective combinations represents a separate repeated measurement using the same two items). All but one (0.54) were above 0.6. Table 1 shows the means of the evaluative index. The index value can range from 1 (unfavorable evaluation) to 5 (favorable evaluation). Again, both groups showed a strong false polarization effect when evaluating supporter behavior from both perspectives and opponent behavior from the supporter perspective. The mean difference was not significant when opponent behavior was evaluated from the opponent perspective, but the means were in the expected direction.

Our study replicated the previous findings showing a false polarization effect when estimating group attitudes. Our main interest, however, was whether the same pattern of results would be found in estimating group ratings of causal explanations. Table 2 shows the mean ratings of stability of causes and ratings of actors' control over the causes of their behavior. All ratings were made on a 1 to 5 Likert type scale; larger values correspond to higher stability and control.

For two dimensions--causes in situation vs. actor and the globality of the causes of target behavior--we found no significant mean differences. Therefore the means of these two variables are not shown in Table 2 (but the means are reported in Table B3 and Table B6 in Appendix B). The false polarization effect appears in estimations of stability of causes of behavior. The strongest overestimation occurs in predicting supporters' ratings of supporter behavior where both groups rated the causes as significantly more stable than the supporters did from their own perspective. In other cases, the significant mean differences are in the same direction. (From the in-group perspective, the causes are seen as more stable.) Interestingly, the responses of supporter and opponent participants themselves do not differ for supporter behavior ($t[138] = 0.12$, $p = .90$) or opponent ($t[136] = -1.83$, $p = .07$) behavior. Nevertheless, both groups expected other supporters and opponents to be biased towards their in-groups.

We found similar results also for ratings of the actors' control over the causes of their behavior. Again, the strongest overestimation of group position occurs when rating the supporter behavior from supporter perspective. For supporter behavior, both groups did not differ in their ratings from own perspective ($t[138] = -0.03$, $p = .98$). For opponent behavior, opponents gave higher control ratings than supporters did ($t[139] = -2.51$, $p < .05$). Nevertheless, the overestimation of group position was stronger and more frequent for the supporter behavior than the opponent behavior.

To summarize, of the four traditional attribution dimensions, the false polarization effect appears in the stability of causes and the actors' control of the causes of attitudinal behavior. For both ratings, the effect was stronger for

supporter behavior than for opponent behavior, and the effect was stronger from the imagined supporter perspective, than the opponent perspective.

Table 2
Mean ratings for the traditional dimensions of causal attribution from own perspective and the imagined group perspectives

	Ratings from own perspective		Ratings from the supporter perspective		Ratings from the opponent perspective	
	By supporters ($n = 92$)	By opponents ($n = 49$)	By supporters	By opponents	By supporters	By opponents
Causes are stable						
Supporter behavior	2.96	2.94	3.76***	3.84***	2.49*	3.02
	(0.91)	(0.84)	(0.97)	(0.99)	(1.04)	(0.88)
Effect size			0.85	0.92	0.47	0.09
Opponent behavior	2.87	3.17	2.59	2.86	3.59*	3.61*
	(0.90)	(0.96)	(1.00)	(1.04)	(1.16)	(0.86)
Effect size			0.29	0.01	0.39	0.48
Actors control the causes of their behavior						
Supporter behavior	3.12	3.13	3.82***	3.86***	2.53**	2.94
	(0.94)	(1.00)	(1.00)	(1.02)	(1.08)	(0.94)
Effect size			0.72	0.75	0.58	0.20
Opponent behavior	2.82	3.29	2.46*	2.84	3.64	3.59
	(1.06)	(1.06)	(0.95)	(1.12)	(1.09)	(1.22)
Effect size			0.36	0.02	0.33	0.26

Note. Standard deviations are given in parentheses. Asterisks indicate that rating from the imagined group perspective is significantly different (2-tailed t-test) from the corresponding rating from own perspective. Means in columns 3 and 4 are compared to means in column 1; means in columns 5 and 6 are compared to means in column 2. The effect size (Cohen's d) for each comparison is given below the respective mean in columns 3-6. * $p < .05$. ** $p < .01$. *** $p < .001$.

To reduce the number of variables, we ran a factor analysis on the items based on the FEB coding scheme. Sums of ratings for each item from all three perspectives for both target behaviors were entered into the analysis. A principal component analysis with Varimax rotation yielded two factors, accounting for 50% of the total variance. The results of the factor analysis are shown in Table3.

Table 3

Results of Principal Components Analysis

Item	h^2	Component 1	Component 2
Target behavior is intentional [a]	.600	.303	.713
Actors are aware of the causes of their behavior [a]	.580	-.077	.758
Target behavior is determined by care for the country's future [a]	.589	.030	.767
Target behavior is determined by actors' own interests [b]	.392	.610	.145
Target behavior is determined by actors' emotions [b]	.533	.728	-.051
Target behavior is determined by actors' desires [b]	.615	.631	.466
Target behavior is determined by actors' valuations of EU [b]	.538	.709	.187
Target behavior is determined by actors' personality traits [b]	.471	.680	.090
Target behavior is determined by the actors' group memberships [b]	.344	.572	.130
Target behavior is determined by actors' beliefs [b]	.597	.537	.556
Target behavior is determined by accepting others' opinions	.252	.432	-.256
Eigenvalues for unrotated solution		3.82	1.69
Variance explained before rotation (%)		34.73	15.38

Note. [a] These items were used to calculate the Rationality index.

[b] These items were used to calculate the Subjectivity index.

Reliability analysis was then conducted for sets of variables with high loading on the same factor. Cronbach's alpha for these sets of variables was calculated separately for ratings of each target from each perspective (because each of these ratings is a separate repeated measurement using the same set of items). Both factors were transformable into reliable indexes. The first index was calculated as an average of three variables: intentionality of the target behavior, actors' awareness of causes of their behavior, and the extent to what behavior is caused by actors' care about the country's future. We labeled this the *Rationality index* because all items are related to rational causes of behavior. All but one (0.58) Alphas for this index were above 0.60. The index values can vary from 1 to 5, higher scores corresponding to higher rationality. The second index was calculated as an average of 7 items, all of which seem to be related to actors'

inner states and psychological characteristics. Therefore, we labeled it the *Subjectivity index*. (The *belief* item, which had similar loading on both factors, was included in Subjectivity index because it increased its reliability, and it decreased the reliability of the Rationality index.) Again, all but one (0.57) Alphas for this index were above 0.60. The index values can vary from 1 to 5, higher scores corresponding to higher subjectivity. One item (extent to which the behavior is caused by accepting others' opinions) that could not be included in any of the indexes was analyzed separately. It was measured on a five-point Likert scale, where higher score indicate more influence from accepting others' opinions. Table 4 shows the corresponding means.

The pattern of means for the Rationality index resembles those discussed above. For supporter behavior, means of supporter and opponent groups did not differ significantly, $t(136) = 1.17$, $p = .25$. However, both groups displayed a false polarization effect when rating the behavior from the imagined supporter perspective, and supporter participants showed the same tendency when estimating the position of opponents. For opponent behavior, both groups differed in their ratings, $t(139) = -3.60$, $p < .001$. Three of the four group perspective means showed a false polarization effect. The results suggest that the participants saw rationality as a favorable cause of attitudinal behavior.

The pattern is less clear for the Subjectivity index. The means of supporter and opponent participants from own perspective did not differ for supporter ($t[134] = -1.38$, $p = .17$) or opponent ($t[139] = 0.09$, $p = .93$) behavior. Supporter and opponent participants saw the causes of both target behaviors as less subjective from both targets' in-group perspectives (supporter behavior from supporter perspective and opponent behavior from opponent perspective). However, no false polarization effect in the opposite direction occurred when the causes were rated from targets' out-group perspectives.

Finally, regarding acceptance of others' opinions as the cause of attitudinal behavior, there was some evidence of a false polarization effect, but the mean differences were relatively small. From own perspective, the means differed significantly for the supporter behavior ($t[136] = -2.38$, $p < .05$), but not opponent behavior ($t[136] = 0.94$, $p = .35$). With one exception, all of the mean differences between own and group perspectives were consistently in the same direction, assigning less influence by others from in-group perspective and more from out-group perspective.

Table 4

Mean ratings for explanations derived from the FEB coding scheme from own perspective and the imagined group perspectives

	Ratings from own perspective		Ratings from the supporter perspective		Ratings from the opponent perspective	
	By supporters ($n = 92$)	By opponents ($n = 49$)	By supporters	By opponents	By supporters	By opponents
"Rationality" index						
Supporter behavior	3.29	3.13	3.99***	3.86***	2.67**	2.96
	(0.78)	(0.70)	(0.83)	(0.85)	(0.91)	(0.82)
Effect size			0.87	0.70	0.57	0.22
Opponent behavior	2.94	3.45	2.68*	2.83	3.79*	3.84*
	(0.78)	(0.83)	(0.93)	(0.83)	(0.94)	(0.83)
Effect size			0.30	0.14	0.38	0.47
"Subjectivity" index						
Supporter behavior	3.68	3.81	3.46**	3.38**	3.54*	3.67
	(0.52)	(0.51)	(0.54)	(0.59)	(0.62)	(0.68)
Effect size			0.42	0.54	0.48	0.23
Opponent behavior	3.80	3.79	3.68	3.80	3.36***	3.54*
	(0.53)	(0.53)	(0.61)	(0.57)	(0.64)	(0.59)
Effect size			0.21	0.00	0.73	0.45
Accepting others' opinions						
Supporter behavior	3.13	3.55	2.75*	2.98	3.33	3.22
	(1.04)	(0.88)	(1.02)	(1.15)	(1.21)	(1.12)
Effect size			0.37	0.14	0.21	0.33
Opponent behavior	3.20	3.02	3.52*	3.29	2.64*	3.06
	(1.11)	(0.92)	(1.05)	(1.14)	(1.13)	(1.03)
Effect size			0.30	0.08	0.39	0.04

Note. Standard deviations are given in parentheses. Asterisks indicate that rating from the imagined group perspective is significantly different (2-tailed *t*-test) from the corresponding rating from own perspective. Means in columns 3 and 4 are compared to means in column 1; means in columns 5 and 6 are compared to means in column 2. The effect size (Cohen's *d*) for each comparison is given below the respective mean in columns 3-6. * $p < .05$. ** $p < .01$. *** $p < .001$.

DISCUSSION

On a number of rating dimensions for explanations of attitudinal behavior, we found significant differences between the actual ratings of the participants and their estimations of others' ratings on the same dimension. Our study confirms that the false polarization effect is a robust phenomenon, which can affect not only estimation of group attitudes but also explanations of attitudinal behavior. On average the magnitude of the false polarization effect was moderate--for most of the reported significant mean differences, the effect size was around medium, according to Cohen's (1988) estimates. However, it should be noted that we found such differences on most of the variables in our study, and that most of the time these differences were in the expected direction. This leads us to believe that, although moderate in effect size, our results are indicators of a robust psychological tendency.

At the same time, the effect was not present on all attributional dimensions. We did not find the false polarization effect in ratings of globality of causes and the person-situation attributions. Moreover, there were no group-serving effects for these two scores at all. One explanation for this finding is that in our study we did not compare positive and negative behaviors by both groups. Although all four dimensions have been shown to yield group-serving biases for comparisons of positive and negative in-group and out-group behaviors (c.f., Islam & Hewstone, 1993), the same effects may be less pronounced when explaining non-specific attitudinal behavior. At the same time, the participants showed clear false polarization effect in group-serving direction on other dimensions where group-serving biases in ratings from own perspective were not pronounced. This fact suggests another possible explanation: whereas stability and control as concepts have positive connotations, globality of causes and the person-situation distinction per se have neither positive nor negative connotations. One may speculate that it is the evaluative connotation of a rating dimension, which triggers the false polarization effects in estimating in-group and out-group ratings, and this perhaps contributes to group-serving biases in general. Asking the participants to evaluate attributional dimensions in positive-negative terms in the future studies could provide more information about the nature of the false polarization bias and about the group-serving effects in intergroup attributions in general.

From the practical point of view, our findings illustrate how the false polarization effect can result in overestimation of group differences. On a number of variables we found very small or insignificant differences between the actual ratings of supporter and opponent participants; however both groups expected much stronger differences when taking the perspective of their in-group and out-group members. If the parties in a controversial issue assume more grounds for disagreement or conflict than is actually the case, it may

impede rational discussions between the groups and cause ungrounded pessimism about the possibility of negotiations. Moreover, the false polarization effect may put additional pressure on group leaders/representatives, who may be reluctant to make concessions in negotiations, fearing discontent of their in-group members. Awareness of the false polarization effect may contribute to conflict prevention and resolution in various intergroup contexts.

In the wider context of attribution research, the results of our study provided some additional information about causal explanations of attitudinal behavior. Stability, control, and rationality were seen as more positive causes of attitudinal behavior (more often ascribed to an in-group target). Subjectivity and influence by other people were seen as less positive. These findings are similar to the results reported by Kenworthy and Miller (2002); they found that more rationality and less externality and emotionality were attributed to in-group attitudes in comparison with out-group attitudes.

Another possible direction for the future research might be using similar perspective-taking studies with specific positive and negative in-group and out-group behaviors. Such designs would allow formulating more precise hypotheses about the expected directions of mean differences and acquire clearer results.

We did not find systematic and interpretable effects of how the groups were defined in our study (supporters/opponents in general vs. student supporters/opponents). However, both definitions were relatively general, relating to a large group whose members' individual opinions were unknown to the participants. The question remains if one would find false polarization effect if the groups were relatively small, and their individual members familiar to the respondents. A study by Kemdal and Montgomery (2000) suggests that under such conditions the false polarization effect may disappear. The effect of group characteristics on the false polarization effect is another prospective direction for future research.

REFERENCES

Abramson, L. Y., Seligmann, M. E. P., & Teasdale, J. D. (1978). Learned helplessness in humans: Critiques and reformulation. *Journal of Abnormal Psychology, 18,* 317-334.

Austers, I. (2002). Ability to take an out-group's perspective in explaining positive and negative behaviors. *Scandinavian Journal of Psychology, 43,* 407-412

Austers, I., & Montgomery, H. (2001). Explaining an out-group's political behavior from four different perspectives. *Journal of Baltic Psychology, 2,* 23-33.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). New York: Academic Press.

Deschamps, J.-C. (1983). Social Attribution. In J. Jaspars, F. D. Fincham, & M. Hewstone (Eds.), *Attribution theory and research: Conceptual, developmental and social dimensions* (pp. 223-240). London: Academic Press.

Diekman, A. B., Eagly, A. H., & Kulesa, P. (2002). Accuracy and bias in stereotypes about the social and political attitudes of women and men. *Journal of Experimental Social Psychology, 38,* 268-282.

Duncan, B. L. (1976). Differential social perception and attribution of intergroup violence: Testing the lower limits of stereotyping of blacks. *Journal of Personality and Social Psychology, 34,* 590-598.

Heider, F. (1958). *The psychology of interpersonal relations.* New York: Wiley.

Hewstone, M. (1989). *Causal attribution: From cognitive processes to collective beliefs.* Oxford: Basil Blackwell.

Hewstone, M. (1990). The "ultimate attribution error"? A review of the literature on intergroup causal attribution. *European Journal of Social Psychology, 20,* 311-335.

Hewstone, M., Rubin, M., & Willis, H. (2002). Intergroup bias. *Annual Review of Psychology, 53,* 575-604.

Hewstone, M., & Ward, C. (1985). Ethnocentrism and causal attribution in Southeast Asia. *Journal of Personality and Social Psychology, 48,* 614-623.

Islam, M. R., & Hewstone, M. (1993). Intergroup attributions and affective consequences in majority and minority groups. *Journal of Personality and Social Psychology, 64,* 936-950.

Keltner, D., & Robinson, R. J. (1997). Defending the status quo: Power and bias in social conflict. *Personality and Social Psychology Bulletin, 23,* 1066-1077.

Kemdal, A. B., & Montgomery, H. (2000). *Seeing it from the other side: Perspectives and attitudes towards animal experimentation.* Unpublished manuscript, Stockholm University.

Kemdal, A. B., & Montgomery, H. (2001). Explaining own and others' behavior in a controversial issue: Animal experimentation. *Journal of Social Psychology, 141,* 693-713.

Kenworthy, J. B., & Miller, N. (2002). Attributional biases about the origins of attitudes: Externality, emotionality, and rationality. *Journal of Personality and Social Psychology, 82,* 693-707.

Lee, F., & Robinson, R. J. (2000). An attributional analysis of social accounts: Implications of playing the blame game. *Journal of Applied Social Psychology, 30,* 1853-1879.

Malle, B. F., & Knobe, J. (1997). The folk concept of intentionality. *Journal of Experimental Social Psychology, 33,* 101-121.

Malle, B. F. (1999). How people explain behavior: A new theoretical framework. *Personality and Social Psychology Review, 3,* 23-48.

Malle, B. F. (2001). Folk explanations of intentional action. In B. F. Malle, L. J. Moses, & D. A. Baldwin (Eds.), *Intentions and intentionality: Foundations of social cognition* (pp. 265-286). Cambridge, MA: MIT Press.

Malle, B. F., Knobe, J., O'Laughlin, M. J., Pearce, G. E., & Nelson, S. E. (2000). Conceptual structure and social functions of behavior explanations: Beyond person-situation attributions. *Journal of Personality and Social Psychology, 79,* 309-326.

Miller, F. D., Smith, E. R., & Uleman, J. (1981). Measurement and interpretation of situational and dispositional attributions. *Journal of Experimental Social Psychology, 17,* 80-95.

Mullen, B., Brown, R., & Smith, C. (1992). In-group bias as a function of salience, relevance, and status: An integration. *European Journal of Social Psychology, 22,* 103-122.

O'Laughlin, M. J., & Malle, B. F. (2002). How people explain actions performed by groups and individuals. *Journal of Personality and Social Psychology, 82,* 33-48.

Pronin, E., Puccio, C., and Ross, L. (2002). Understanding misunderstanding: Social psychological perspectives. In T. Gilovich, D. Griffin, & D. Kahneman (Eds.), *Heuristics and biases: The psychology of intuitive judgment* (pp. 636-665). Cambridge: Cambridge University Press.

Robinson, R. J., & Friedman, R. A. (1995). Mistrust and misconstrual in union-management relationships: causal accounts in adversarial contexts. *International Journal of Conflict Management, 3,* 312-327.

Robinson, R. J., Keltner, D., Ward, A., & Ross, L. (1995). Actual versus assumed differences in construal: "Naïve realism" in intergroup perception and conflict. *Journal of Personality and Social Psychology, 68,* 404-417.

Rosenberg, S. W., & Wolfsfeld, G. (1977). International conflict and the problem of attribution. *Journal of Conflict Resolution, 21,* 75-103.

Ross, L. (1977). The intuitive psychologist and his shortcomings: Distortions in the attribution process. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 10, pp. 173-220). New York: Academic Press.

Rouhana, N. N., O'Dwyer, A., Morrison Vaso, S. K. (1997). Cognitive biases and political party affiliation in intergroup conflict. *Journal of Applied Social Psychology, 27,* 37-57.

Taylor, D. M.., & Jaggi, V. (1974). Ethnocentrism and causal attribution in a South Indian context. *Journal of Cross-Cultural Psychology, 119,* 162-171.

Weiner, B. (1986). *An attributional theory of motivation and emotion.* New York: Springer Verlag.

Wilder, D. A., Simon, A. F., & Faith, M. (1996). Enhancing the impact of counterstereotypic information: Dispositional attributions for deviance. *Journal of Personality and Social Psychology, 71,* 276-287.

APPENDIX A. THE SURVEY INSTRUMENT

First Part of the Questionnaire (Cover Page)

By answering the questions in this survey you are taking part in a study of political attitudes carried out by Stockholm University and University of Latvia reserachers. Your responses are confidential and will be used for research purposes only.

The aim of the study is to explore how people evaluate the behavior of supporters and opponents of the European Union from their own and imagined point of view. Therefore we ask you to answer a number of similar questions several times. It is very important that you answer all the questions.

Please, tick the appropriate response or fill in the information!

Your age: [*participant instructed to indicate the age*]

Sex: M F [*participant instructed to check one*]

Ethnicity: [*participant instructed to indicate the ethnicity*]

Generally you:

support Latvia's membership of the European Union

oppose Latvia's membership of the European Union

[*participant instructed to check one*]

In your opinion:

EU supporters in Latvia are in a significant majority; EU opponents are in a minority

there are approximately equal number of EU supporters and opponents in Latvia

EU opponents in Latvia are in a significant majority; EU supporters are in a minority

[*participant instructed to check one*]

In the rest of the questionnaire, please, answer the questions by circling the response that corresponds to your opinion!
[*end of cover page*]

Second Part of the Questionnaire

Description of supporter behavior:

There are many people in Latvia who actively support Latvia entering the European Union

Description of opponent behavior:

There are many people in Latvia who actively protest against Latvia entering the European Union

Note. In the second part of the questionnaire, each of the target behaviors was rated on 17 items, provided below. Each target behavior appeared three times in the questionnaire, and correspondingly was rated from three perspectives: from own perspective, and from two imagined group perspectives--supporter perspective and opponent perspective.

Items for rating causal explanations and evaluating target behavior from own perspective:

In your opinion, to what extent do the causes of the supporters'/opponents' behavior originate in the situation rather than within themselves?
[*participant instructed to indicate the response on a 5-point scale where* 1 = not in the least *and* 5 = to the highest extent]

In your opinion, to what extent are the causes of the supporters'/opponents' behavior stable (invariable)?
[1 = not in the least; 5 = to the highest extent]

In your opinion, to what extent do the supporters/opponents control their behavior?
[1 = not in the least; 5 = to the highest extent]

In your opinion, to what extent do the causes of the supporters'/opponents' behavior influence their behavior in all situations (also in those not related to Latvia's EU membership)?
[1 = not in the least; 5 = to the highest extent]

In your opinion, to what extent is the supporters'/opponents' behavior intentional?
[1 = not in the least; 5 = to the highest extent]

In your opinion, to what extent are the supporters/opponents aware of the causes of their behavior?
[1 = not in the least; 5 = to the highest extent]

In your opinion, to what extent are the causes of the supporters'/opponents' behavior determined by their personality traits?
[1 = not in the least; 5 = to the highest extent]

In your opinion, to what extent are the causes of the supporters'/opponents' behavior determined by their own interests?
[1 = not in the least; 5 = to the highest extent]

In your opinion, to what extent are the causes of the supporters'/opponents' behavior determined by their membership in various social groups?
[1 = not in the least; 5 = to the highest extent]

In your opinion, to what extent are the causes of the supporters'/opponents' behavior determined by their care about the country's future?
[1 = not in the least; 5 = to the highest extent]

In your opinion, to what extent are the causes of the supporters'/opponents' behavior determined by accepting other people's opinions?
[1 = not in the least; 5 = to the highest extent]

In your opinion, to what extent are the causes of the supporters'/opponents' behavior determined by their beliefs?
[1 = not in the least; 5 = to the highest extent]

In your opinion, to what extent are the causes of the supporters'/opponents' behavior determined by their desires (aspiration for a specific result)?
[1 = not in the least; 5 = to the highest extent]

In your opinion, to what extent are the causes of the supporters'/opponents' behavior determined by their subjective liking or disliking for various aspects of the EU?
[1 = not in the least; 5 = to the highest extent]

In your opinion, to what extent are the causes of the supporters'/opponents' behavior determined by their emotions?
[1 = not in the least; 5 = to the highest extent]

In your opinion, to what extent does the supporters'/opponents' behavior correspond to the interests of the Latvian society?
[participant instructed to indicate the response on a 5-point scale where
-2 = does not correspond at all;
-1 = does not correspond rather than corresponds;
0 = is neutral;
+1 = corresponds rather than does not correspond;
+2 = fully corresponds]

To what extent do you like or dislike the supporters'/opponents' behavior?
[participant instructed to indicate the response on a 5-point scale where
-2 = dislike very much;
-1 = dislike rather than like;
0 = neither like nor dislike;
+1 = like rather than dislike;
+2 = like very much]

Note: These items appeared twice in the questionnaire: once for the supporter behavior and once for the opponent behavior.

Items for rating causal explanations and evaluating target behavior from the group perspectives:

How would people who support/oppose Latvia entering the European Union answer the question to what extent the causes of the supporters'/opponents' behavior originate in the situation rather than within themselves?
[1 = not in the least; 5 = to the highest extent]

How would people who support/oppose Latvia entering the European Union answer the question to what extent the causes of the supporters'/opponents' behavior are stable (invariable)?
[1 = not in the least; 5 = to the highest extent]

How would people who support/oppose Latvia entering the European Union answer the question to what extent the supporters/opponents control their behavior?
[1 = not in the least; 5 = to the highest extent]

How would people who support/oppose Latvia entering the European Union answer the question to what extent the causes of the supporters'/opponents' behavior influence their behavior in all situations (also in those not related to Latvia's EU membership)?
[1 = not in the least; 5 = to the highest extent]

How would people who support/oppose Latvia entering the European Union answer the question to what extent the supporters'/opponents' behavior is intentional?
[1 = not in the least; 5 = to the highest extent]

How would people who support/oppose Latvia entering the European Union answer the question to what extent the supporters/opponents are aware of the causes of their behavior?
[1 = not in the least; 5 = to the highest extent]

How would people who support/oppose Latvia entering the European Union answer the question to what extent the causes of the supporters'/opponents' behavior are determined by their personality traits?
[1 = not in the least; 5 = to the highest extent]

How would people who support/oppose Latvia entering the European Union answer the question to what extent the causes of the supporters'/opponents' behavior are determined by their own interests?
[1 = not in the least; 5 = to the highest extent]

How would people who support/oppose Latvia entering the European Union answer the question to what extent the causes of the supporters'/opponents' behavior are determined by their membership in various social groups?
[1 = not in the least; 5 = to the highest extent]

How would people who support/oppose Latvia entering the European Union answer the question to what extent the causes of the supporters'/opponents' behavior are determined by their care about the country's future?
[1 = not in the least; 5 = to the highest extent]

How would people who support/oppose Latvia entering the European Union answer the question to what extent the causes of the supporters'/opponents' behavior are determined by accepting other people's opinions?
[1 = not in the least; 5 = to the highest extent]

How would people who support/oppose Latvia entering the European Union answer the question to what extent the causes of the supporters'/opponents' behavior are determined by their beliefs?
[1 = not in the least; 5 = to the highest extent]

How would people who support/oppose Latvia entering the European Union answer the question to what extent the causes of the supporters'/opponents' behavior are determined by their desires (aspiration for a specific result)?
[1 = not in the least; 5 = to the highest extent]

How would people who support/oppose Latvia entering the European Union answer the question to what extent the causes of the supporters'/opponents' behavior are determined by their subjective liking or disliking for various aspects of the EU?
[1 = not in the least; 5 = to the highest extent]

How would people who support/oppose Latvia entering the European Union answer the question to what extent the causes of the supporters'/opponents' behavior are determined by their emotions?
[1 = not in the least; 5 = to the highest extent]

How would people who support/oppose Latvia entering the European Union answer the question to what extent the supporters'/opponents' behavior corresponds to the interests of the Latvian society?
[-2 = does not correspond at all;
-1 = does not correspond rather than corresponds;
0 = is neutral;
+1 = corresponds rather than does not correspond;
+2 = fully corresponds]

How would people who support/oppose Latvia entering the European Union answer the question to what extent they like or dislike the supporters'/opponents' behavior?
[-2 = dislike very much;
-1 = dislike rather than like;
0 = neither like nor dislike;
+1 = like rather than dislike;
+2 = like very much]

Note: These items appeared four times in the questionnaire--once for each possible combination of target behavior (supporter or opponent) and group perspective (supporter or opponent). For half of the participants, the wording in all items was "University of Latvia students" instead of "people".

Third Part of the Questionnaire

Instruction for the own perspective:
Please indicate to what extent you agree or disagree with the following statements. Indicate your response by circling the appropriate number next to each statement!

Instruction for the group perspectives:
How would people who support/oppose Latvia entering the European Union agree or disagree with the following statements.? Indicate your response by circling the appropriate number next to each statement!

Items for measuring the strength of support to Latvia's EU membership:

Latvia should enter the European Union.
[participant instructed to indicate the response on a 5-point scale where
1 = strongly disagree;
2 = disagree rather than agree;
3 = neither agree nor disagree;
4 = agree rather than disagree;
5 = strongly agree]

The European Union membership will do more harm than good to Latvia. *
[1 = strongly disagree;
2 = disagree rather than agree;
3 = neither agree nor disagree;
4 = agree rather than disagree;
5 = strongly agree]

Latvia will only gain by remaining outside the European Union. *
[1 = strongly disagree;
2 = disagree rather than agree;
3 = neither agree nor disagree;
4 = agree rather than disagree;
5 = strongly agree]

I am personally against Latvia's membership into the European Union. *
[1 = strongly disagree;
2 = disagree rather than agree;
3 = neither agree nor disagree;
4 = agree rather than disagree;
5 = strongly agree]

It would be better for majority of Latvians if Latvia became a member of the European Union.
[1 = strongly disagree;
2 = disagree rather than agree;
3 = neither agree nor disagree;
4 = agree rather than disagree;
5 = strongly agree]

European Union membership threatens the sovereignty of Latvia. *
[1 = strongly disagree;
2 = disagree rather than agree;
3 = neither agree nor disagree;
4 = agree rather than disagree;
5 = strongly agree]

I personally support Latvia's membership into the European Union.
[1 = strongly disagree;
2 = disagree rather than agree;
3 = neither agree nor disagree;
4 = agree rather than disagree;
5 = strongly agree]

Note: These items appeared three times in the questionnaire, and correspondingly were rated from three perspectives: own, supporter, and opponent. Asterisked items are reverse-scored. Reliability measures (Cronbach's Alpha): for ratings from own perspective, Alpha = .90, for ratings from the supporter perspective, Alpha = .78, for ratings from the opponent perspective, Alpha = .82.

APPENDIX B. CORRELATION MATRIXES WITH MEANS AND STANDARD DEVIATIONS FOR ALL VARIABLES

Table B1

Attitude towards EU membership

Supporter participants

	n	M	SD	(1)	(2)
(1) Own perspective	90	3.88	0.55		
(2) Supporter perspective	85	4.30	0.61	.39**	
(3) Opponent perspective	87	1.91	0.85	.03	-.48**

Opponent participants

	n	M	SD	(1)	(2)
(1) Own perspective	46	2.32	0.56		
(2) Supporter perspective	47	4.29	0.65	-.34*	
(3) Opponent perspective	49	1.94	0.74	.33*	-.71**

* $p < 0.05$, ** $p < 0.01$.

Table B2

Evaluation of the target behavior

Supporter participants

	n	M	SD	(1)	(2)	(3)	(4)	(5)
(1) SB, own perspective	92	3.53	0.67					
(2) SB, supporter persp.	91	4.13	0.73	.14				
(3) SB, opponent persp.	91	1.96	0.82	.08	-.33**			
(4) OB, own perspective	92	2.63	0.81	-.07	-.13	.19		
(5) OB, supporter persp.	91	1.99	0.86	.03	-.32	.43**	.30**	
(6) OB, opponent persp.	91	3.85	1.06	-.27*	.30**	-.41**	.06	.35**

Opponent participants

	n	M	SD	(1)	(2)	(3)	(4)	(5)
(1) SB, own perspective	48	2.77	0.76					
(2) SB, supporter persp.	49	3.96	0.87	-.06				
(3) SB, opponent persp.	49	2.36	0.91	.23	-.22			
(4) OB, own perspective	48	3.52	0.82	-.01	.09	.23		
(5) OB, supporter persp.	49	2.13	0.81	.18	-.43**	.30*	.17	
(6) OB, opponent persp.	48	3.80	0.97	.02	.68**	.01	.37**	-.28

SB = supporter behavior; OB = opponent behavior

* $p < 0.05$, ** $p < 0.01$.

Table B3

Situation vs. actor ratings

Supporter participants

	n	M	SD	(1)	(2)	(3)	(4)	(5)
(1) SB, own perspective	92	3.42	0.97					
(2) SB, supporter persp.	92	3.48	0.95	.29**				
(3) SB, opponent persp.	91	3.11	1.09	.31**	.00			
(4) OB, own perspective	92	3.46	1.02	.06	.08	.33**		
(5) OB, supporter persp.	92	3.14	1.13	.08	-.04	.42**	.23*	
(6) OB, opponent persp.	91	3.57	1.05	-.04	.12	-.06	-.03	.07

Opponent participants

	n	M	SD	(1)	(2)	(3)	(4)	(5)
(1) SB, own perspective	48	3.46	0.92					
(2) SB, supporter persp.	49	3.63	1.11	.34*				
(3) SB, opponent persp.	49	3.55	0.89	.46**	.29*			
(4) OB, own perspective	49	3.39	1.00	.28	.47**	.60**		
(5) OB, supporter persp.	49	3.43	1.12	.05	-.09	-.10	-.10	
(6) OB, opponent persp.	49	3.47	0.98	.36*	.54**	.41**	.51**	-.07

SB = supporter behavior; OB = opponent behavior

$* p < 0.05$, $** p < 0.01$.

Table B4

Stability ratings

Supporter participants

	n	M	SD	(1)	(2)	(3)	(4)	(5)
(1) SB, own perspective	92	2.96	0.91					
(2) SB, supporter persp.	91	3.76	0.97	.29**				
(3) SB, opponent persp.	92	2.49	1.04	.30**	.12			
(4) OB, own perspective	91	2.87	0.90	.01	-.04	.19		
(5) OB, supporter persp.	91	2.59	1.00	.19	-.10	.48**	.32**	
(6) OB, opponent persp.	91	3.59	1.16	-.03	.09	.16	.19	.06

Opponent participants

	n	M	SD	(1)	(2)	(3)	(4)	(5)
(1) SB, own perspective	48	2.94	0.84					
(2) SB, supporter persp.	49	3.84	0.99	.15				
(3) SB, opponent persp.	49	3.02	0.88	.26	.05			
(4) OB, own perspective	47	3.17	0.96	-.09	.37*	.32*		
(5) OB, supporter persp.	49	2.86	1.04	.14	-.19	.09	-.08	
(6) OB, opponent persp.	49	3.61	0.86	.24	.34*	.12	.42**	.10

SB = supporter behavior; OB = opponent behavior * $p < 0.05$, ** $p < 0.01$.

Table B5

Control ratings

Supporter participants

	n	M	SD	(1)	(2)	(3)	(4)	(5)
(1) SB, own perspective	92	3.12	0.94					
(2) SB, supporter persp.	92	3.82	1.00	.33**				
(3) SB, opponent persp.	91	2.53	1.08	.09	.05			
(4) OB, own perspective	92	2.82	1.06	.19	.21*	-.08		
(5) OB, supporter persp.	92	2.46	0.95	-.16	-.23*	.27**	.26*	
(6) OB, opponent persp.	92	3.64	1.09	.11	.34**	-.03	.12	-.11

Opponent participants

	n	M	SD	(1)	(2)	(3)	(4)	(5)
(1) SB, own perspective	48	3.13	1.00					
(2) SB, supporter persp.	49	3.86	1.02	.03				
(3) SB, opponent persp.	49	2.94	0.94	.32*	-.10			
(4) OB, own perspective	49	3.29	1.06	.17	.21	.27		
(5) OB, supporter persp.	49	2.84	1.12	.02	-.16	.23	.08	
(6) OB, opponent persp.	49	3.59	1.22	.23	.40**	.12	.41**	-.16

SB = supporter behavior; OB = opponent behavior

$* p < 0.05$, $** p < 0.01$.

Table B6

Globality ratings

Supporter participants

	n	M	SD	(1)	(2)	(3)	(4)	(5)
(1) SB, own perspective	92	3.28	0.82					
(2) SB, supporter persp.	92	3.33	0.88	.25*				
(3) SB, opponent persp.	92	3.13	0.92	.01	.10			
(4) OB, own perspective	91	3.10	0.96	.18	.05	.09		
(5) OB, supporter persp.	92	3.41	0.92	.03	-.02	.11	.42**	
(6) OB, opponent persp.	92	3.16	0.76	.21*	.22*	.19	.07	.08

Opponent participants

	n	M	SD	(1)	(2)	(3)	(4)	(5)
(1) SB, own perspective	48	3.25	0.91					
(2) SB, supporter persp.	49	3.06	1.05	.23				
(3) SB, opponent persp.	49	3.35	0.97	.31*	.14			
(4) OB, own perspective	49	3.14	0.89	.16	.24	.04		
(5) OB, supporter persp.	49	3.31	0.94	.25	-.04	.20	.12	
(6) OB, opponent persp.	48	3.19	0.94	.11	.44**	.00	.35*	-.14

SB = supporter behavior; OB = opponent behavior

* $p < 0.05$, ** $p < 0.01$.

Table B7

Rationality index

Supporter participants

	n	M	SD	(1)	(2)	(3)	(4)	(5)
(1) SB, own perspective	92	3.29	0.78					
(2) SB, supporter persp.	91	3.99	0.83	.34**				
(3) SB, opponent persp.	91	2.67	0.91	.06	-.14			
(4) OB, own perspective	92	2.94	0.78	.22*	.23*	.08		
(5) OB, supporter persp.	91	2.68	0.93	.04	-.25*	.36**	.30**	
(6) OB, opponent persp.	90	3.79	0.94	-.06	.46**	-.05	.09	-.04

Opponent participants

	n	M	SD	(1)	(2)	(3)	(4)	(5)
(1) SB, own perspective	46	3.13	0.70					
(2) SB, supporter persp.	49	3.86	0.85	.37*				
(3) SB, opponent persp.	47	2.96	0.82	.21	-.08			
(4) OB, own perspective	49	3.45	0.83	.41**	.34*	.30*		
(5) OB, supporter persp.	49	2.83	0.83	.24	-.06	.08	-.07	
(6) OB, opponent persp.	48	3.84	0.84	.40**	.47*	.16	.60**	-.04

SB = supporter behavior; OB = opponent behavior

$* p < 0.05$, $** p < 0.01$.

Table B8

Subjectivity index

Supporter participants

	n	M	SD	(1)	(2)	(3)	(4)	(5)
(1) SB, own perspective	91	3.68	0.52					
(2) SB, supporter persp.	90	3.46	0.54	.39**				
(3) SB, opponent persp.	89	3.54	0.62	.43**	.40**			
(4) OB, own perspective	92	3.80	0.53	.48**	.21*	.44**		
(5) OB, supporter persp.	89	3.68	0.61	.50**	.26*	.48**	.57**	
(6) OB, opponent persp.	89	3.36	0.64	.35**	.41**	.22*	.40**	.06

Opponent participants

	n	M	SD	(1)	(2)	(3)	(4)	(5)
(1) SB, own perspective	45	3.81	0.51					
(2) SB, supporter persp.	47	3.38	0.59	.31*				
(3) SB, opponent persp.	49	3.67	0.68	.48**	.41**			
(4) OB, own perspective	49	3.79	0.53	.56**	.32*	.56**		
(5) OB, supporter persp.	49	3.80	0.57	.48**	.17	.60**	.78**	
(6) OB, opponent persp.	48	3.54	0.59	.44**	.36*	.41**	.60**	.44**

SB = supporter behavior; OB = opponent behavior

* $p < 0.05$, ** $p < 0.01$.

Table B9

Accepting others' opinions ratings

Supporter participants

	n	M	SD	(1)	(2)	(3)	(4)	(5)
(1) SB, own perspective	91	3.13	1.04					
(2) SB, supporter persp.	92	2.75	1.02	.14				
(3) SB, opponent persp.	92	3.33	1.21	.25*	.04			
(4) OB, own perspective	91	3.20	1.11	.18	-.01	.06		
(5) OB, supporter persp.	92	3.52	1.05	.25*	.02	.11	.38**	
(6) OB, opponent persp.	92	2.64	1.13	-.02	.45**	-.16	.11	.13

Opponent participants

	n	M	SD	(1)	(2)	(3)	(4)	(5)
(1) SB, own perspective	47	3.55	0.88					
(2) SB, supporter persp.	49	2.98	1.15	.14				
(3) SB, opponent persp.	49	3.22	1.12	.42**	.02			
(4) OB, own perspective	47	3.02	0.92	.20	-.19	.20		
(5) OB, supporter persp.	49	3.29	1.14	.27	-.32*	.23	.27	
(6) OB, opponent persp.	49	3.06	1.03	.09	.30*	-.12	.25	.16

SB = supporter behavior; OB = opponent behavior

* $p < 0.05$. ** $p < 0.01$.

AUTHORS' BIOGRAPHIES

Girts Dimdins is currently a doctoral student at the Department of Psychology, Stockholm University. His research is focusing on the role of perspective taking in intergroup perception. E-mail: gds@psychology.su.se

Henry Montgomery is a professor at the Department of Psychology, Stockholm University. His research deals with mental processes and structures associated with people's values, attitudes and decisions. A particular interest is the role of perspectives in people's interpretations of actions and action alternatives. E-mail: hmy@psychology.su.se

Ivars Austers is an associate professor at the Department of Psychology, Faculty of Education and Psychology, University of Latvia. His research interests are social cognition and political psychology. E-mail: austers@lanet.lv

STUDY 3

False polarisation when imagining others' explanations of attitudinal behaviour

Girts Dimdins and Henry Montgomery
Stockholm University, Stockholm, Sweden
Ivars Austers
University of Latvia, Riga, Latvia

Abstract

Supporters and opponents of Latvia's EU membership explained attitudinal behaviour of both groups from their own perspective and from an imagined in-group or out-group perspective. Content analysis of data show that explanations given from the imagined group perspectives differed more between opposing attitudinal behaviours (particularly in terms of reason versus causal history of reason explanations) than those given from participants' own perspective. The results indicate that the false polarisation effect influences not only quantitative ratings, but also the qualitative contents of free-response explanations. This study demonstrates the applicability of Malle's (1999) Folk Explanations of Behaviour coding scheme to the analysis of intergroup attributions.

Key words: Intergroup attribution, group attribution, false polarisation effect, perspective taking

Theoretical Background

Hardin and Higgins (1996) note that even basic cognitive processes are defined by the social activities in which they are manifested. People construct the reality they live in through social verification, and nowhere else is this phenomenon more vivid than in the context of intergroup relations. When reality is constructed through the process of social verification, it is subject to various interpretations by the groups involved in the process. The same "objective" and physical facts can be interpreted very differently depending on the perspective from which they are seen (Montgomery, 1994). When the interests of social groups clash, these interests determine their perspectives. If one or all of these groups fail to acknowledge that the other side sees the world differently than they do, this may lead to erroneous attributions, misunderstandings, and conflicts between the parties (Griffin & Ross, 1991).

Ichheiser (1949) notes that people are aware of differences between opinions of different social groups, explaining these differences by attributing psychological biases to others. Pronin, Puccio, and Ross (2002) conclude that the real source of misunderstanding arises from people's failure to recognise the operation of the same biases in their own judgements and decisions. This represents the so-called "naïve realism" view of the social perceiver (Ross & Ward, 1996). According to this view, social perceivers assume they have access to unbiased and objective reality, which others should perceive the same way as long as their view is not biased by ideology or self-interest. This assumption often results in an overestimation of psychological biases in other people (Kruger & Gilovich, 1999; Pronin, Gilovich, & Ross, in press; Pronin, Lin, & Ross, 2002).

One of the consequences of the "naïve realism" of the social perceiver is the so-called false polarisation effect (Pronin et al., 2002)--an overestimation of a group score on a certain attribute in comparison with the actual group score on that attribute. Robinson, Keltner, Ward, and Ross (1995) found that both liberals and conservatives tend to overestimate the differences in both groups' interpretations of a controversial incident. Respondents of both groups tend to see both in-group and out-group members as more extreme in their attitudes than themselves. Similar findings in a study of book preferences among traditionalist and revisionist English professors by Keltner and Robinson (1997) lead them to conclude that representatives of opposing parties tend to rely on a polarisation heuristic when judging their conflicts, accentuating the differences between both sides' positions and underestimating the common ground. In some studies, the false polarisation effect has been found in estimating out-group attitudes (Robinson & Friedman, 1995; Rouhana, O'Dwyer, & Morrison Vaso, 1997). Other studies report similar overestimation of both out-group and in-group attitudes (Diekman, Eagly, & Kulesa, 2002; Keltner & Robinson, 1997; Puccio, 2003; Robinson et al., 1995).

2

Studies of intergroup attribution constitute one area of social psychological research that highlights differences between groups in construing the same observable events. Hewstone (1989, 1990) reports a number of findings showing that individuals prefer in-group-serving attributions relative to out-group-serving attributions. Perspective-taking studies of intergroup attribution have found evidence of the false polarisation effect in explanations of group attitudes and behaviour. Robinson et al. (1995) tested some aspects of group attribution by asking their participants to rate the basis of their own judgements and that of their in-group and out-group members. The respondents indicated that they had been less influenced by ideology or political orientation than either their peers or opponents. Dimdins, Montgomery, and Austers (2003) asked European Union supporters and opponents in Latvia to rate various causal explanations of supporters' and opponents' attitudinal behaviour. From a target's in-group perspective, they found that both groups rated the causes of behaviour as more stable, controllable, and rational, and less subjective and less influenced by others than the respective group did from their own perspective. Austers (2002) asked ethnic Latvian and Russian schoolteachers to rate positive and negative behaviours from their own and out-group perspective on a number of attributional dimensions. The results indicated some false polarisation effect in predicting out-group's responses, although the results were not consistent. The cited studies measured attribution on various quantitative dimensions. This paper explores whether the false polarisation effect would appear in free response explanations of attitudinal behaviour.

Traditionally, intergroup attribution studies have been based on two classifications of causal factors. The older classification, based on Heider's (1958) model of interpersonal attributions, distinguishes between internal (dispositional) and external (situational) attributions. Because of several limitations of this classification (Miller, Smith, & Uleman, 1981), an alternative four-dimensional model was proposed by Islam and Hewstone (1993), measuring attributions along the continuums of causal locus, stability, controllability (Weiner, 1986), and globality (Abramson, Seligman, & Teasdala, 1978). A number of successful intergroup attribution studies have been carried out using this model with the four dimensions serving as the basis for quantitative rating scales (Austers, 2002; Islam & Hewstone, 1993), coding categories of free responses (Austers & Montgomery, 2001, Lee & Robinson, 2000), and stimulus material in different experimental conditions (Wilder, Simon & Faith, 1996).

Bertram Malle and his colleagues (Malle, 1999; Malle & Knobe, 1997; Malle, Knobe, O'Laughlin, Pearce, & Nelson, 2000; O'Laughlin & Malle, 2002) have recently proposed a different alternative to the person/situation classification. The main idea behind their model -- Folk Explanations of Behaviour (FEB) -- is that rather than concentrating on the explicit contents of causal attributions people's behaviour explanations must instead be considered

in terms of their conceptual structure, the network of concepts and assumptions on which these explanations are based.

Central to this structure is the folk concept of intentionality (Malle & Knobe, 1997). According to the classification suggested by Malle (1999), explanations that interpret an agent's behaviour as unintentional are considered cause explanations. Explanations that interpret the behaviour as intentional are divided into three major groups: (a) causal history of reasons (factors from agent's personal history that cause the intentional behaviour without the agent being aware of them); (b) reasons (factors that the agent considered when forming the intention to act); and (c) enabling factors (factors that clarify how it was possible that the agent completed the intended action). Reasons are further divided into three types: desires, beliefs, and valuings. A distinction is made between reasons marked with a mental state description and those unmarked. In addition, the coding scheme developed by Malle (2000) uses a number of secondary codes to reflect the contents of each explanation. Some examples of coding categories can be found in Appendix A.

Perhaps the most appealing aspect of using the folk explanations of behaviour model in coding of attributions is that it uses categories that take into account both linguistic and conceptual features of the explanations. Such categories are not as sensitive to subtle linguistic differences as those traditionally used in person/situation classification. At the same time, the coding scheme provides strict rules based on a consistent inner logic to avoid speculative interpretations during analysis.

Of special interest to the studies of intergroup attributions is the distinction between causal history of reasons (CHR) and reasons. Both categories explain intentional action; however, using CHR (such as an actor's personality traits, social category membership, and the state of being deprived of relevant information) shows, in the explainer's opinion, that the actor is not aware of what factors have caused his or her behaviour. Locke and Pennington (1982, p. 213) make a philosophical distinction between explaining and justifying behaviour. The same reasoning can be applied to the distinction between reason and CHR explanations. CHR explanations can explain the action but cannot justify it. Reasons, on the contrary, both explain the action and justify it. (In the observer's opinion, actors themselves could explain why they acted in a certain way if asked.) According to Locke and Pennington, "Behaviour that is so justified can be said to be rational in Simon's (1957) sense of 'bounded rationality'.... If behaviour cannot be justified this way, then we have to look for other explanations, for example, psychopathologies" (p. 213). In other words, awareness of one's causes of behaviour implies more rationality than lack of such awareness.

Previous studies have shown that people perceive rationality as a positive factor in intergroup attributions (Dimdins et al., 2003; Kenworthy & Miller, 2002). Consequently, one may expect to find systematic differences in people's

4

use of reason and CHR explanations when explaining positive and negative behaviour or in-group and out-group behaviour. It is reasonable to expect that if the observer finds an actor's behaviour personally acceptable, positive, and congruent with one's own attitudes, he or she will tend to use more reason explanations relative to CHR explanations; the opposite should be true for negatively evaluated behaviour.

Our study has two aims. First, we examine if the false polarisation effect would appear in the free response explanations of attitudinal behaviour. (By attitudinal behaviour, we understand general, non-specific behaviour that expresses the actor's attitude towards a certain issue, acting in line with one's attitudes. We chose attitudinal behaviour instead of more specific group behaviours because it allowed us to construct simple, context-independent, and unambiguous stimuli for our study.) Second, we explore the possibilities of using FEB scheme in studies of intergroup attributions.

We made two predictions regarding the false polarisation effect. First, following our reasoning described above, we expected that the ratio between the number of reasons and the CHR explanations could elicit false polarisation. More reason explanations relative to CHR explanations can be expected when members of a target's in-group explain its behaviour than if members of the target's out-group explain the same behaviour. (We understand in-group as the participants who share the attitude expressed in the target behaviour, and we understand out-group as the participants who hold the opposite attitude.) Thus, if the participants' responses are influenced by the false polarisation, participants should provide more reason (relative to CHR) explanations when taking the (imagined) in-group perspective of the target than the target's in-group members have provided from their own perspective. In addition, simulated explanations given from a target's out-group perspective should contain more CHR (relative to reason) explanations compared to the explanations given by that target's out-group members.

Our second prediction concerns the linguistic aspects of the FEB coding scheme, the interpretation of mental state markers in reason explanations. Malle (1999, p. 41) notes that mental state markers are likely to be used by observers to distance themselves from the agent's subjective reasons. By extending this assertion to an intergroup context, one should expect more marked (relative to unmarked) reason explanations when a target's out-group members explain the target's behaviour than if the same behaviour is explained by the target's in-group members. In this way, explainers would try to distance themselves from reasons they might find personally less acceptable. We wanted to test if this pattern would indeed appear in an intergroup context. If so, we wanted to test if the use of mental state markers would be affected by false polarisation. We predicted that participants should use more mental state markers when taking the (imagined) out-group perspective of the target than the target's out-group

have used from their own perspective; the opposite should be true regarding the target's in-group perspective.

After a pilot study of political attitudes among Latvian students, we decided that Latvia's planned membership in the European Union was the most appropriate issue for our study because it elicited the strongest and most polarised attitudes. Defining the group membership in terms of attitudes is a valid practice in studies of controversial political issues (Kenworthy & Miller, 2002).

Method

Sample

Eighty-four students (75 women) at the University of Latvia participated in the study. Participants were first, second, and third year students of psychology, education, and languages. The mean age of the participants was 20 years.

Questionnaire

The participants were first asked to indicate their age, gender, ethnicity, and their general attitude towards Latvia's membership into the European Union (for vs. against).

The second part of the questionnaire consisted of descriptions of two opposite attitudinal behaviours, which the participants were asked to explain from their own perspective and from the perspective of either EU supporters or EU opponents. The opposite behaviours were described with the following statements: (a) *"There are many people in Latvia who actively support Latvia entering the European Union"*; and (b) *"There are many people in Latvia who activelyprotest against Latvia entering the European Union"*. Each statement was on top of a separate sheet. All respondents explained both behaviours. The order of both target behaviours was counterbalanced across the questionnaires.

After reading each of the two statements, the respondents were asked to explain the respective behaviour from two perspectives. The first question was the same in all questionnaires: *"Why do you believe people who support Latvia entering the European Union/people who oppose Latvia entering the European Union act the way they do? Please, write the answer in the space below."* The second question varied depending on which group perspective was assigned to the respective questionnaire. Half of the respondents were asked to explain both behaviours from the perspective of EU supporters, whereas the other half were asked to explain both behaviours from the perspective of EU opponents. For example, for supporter behaviour to be explained from the EU opponent perspective, the question reads: *"How would people who oppose Latvia entering the European Union answer this question: "Why do supporters of European Union act this way?"* The wording was developed through several rounds of pre-testing to ensure that the participants understood the perspective-taking questions correctly.

In the third part of the questionnaire, the participants were asked to indicate their agreement/disagreement with seven items using a 5-point Likert scale. This was designed to measure the strength of support of Latvia's membership into the European Union. The list of items can be found in Appendix B. The participants were also asked to indicate to what extent in their opinion people who support (or, depending on the group perspective, oppose) Latvia entering the European Union would agree or disagree with the 7 items. The group perspective indicated in this question was consistent with that indicated in the previous questions.

To reiterate, the group perspective was manipulated between subjects. Each respondent was asked to explain both attitudinal behaviours (supporter and opponent in random order) from their perspective and from the perspective of one of the groups -- either EU supporters or EU opponents. This design was chosen to minimise the possibility that participants might guess the research hypotheses.

Content Analysis of Data

Two independent judges first divided all responses into codeable units. A codeable unit was defined as a unit of text that alone would be sufficient to explain the behaviour in question. The initial agreement between the judges on the division into units was 87.7%. After discussing the differences, both judges reached a 100% agreement on the units that had to be coded.

Afterwards, the same judges independently coded all the units in accordance with the F.EX Coding Scheme for Folk Explanations of Behaviour (Malle, 2000). The explanation type (cause explanations, causal history of reason [CHR] explanations, reason explanations [marked and unmarked], and enabling factor explanations) served as the primary code. The contents of each explanation were coded using a number of secondary codes (more than 20 different codes for each explanation type). Examples of the codes used are given in Appendix A.

The initial agreement between judges for the primary codes was 88.3%. After discussing the differences, the judges reached an agreement on 99.4% of the units. The units that the coders could not agree upon were dropped from the analysis. For secondary codes (i.e., how the contents of the explanations were classified), the agreement was relatively low (65.7%), and these categories were not used in the data analysis.

None of the explanations were coded as enabling factor explanations, and only two of the 796 units were coded as cause explanations. These two types of explanations will not appear in the results and discussion sections below.

Results

On average, the participants provided 2.35 explanations per behaviour ($SD = 1.29$). There was a tendency that slightly more explanations were provided from own perspective (for opponent behaviour $M = 2.79$, $SD = 1.41$;

for supporter behaviour $M = 2.54$, $SD = 1.35$) than from group perspective (for opponent behaviour $M = 2.11$, $SD = 1.18$; for supporter behaviour $M = 2.05$, $SD = 1.21$). Sixty-one participants (73%) identified themselves as supporters of Latvia's EU membership, and twenty-three (27%) identified themselves as opponents.

An index measuring the support of Latvia's membership into the European Union was calculated for each participant by summing the responses to the seven items displayed in Appendix B. An index for predicted strength of support from the imagined group perspective was calculated in the same way.

As we predicted, the numbers of reason and CHR explanations were negatively correlated for each target behaviour (for supporter behaviour, own perspective $r(84) = -.33$, $p < .01$, group perspective $r(84) = -.47$, $p < .001$; for opponent behaviour, own perspective $r(84) = -.46$, $p < .001$, group perspective $r(84) = -.49$, $p < .001$). This allowed us to calculate a Reason-CHR index for each participant by subtracting the number of CHR explanations from the number of Reason explanations and dividing the difference by the total number of explanations given. An index value of -1.00 means that only CHR explanations were given, and +1.00 indicates that only Reason explanations were given. An index value of zero results from an equal number of both types of explanations.

Similarly, the numbers of marked and unmarked explanations were negatively correlated for each target behaviour (for supporter behaviour, own perspective $r(78) = -.23$, $p < .05$, group perspective $r(65) = -.54$, $p < .001$; for opponent behaviour, own perspective $r(76) = -.32$, $p < .01$, group perspective $r(62) = -.59$, $p < .001$). We calculated an Unmarked-Marked index for each participant by subtracting the number of marked reason explanations from the number of unmarked reason explanations and dividing the difference by the total number of reason explanations given. A more positive index thus indicates a prevalence of unmarked reason explanations over marked explanations.

We based our data analysis on planned pairwise comparisons (one-tailed t-tests) between the rating means from own perspective and the corresponding means from both group perspectives. Such a direct test of the false polarisation effect increased the power of the analysis in comparison with a full ANOVA. For each comparison, we computed effect size (Cohen's d) to provide an estimate of the magnitude of mean differences.

The top row of Table 1 shows the strength of support index means. The means for supporter and opponent participants from their own perspective differed significantly, $t(79) = 11.42$, $p < .05$. When answering the same questions from the imagined supporter perspective, both supporter and opponent participants showed a strong false polarisation effect. When taking the opponents' perspective, only the opponent participants showed moderate false polarisation. Supporters, however, accurately predicted the opponents'

responses. The means for either of the imagined group perspectives did not differ significantly between supporter and opponent participants.

Table 1

Mean ratings from own perspective and the imagined group perspectives

	Ratings from own perspective		Ratings from the supporter perspective		Ratings from the opponent perspective	
	By supporters (n = 61)	By opponents (n = 23)	By supporters (n = 33)	By opponents (n = 10)	By supporters (n = 28)	By opponents (n = 13)
Support to EU membership	26.19	15.14	30.73***	29.70**	14.22	12.25*
	(4.00)	(3.51)	(3.76)	(2.71)	(4.72)	(5.49)
Effect size			1.17	1.03	0.22	0.63
Reason/CHR index:						
Supporter behaviour	0.84	0.68	1.00*	1.00	-0.06***	0.36
	(0.46)	(0.63)	(0.00)	(0.00)	(0.89)	(0.92)
Effect size			0.49	0.49	0.96	0.40
Opponent behaviour	0.46	0.80	0.08*	-0.33**	0.92	0.85
	(0.71)	(0.39)	(0.84)	(0.94)	(0.39)	(0.55)
Effect size			0.49	0.95	0.31	0.10
Unmarked/marked index						
Supporter behaviour	0.06	-0.02	0.60***	0.43	-0.65**	-0.75*
	(0.71)	(0.81)	(0.72)	(0.83)	(0.60)	(0.46)
Effect size			0.76	0.48	0.88	1.11
Opponent behaviour	-0.40	0.08	-0.17	-0.17	0.45	0.17
	(0.69)	(0.79)	(0.88)	(1.00)	(0.80)	(1.03)
Effect size			-0.29	-0.27	0.47	0.10

Note. Standard deviations are given in parentheses. Asterisks indicate that rating from the imagined group perspective is significantly different (1-tailed *t*-test) from the corresponding rating from own perspective. Means in columns 3 and 4 are compared to means in column 1; means in columns 5 and 6 are compared to means in column 2. The effect size (Cohen's *d*) for each comparison is given below the respective mean in columns 3-6. * $p < .05$. ** $p < .01$. *** $p < .001$.

Regarding the CHR-Reason index, for the supporter behaviour both groups of participants did not differ significantly -- $t(80) = -1.26$, $p = .21$ -- when giving explanations from own perspective. The predicted pattern of means emerges in explanations from the imagined group perspectives. From the

supporter perspective, both groups used only reason explanations (note that the same mean difference is significant for supporter participants, but it is not significant for the opponent participants because of a different number of degrees of freedom). A stronger false polarisation effect occurred when explanations were provided from the opponent perspective, especially for the supporter participants. The mean pattern for the opponent behaviour is similar. Here, the means for both groups from own perspective differed significantly, $t(80) = 2.17$, $p < .05$. The false polarisation effect can be seen in explanations provided by both groups from the imagined supporter perspective but not from the opponent perspective. For both behaviours, the means for either of the group perspectives did not differ significantly between supporter and opponent participants. To summarise, a target's in-group members showed the strongest false polarisation (in terms of using more CHR explanations relative to reason explanations) when explaining the target behaviour from its out-group perspective.

Regarding the Unmarked-Marked index, the means in Table 1 show that for own perspective, the expected pattern of means appears in explanations for the opponent behaviour; supporter participants used more mental state markers than the opponent participants, $t(73) = 2.64$, $p < .05$. When explaining the supporter behaviour, both groups used approximately equal number of marked and unmarked explanations, $t(76) = -0.43$, $p = 0.67$. The expected pattern of means appears in explanations provided from the imagined group perspectives. When explaining supporter behaviour from the supporter perspective, both groups used fewer mental state markers than supporters did from their own perspective; from the opponent perspective, both groups used more mental state markers than opponents themselves used. For opponent behaviour, the respective mean differences are small and insignificant. For supporter perspective, the mean differences were not in the expected direction. Again, for both behaviours, the means for either of the group perspectives did not differ significantly between supporter and opponent participants. We conclude that there is some evidence of false polarisation regarding the use of mental state markers when explaining attitudinal behaviour from the imagined group perspectives, but this evidence is not conclusive.

Discussion

Evidence confirms both of our predictions. Prediction 1 (regarding the use of CHR and reason explanations) received stronger support from the results. Although not all mean differences were significant, they were always in the expected direction for both supporter and opponent behaviour. Support for Prediction 2 (regarding the use of mental state markers) was less consistent, because we found the expected mean pattern only for explanations to supporter behaviour but not opponent behaviour. However, generally our results show that the false polarisation effect is a robust phenomenon that influences the

qualitative contents of free-response explanations. In this study, we replicated the results of Dimdins et al. (2003), using a different method of data gathering and analysis.

An interesting pattern emerging from the results was that both groups of judges similarly overestimated the predicted responses from the group perspectives. Supporters and opponents agreed on how both groups should see the attitudes and explain the behaviour of their in-group and out-group members. The results reflect a sort of "shared pseudo-reality" as opposed to the "actual reality" in which both groups were significantly less extreme in their ratings and explanations. As Robinson et al. (1995) note, one possible explanation for this phenomenon deals with the salience of social debate and availability of information about both groups' positions. Most information about the European Union membership in Latvia comes from mass media, where most of the debate on the topic takes place. People expressing their opinions in such debate are typically politicians or socially active members of non-governmental organisations. Thus their opinions are more extreme than that of an "average person", and regular exposure to such extreme opinions may lead to development of stereotypes about both groups. In other words, people are relying on the availability heuristic (Taylor, 1982; Tversky & Kahneman, 1973) when trying to predict a person's responses. As the analysis of free responses indicates, this explanation is supported by the fact that both groups tended to use similar rhetoric when explaining both target behaviours from the same perspective.

Puccio (2003) goes a step further and argues that because of the correspondence bias (Gilbert & Malone, 1995; Jones & Davis, 1965) people tend to assume that the salient behaviour of the biased sample represents an accurate reflection of the views of the particular group. This assumption may lead to pluralistic ignorance (Miller, Monin, & Prentice, 2000) -- a widespread misperception of group norm that leads to a situation where a majority of group members privately disagree with the norm, but erroneously assume that majority of other group members support it. It is possible that the "shared pseudo-reality" that we found in our data reflects a sort of pluralistic ignorance about the mutual perceptions of EU supporters and EU opponents in Latvian society.

Findings by Van Boven, Kamada, and Gilovich (1999) point at another possible link between the correspondence bias and the false polarisation effect. They found that actors overestimated the magnitude of correspondence bias displayed by observers. In several experiments with highly salient situational constraints on actors' behaviour, the observers indeed displayed the correspondence bias. However, the actors predicted that the observers would make more favourable (or unfavourable, depending on the experimental condition) dispositional inferences than the observers actually did. One may speculate that the same "biased perception of bias" could explain the

exaggerated attributions from group perspectives that we found in our study. We believe that the cognitive mechanisms behind the false polarisation effect, and in particular the relation between the correspondence bias and the false polarisation effect, still produce possibilities for further research, especially in the area of intergroup attribution.

Our study proved that the FEB scheme is an efficient and convenient tool for studies of group attributions. In line with our expectations, we found that the explainers tended to use more reason explanations for attitude-congruent (in-group) behaviour, and more causal history of reason explanations for attitude-incongruent (out-group) behaviour. We think that the ratio between both types of explanations is worth adding to the set of methods for estimating group-serving biases in intergroup attribution studies.

From a methodological respect, the application of the FEB scheme in our study was also successful. The initial agreement between the coders on the primary codes (i.e., the four basic explanation types) was high despite the fact that both judges used the coding scheme for the first time. This suggests that these coding categories are stable, valid, and relatively easy to recognise. The categories proved applicable to an intergroup context, and the results yielded significant effects. The application of secondary codes (mostly those related to the explanation context) was more problematic. Because of the relatively low initial agreement between judges, none of the data based on secondary codes were included in the analysis. The low agreement perhaps can be partially explained by the lack of experience of the coders; however, it may also indicate that the secondary coding categories as they are might be less applicable in an intergroup context than they are in an interpersonal context where the scheme has been developed. The large number of secondary categories also poses a problem of how to reduce the number of variables in the analysis. Perhaps after accumulating a larger amount of data, a factor analysis might yield some insights in this respect. We are convinced, however, that the FEB scheme can provide unexplored possibilities for intergroup attribution research.

In our study, a novel aspect of applying the FEB scheme was that it was used in respect to explanations provided in Latvian. To our knowledge, this was the first attempt to use this coding scheme in a language other than English. Latvian belongs to the Baltic branch of the Indo-European language tree, and it is very different from English in both vocabulary and grammar. The fact that the same coding categories could be successfully applied in a completely different linguistic context offers strong support to the validity of these categories and points to possibilities of using the FEB scheme in cross-cultural research in the future. In general, our findings suggest that the most interesting results in the future studies of intergroup attribution can be expected from the use of free-response formats and qualitative methods of data analysis.

References

Abramson, L. Y., Seligmann, M. E. P., & Teasdale, J. D. (1978). Learned helplessness in humans: Critique and reformulation. *Journal of Abnormal Psychology, 87,* 49-74.

Austers, I. (2002). Ability to take an out-group's perspective in explaining positive and negative behaviors. *Scandinavian Journal of Psychology, 43,* 407-412

Austers, I., & Montgomery, H. (2001). Explaining an out-group's political behavior from four different perspectives. *Baltic Journal of Psychology, 2,* 23-33.

Diekman, A. B., Eagly, A. H., & Kulesa, P. (2002). Accuracy and bias in stereotypes about the social and political attitudes of women and men. *Journal of Experimental Social Psychology, 38,* 268-282.

Dimdins, G., Montgomery, H., & Austers, I. (2003). The false polarization effect in explanations of attitudinal behavior. *Current Research in Social Psychology, 8,* 275-301.

Gilbert, D. T., & Malone, P. S. (1995). The correspondence bias. *Psychological Bulletin, 117,* 21-38.

Griffin, D.W., & Ross, L. (1991). Subjective construal, social inference, and human misunderstanding. In M. P. Zanna (Ed.), *Advances in Experimental Social Psychology* (Vol. 24, pp. 319-359). New York: Academic Press.

Hardin, C. D., & Higgins, E. T. (1996). Shared reality: How social verification makes the subjective objective. In R. M. Sorrentino, & E. T. Higgins (Eds.), *Handbook of motivation and cognition.* (Vol. 3, pp. 28-84). New York: The Guilford.

Hewstone, M. (1989). *Causal attribution: From cognitive processes to collective beliefs.* Oxford: Basil Blackwell.

Hewstone, M. (1990). The "ultimate attribution error"? A review of the literature on intergroup causal attribution. *European Journal of Social Psychology, 20,* 311-335.

Ichheiser, G. (1949). Misunderstandings in human relations: A study in false social perception. *American Journal of Sociology, 55 (Suppl.).*

Islam, M. R., & Hewstone, M. (1993). Intergroup attributions and affective consequences in majority and minority groups. *Journal of Personality and Social Psychology, 64,* 936-950.

Jones, E. E., & Davis, K. E. (1965). A theory of correspondent inferences: From acts to dispositions. In L. Berkowitz (Ed.), *Advances in Experimental Social Psychology* (Vol. 2, pp. 219-266). New York: Academic Press.

Keltner, D., & Robinson, R. J. (1997). Defending the status quo: Power and bias in social conflict. *Personality and Social Psychology Bulletin, 23,* 1066-1077.

Kemdal, A. B., & Montgomery, H. (2001). Explaining own and others' behavior in a controversial issue: Animal experimentation. *Journal of Social Psychology, 141,* 693-713.

Kenworthy, J. B., & Miller, N. (2002). Attributional biases about the origins of attitudes: Externality, emotionality, and rationality. *Journal of Personality and Social Psychology, 82,* 693-707.

Kruger, J., & Gilovich, T. (1999). "Naïve cynicism" in everyday theories of responsibility assessment: On biased assumptions of bias. *Journal of Personality and Social Psychology, 79,* 309-326.

Lee, F., & Robinson, R. J. (2000). An attributional analysis of social accounts: Implications of playing the blame game. *Journal of Applied Social Psychology, 30,* 1853-1879.

Locke, D., & Pennington, D. (1982). Reasons and other causes: Their role in attribution process. *Journal of Personality and Social Psychology, 42,* 212-223.

Malle, B. F. (1997). The folk concept of intentionality. *Journal of Experimental Social Psychology, 33,* 101-121.

Malle, B. F. (1999). How people explain behavior: A new theoretical framework. *Personality and Social Psychology Review, 3,* 23-48.

Malle, B. F. (2000). *F. Ex: A coding scheme for folk explanations of behavior (Version 4.0)* [On-line]. Retrieved December 17, 2002, from: http://darkwing.uoregon.edu/~interact/fex.html

Malle, B. F., Knobe, J., O'Laughlin, M. J., Pearce, G. E., & Nelson, S. E. (2000). Conceptual structure and social functions of behavior explanations: Beyond person-situation attributions. *Journal of Personality and Social Psychology, 79,* 309-326.

Miller, D. T., Monin, B., & Prentice, D. A. (2000). Pluralistic ignorance and inconsistency between private attitudes and public behaviors. In D. J. Terry & M. A. Hogg (Eds.), *Attitudes, behavior, and social context: The role of norms and group membership* (pp. 95-113). Mahwah, NJ: Erlbaum.

Miller, F. D., Smith, E. R., & Uleman, J. (1981). Measurement and interpretation of situational and dispositional attributions. *Journal of Experimental Social Psychology, 17,* 80-95.

Montgomery, H. (1994). Towards a perspective theory of decision making and judgement. *Acta Psychologica, 87,* 155-178.

O'Laughlin, M. J., & Malle, B. F. (2002). How people explain actions performed by groups and individuals. *Journal of Personality and Social Psychology, 82,* 33-48.

Pettigrew, T. (1979). The ultimate attribution error: Extending Allport's cognitive analysis of prejudice. *Personality and Social Psychology Bulletin, 5,* 461-476.

Pronin, E., Gilovich, T., & Ross, L. (in press). Objectivity in the eye of the beholder: Divergent perceptions of bias in self versus others. *Psychological Review*

Pronin, E., Lin, D. Y., & Ross, L. (2002). The bias blind spot: Perceptions of bias in self versus others. *Personality and Social Psychology Bulletin, 28,* 369-381.

Pronin, E., Puccio, C., and Ross, L. (2002). Understanding misunderstanding: Social psychological perspectives. In T. Gilovich, D. Griffin, & D. Kahneman (Eds.), *Heuristics and biases: The psychology of intuitive judgment* (pp. 636-665). Cambridge: Cambridge University Press.

Puccio, C. (2003). *The search for common ground: Overcoming false polarization and unwarranted pessimism about ideological differences.* Unpublished doctoral dissertation, Stanford University.

Robinson, R. J., & Friedman, R. A. (1995). Mistrust and misconstrual in union-management relationships: causal accounts in adversarial contexts. *International Journal of Conflict Management, 3,* 312-327.

Robinson, R. J., Keltner, D., Ward, A., & Ross, L. (1995). Actual versus assumed differences in construal: "Naïve realism" in intergroup perception and conflict. *Journal of Personality and Social Psychology, 68,* 404-417.

Ross, L., & Ward, A. (1996). Naive realism in everyday life: Implications for social conflict and misunderstanding. In T. Brown, E. Reed, & E. Turiel (Eds.), *Values and knowledge* (pp. 103-135). Hillsdale, NJ: Erlbaum.

Rouhana, N. N., O'Dwyer, A., & Morrison Vaso, S. K. (1997). Cognitive biases and political party affiliation in intergroup conflict. *Journal of Applied Social Psychology, 27,* 37-57.

Taylor, S. E. (1982). The availability bias in social perception and interaction. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgement under uncertainty: Heuristics and biases* (pp. 190-200). Cambridge: Cambridge University Press.

Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology, 5,* 207-232.

Van Boven, L., Kamada, A., & Gilovich, T. (1999). The perceiver as perceived: Everyday intuitions about the correspondence bias. *Journal of Personality and Social Psychology, 77,* 1188-1199.

Weiner, B. (1986). *An attributional theory of motivation and emotion.* New York: Springer Verlag.

Wilder, D. A., Simon, A. F., & Faith, M. (1996). Enhancing the impact of counterstereotypic information: Dispositional attributions for deviance. *Journal of Personality and Social Psychology, 71,* 276-287.

Coding examples

Coding category	Example
Causal History of Reason	
Agent stable propositional state	They do that because of their overwhelming patriotism.
Agent category membership	Those are mostly middle-aged and old people.
Agent trait	Unable to act independently.
Agent passive	They do not receive enough information about the EU.
Situation	There is not enough information available about EU.
Reason	
Desire	
Agent content (marked)	They want security for themselves.
Situation content (marked)	They do not want the country to become a member in any unions.
Agent + situation interaction (marked)	They want to enjoy the country's independence.
Agent content (unmarked)	(They support the EU membership) to achieve a better life quality for themselves.
Situation content (unmarked)	(They support the EU membership) to improve the living standard in the country
Belief	
Agent content (marked)	They think they are not going to lose anything.
Situation content (marked)	They think it will be safer for Latvia.
Agent + situation interaction (marked)	They think they are going to live like in Western countries.
Agent content (unmarked)	(They support the EU membership) because it will be safer for them.
Situation content (unmarked)	(They protest against the EU membership) because it will harm the agriculture.
Agent + situation interaction (unmarked)	EU membership will increase their chances to work in other countries.

Valuing	
Agent content (marked)	They are afraid of changes.
Situation content (marked)	They dislike globalisation.
Agent + situation interaction (marked)	They are afraid to be subdued by foreign powers.

Appendix B

Items that measure the strength of support for Latvia's EU membership

1. Latvia should enter the European Union.
2. The European Union membership will do more harm than good to Latvia (R).
3. Latvia will only gain by remaining outside the European Union (R).
4. I am personally against Latvia's membership into the European Union (R).
5. It would be better for majority of Latvians if Latvia became a member of the European Union.
6. European Union membership threatens the sovereignty of Latvia (R).
7. I personally support Latvia's membership into the European Union.

Note. Items 2, 3, 4, and 6 are reverse-scored. Cronbach's Alpha = 0.91

STUDY 4

Differentiating Explanations of Attitude-Consistent Behavior:
The Role of Perspectives and Mode of Perspective Taking

Girts Dimdins and Henry Montgomery
Department of Psychology, Stockholm University, Sweden
Ivars Austers
University of Latvia, Riga, Latvia

Abstract

We examined whether participants could differentiate between explanations of attitude-consistent behavior related to EU membership given from two perspectives (EU supporter and EU opponent) by means of three perspective taking modes (the explainer's own perspective, imagined in-group members' perspective, and imagined out-group members' perspective). Participants were presented with explanations provided from different perspectives and perspective taking modes, and they were asked to judge the extent to which they agreed with each explanation, to guess the attitude of the provider of each explanation, and to rate the quality of each explanation in various respects. Participants could not differentiate between explanations given by in-group members and out-group members who imagined the same perspective. They responded more favorably to explanations given from own perspective than from the imagined perspectives. The results suggest that there exists a shared understanding about how both sides should explain attitude-consistent behavior, but this understanding is measurably different from the actual explanations.

Key words: Shared reality, perspective taking, causal explanation, false polarization

A substantial part of intergroup research in social psychology has dealt with how people explain the behaviors of their in-group and out-group members. In his review of research of intergroup attribution, Hewstone (1989, 1990) reports a number of findings, which indicate robust effects of social categorization on explanations of behavior. Research designs and patterns of results have varied from study to study, but in most cases the underlying general finding has been a preference for in-group-serving attributions relative to out-group serving attributions (Hewstone, 1990). Perhaps this can be seen as a natural consequence of social categorization, where a favorable comparison of one's in-group to other groups serves as a powerful instrument of self-enhancement (Tajfel, 1978).

The robust nature of group differences in explaining social behavior has prompted a number of attribution studies where participants are asked to provide explanations from perspectives of other groups or individuals. A number of these studies have found that when making attributions from the imagined out-group perspective, participants simulate the out-group's motivation to have a positive and benevolent self-image, and accordingly bias their responses in favor of the out-group (Austers, 2002; Austers & Montgomery, 2001; Dimdins, Montgomery, & Austers, 2003a).

Other studies have looked at the similarities between the contents of attributions provided by members of various groups taking each other's perspectives. Using the number of internal and external causes as the criteria for comparison, Kemdal and Montgomery (2001) found that animal rights supporters and animal experimenters could fairly accurately take the other party's perspective in attributing causes to behavior of both groups. Dimdins, Montgomery, and Austers (2003b) asked EU supporters and opponents to explain the behavior of people supporting and opposing Latvia's EU membership (i.e., behavior expressing a positive or negative attitude) from their own perspectives and the perspectives of both groups. Explanations provided by supporters and opponents from the same group perspective were remarkably similar (the criterion for comparison was the ratio between two types of explanations, taken from the coding scheme suggested by Malle [1999, 2000]: reason explanations and causal history of reason explanations).

In the quoted studies, the effectiveness of the perspective taking has been explained as a result of a *shared reality* (Hardin & Higgins, 1996). According to this theoretical perspective, it is not surprising to find such a consensus regarding a topic of public debate. When two groups share the same social universe, both have access to largely the same information about the debated issues and each other, and through direct or indirect communication they exchange opinions on a regular basis. This is especially true in regard to attitudes about controversial issues. As Billig (1991) has pointed out, to have an attitude and to follow it in behavior means placing oneself in a public debate, where attitude-consistent behavior has to be justified and defended against

holders of other attitudes. In this context, taking an opponent's perspective helps the groups to find persuasive arguments against the opponent's position by predicting possible assaults on their own position and formulating arguments to defend their positions.

However, one aspect of the cited studies needs to be addressed. In all these papers, the researchers measured the accuracy of perspective taking. This was done either on the basis of similarities in quantitative ratings provided from different perspectives (Austers, 2002, Dimdins et al., 2003a), or similarities in frequency with which certain categories of content analysis occurred in responses provided from different perspectives (Austers & Montgomery, 2001; Kemdal & Montgomery, 2001, Dimdins et al., 2003b). It is likely that the broad theoretical and/or linguistic categories used as the measurement dimensions in these studies have failed to take into account subtler differences in the contents of behavioral explanations provided from different perspectives. For example, when explaining a certain out-group's behavior from that out-group's perspective, respondents might have used predominantly internal attributions, just like the members of the respective group itself. However, in regard to the specific contents, both groups could provide totally different explanations (e.g., "they are smart" and "they are not very smart"), yet these differences would not be captured in the content analysis using the broad internal versus external categories (both explanations would be coded as internal). This is an exaggerated example unlikely to occur in real life, but it serves to illustrate the shortcomings of the content analysis approach. Moreover, social and cognitive psychologists are far from being in agreement about what would be the appropriate dimensions for measuring attributions (e.g., Sabini, Siepmann, & Stein, 2001, Gilovich & Eibach, 2001; Ross, 2001; see also Malle, Knobe, O'Laughlin, Pearce, & Nelson, 2000).

An interesting and so far little explored aspect of group perspective taking is the question of what the group members themselves think about the explanations given by out-group representatives trying to take their group's perspective. To our knowledge, no published research addresses this question. But an early study of attitudes by Dawes, Singer, and Lemons (1972) showed that partisans in a political controversy were less successful in taking an out-group's perspective than in-group's perspective. Dawes et al. asked supporters and opponents of the Vietnam War to write statements that would be endorsed by members of one or the other group. When these statements where presented to the corresponding groups, both supporters and opponents rejected more statements written by their out-group members than by in-group members. This finding is not surprising, given that the previous research has documented a number of psychological mechanisms that may work against accurately taking an out-group perspective. Such mechanisms include *adversary's extremity bias* (Rouhana, O'Dwyer, & Morrison Vaso, 1997), influence of in-group norms

(Turner, 1991; Abrams, Wetherell, Cochrane, Hogg, & Turner, 2001), and motivational biases in seeking and evaluating information (Kunda, 1990).

The above arguments point to the importance of replicating the previous findings (that interpret the effectiveness of perspective taking as a consequence of shared reality) by using a design where the accuracy of perspective taking is not assessed by researchers but by the group members themselves. If the shared reality hypothesis is true, the simulated opinions produced by out-group members should be highly similar to the ones provided by in-group members themselves, reflecting the fact that both originate in the same shared reality. Consequently, members of a particular group should see no difference between arguments provided by their in-group members and those provided by out-group members from the perspective of the particular group. If, however, the group members would see such a difference (e.g., by favoring opinions expressed by in-group members over those expressed by out-group members taking the particular group's perspective), it would indicate that the out-group members are missing some subtle characteristics when simulating the opinions of the particular group. Such a result would challenge the shared reality hypothesis and the results of many studies demonstrating the efficiency of perspective taking (as estimated by the researchers).

Apart from validating the previous findings regarding the causal role of shared reality, such a design will also allow a test of a recently formulated explanation of the *false polarization effect* (overestimation of group differences on a certain attribute in comparison with the actual differences). Dimdins et al. (2003a, 2003b) found that both supporters and opponents of EU membership expected their in-group and out-group members to be more extreme in their explanations than the actual results showed. This was a replication of a well-documented phenomenon previously known from studies of beliefs and attitudes (see Pronin, Puccio, & Ross, 2002, for a review). Pronin et al. (2002) argue that the cognitive mechanism behind the false polarization effect is the "naïve realism" of the social perceiver (Ross & Ward, 1996). The social perceiver assumes that he or she has unbiased access to the objective reality. Others should perceive the reality the same way as the social perceiver does as long as their view is not biased by ideology or self-interest. Because the perceptions of others often differ from those of the social perceiver, this assumption results in an overestimation of psychological biases in other people (Pronin, Gilovich, & Ross, in press; Pronin, Lin, & Ross, 2002; Kruger & Gilovich, 1999). Consequently, when asked to provide an estimate of other people's opinions/attitudes, the social perceiver tends to come up with simplified, stereotypical, and exaggerated responses.

If the false polarization indeed is a result of assuming bias in others (in-group and out-group members alike), people should be able to see some differences between the opinions provided by their in-group and out-group members from their own perspective, and the simulated opinions provided from

4

the imagined group perspectives. Because the simulated opinions should reflect the expected bias, they should sound less rational and reasonable than the opinions from own perspective. Consequently, people should agree more to the opinions provided from own perspective than to the simulated opinions provided from the imagined group perspectives. If, however, there were no differences in agreement between the opinions expressed from own and imagined perspectives, it would indicate that people do not see the quality of the simulated arguments as inferior to that of the real arguments provided from own perspective (even though researchers may find signs of bias through content analysis or using quantitative rating scales). Such a result would challenge the assumption that the false polarization effect is a result of expectation of bias in other people. In that case, one might wish to test an alternative hypothesis that the false polarization effect is a result of moderating respondents' own opinions (e.g., because of social desirability) rather than exaggerating the group's position.

This article examines these assumptions in two studies regarding a controversial political issue (Latvia's membership in the European Union). The topic provides an excellent context for examining perspective taking in an intergroup context. Over the last two years, the debate on the issue has grown in force, with the opinions of both supporters and opponents well represented in mass media and highly salient.

In Study 1, we explored to what extent supporters and opponents of Latvia's EU membership agreed or disagreed with explanations of both groups' attitude-consistent behavior provided by other supporters and opponents from their own perspective, their in-group perspective, and their out-group perspectives. By attitude-consistent behavior we understand general, non-specific behavior that expresses the actor's attitude towards a certain issue. In other words, actors behaving in line with one's attitude rather than simply having a certain attitude. We chose attitude-consistent behavior instead of more specific group behaviors or attitudes in general because it allowed us to construct simple, context-independent, and unambiguous stimuli for our study. In Study 2, we examined in more detail what differences the participants could discern between the various explanations.

Study 1

Method

Sample. In this study, 137 students at the University of Latvia (107 women [78%]) participated. The participants were students of psychology, education, and languages. The mean age of participants was 21 years, and 118 participants (86%) were Latvians.

Questionnaire. The participants filled in a three-part questionnaire. First, the participants were asked to indicate their age, gender, ethnicity, as well as their general attitude towards Latvia's membership into the European Union

5

(pro vs. against). In the second part of the questionnaire, the participants were presented with descriptions of two opposite attitude-consistent behaviors and a number of explanations to these behaviors. The participants were then asked to indicate how much they agreed or disagreed with each explanation.

The descriptions of target behavior were as follows: (a) "There are many people in Latvia who actively protest against Latvia entering the European Union" and (b) "There are many people in Latvia who actively support Latvia entering the European Union". Each description was provided on top of a separate sheet. In half of the questionnaires, the description of supporter behavior was described first, followed by opponent behavior, and in the other half the order was reversed. After each description of behavior, 12 different explanations to this behavior were provided. All of these explanations were taken from the results of the earlier study by Dimdins, Montgomery, and Austers (2003b). In this study, EU opponents and supporters were asked to provide explanations (in a free response format) to the same two target behaviors from their own, imagined in-group, and imagined out-group perspectives. A number of these explanations were selected as stimuli for the present study.

The 12 explanations came from the six possible combinations of perspective (supporter or opponent) and mode of perspective taking (own perspective, imagined in-group perspective, and imagined out-group perspective). Thus, for each behavior there were 2 explanations provided from each combination of perspective and mode of perspective taking. The origin of each explanation (i.e., the perspective, from which the explanation had been provided and the mode of perspective taking) was not indicated in the questionnaire, and the order of explanations was arranged to separate those given from the same perspective.

The explanations included in the questionnaire had been selected according to the following scheme. All the participants in the study by Dimdins et al. (2003b) were ranked in the order of their strength of support for EU membership. Only explanations provided by self-labeled opponents ranking below the 25th percentile and self-labeled supporters above the 75th percentile of the resulting distribution were used. In this way it was ensured that the explanations came from participants with a distinct (stronger than average) attitude towards the issue. Explanations to the same behavior provided from the same perspective were then pooled together and randomly assigned to one of three different versions of the questionnaire. The assignment of items was not controlled regarding the contents of the explanations or the coding of their contents known from the previous content analysis of the items in order to avoid introducing method bias. All together, 72 different explanations were used, 12 from each combination of perspective and mode of perspective taking. Some examples of explanations used are provided in Table 1.

Table 1

Examples of explanations used as stimuli in the presented studies

Perspective and mode of perspective taking	Explanation
	Target: supporter behavior
Supporter perspective	
Own (i.e., supporter from own perspective)	"They see perspectives for development of the country, through cooperating with other European states." "Possibly they will feel safer from the threats of Russia, feel protected."
Imagined in-group (i.e., supporter taking supporter perspective)	"They believe in the EU." "They think: 'we want a better life, we want Latvia to be noticed'."
Imagined out-group (i.e., opponent taking supporter perspective)	"May be material values are more important to them than the moral ones." "They want to submit to someone, or cannot act independently of others."
Opponent perspective	
Own (i.e., opponent from own perspective)	"They do not want to work. [They act] according to the principle: 'EU will come and present it to us'." "Those, who will profit, support [the EU membership], forgetting about those who will be destroyed by the [European] Union.
Imagined in-group (i.e., opponent taking opponent perspective)	"May be they don't care what will happen to Latvia." "They do not have enough information about the negative [aspects of] EU."
Imagined out-group (i.e., supporter taking opponent perspective)	"The EU will support us both economically and militarily." "They want a living standard like that in the Western European countries."
	Target: opponent behavior
Supporter perspective	
Own	"May be these people are afraid of changes, afraid of financial loss." "There is not enough information about the EU model in general, a lack of general information in mass media."

Table 1 (continued)

	"They are not convinced that the EU will bring only good [things]. Afraid to lose something important."
Imagined in-group	"They don't understand the risk factor of Latvia not joining the EU. Latvia do not have resources like the Norwegians [do] to exist successfully outside the EU."
Imagined out-group	"They don't want Latvia to fall under someone's rule again." "They think Latvia should be independent, we need no masters."
Opponent perspective	
Own	"They think we will lose our identity and culture. From one union we will fall into another." "People want to take no risks, are afraid that the changes will harm them."
Imagined in-group	"They think: 'It is another Soviet Union'." "They act like this because they want to maintain Latvian identity -- [they] are patriots." "They are incompetent."
Imagined out-group	"Conservative people. Don't want to develop their country."

Table 2

Items for measuring the strength of support to Latvia's EU membership

1. Latvia should enter the European Union.
2. The European Union membership will do more harm than good to Latvia (R).
3. Latvia will only gain by remaining outside the European Union (R).
4. I am personally against Latvia's membership into the European Union (R).
5. It would be better for majority of Latvians if Latvia became a member of the European Union.
6. European Union membership threatens the sovereignty of Latvia (R).
7. I personally support Latvia's membership into the European Union.

Note. Items 2, 3, 4, and 6 are reverse-scored.
Reliability measures (Cronbach's Alpha): $\alpha_{Study\ 1} = .92$, $\alpha_{Study\ 2} = .88$.

The participants were asked to indicate their agreement or disagreement with each explanation on a 7-point Likert scale, with "1" labeled *"Strongly disagree"* and "7" labeled *"Strongly agree"*.

In the third part of the questionnaire, the participants were asked to indicate their agreement/disagreement with seven items (provided in Table 2) to

measure the strength of support to Latvia's membership into the European Union on a 5-point Likert scale.

Results and Discussion

An average agreement score was calculated for each perspective as the mean of the agreement scores with explanations provided from the same perspective for both target behaviors. Six agreement scores corresponding to the six perspectives described above were acquired for each participant. Figure 1 shows the means.

An index measuring the strength of support to Latvia's membership of the European Union was calculated for each participant by summing the responses to the seven items displayed in Table 2. The index values could thus vary between 5 and 35, higher scores indicating stronger support.

Seventy-eight participants (57%) identified themselves as supporters of Latvia's EU membership, and 57 (42%) identified themselves as opponents. Both groups differed significantly in their support for EU membership as measured by the mean support index (for supporter participants $M = 27.88$, $SD = 3.49$, and for opponents $M = 16.16$, $SD = 4.35$, $t(126) = 16.93$, $p < .001$).

A 2 (participant attitude: supporter vs. opponent) X 2 (perspective: supporter vs. opponent) X 3 (mode of perspective taking: own perspective vs. imagined in-group vs. imagined out-group) mixed ANOVA with repeated measures on the last two factors was carried out on the agreement index scores. We define perspective as the view from which the explanation had been provided. We define mode as the way the particular perspective had been taken. For the supporter perspective, the mode of perspective taking corresponds to *own perspective* when a supporter has given an explanation from his or her own perspective, to *imagined in-group* when a supporter has given an explanation from the supporter perspective (i.e., imagining how other supporters would explain the behavior), and to *imagined out-group* when an opponent has given an explanation from the supporter perspective (i.e., imagining how supporters would explain the behavior). Correspondingly, for the opponent perspective the mode of perspective taking corresponds to own perspective when an opponent has given an explanation from his or her own perspective, to imagined in-group when an opponent has given an explanation from the opponent perspective (i.e., imagining how other opponents would explain the behavior), and to imagined out-group when a supporter has given an explanation from the opponent perspective (i.e. imagining how opponents would explain the behavior).

Our first prediction was that the agreement with explanations would be determined to a much larger extent by the perspective from which the explanation had been provided (supporter or opponent perspective) than by the attitude of the explainer (as reflected in Participant Attitude X Mode of Perspective Taking interaction). As expected, there was a highly significant Participant Attitude X Perspective interaction, $F(1; 125) = 91.29$, $p < .001$,

9

partial $\eta^2 = .42$, accounting for the pattern that the participants agreed more with explanations provided from an in-group perspective than from an out-group perspective. The Participant Attitude X Mode of Perspective Taking interaction was not significant, $F(2, 125) = 2.69$, $p > .05$, partial $\eta^2 = .02$, and neither was the Attitude X Mode X Perspective interaction, $F(2, 125) = 1.45$, $p > .05$, partial $\eta^2 = .01$.
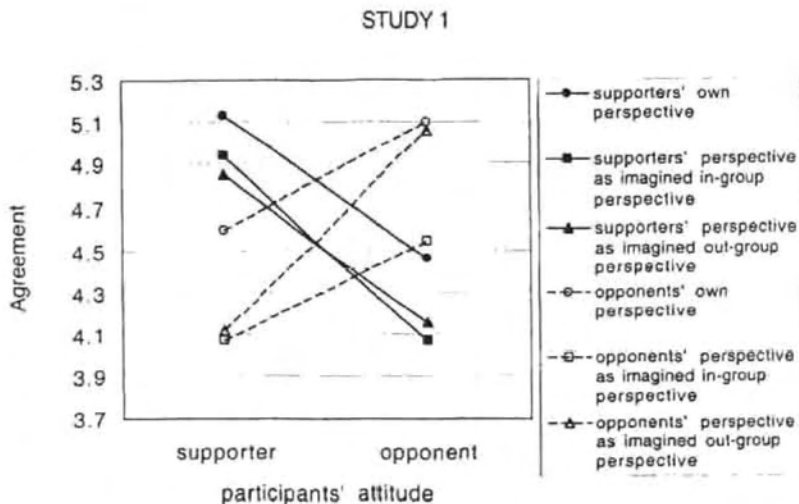
STUDY 1

Figure 1. Mean agreement with explanations provided from EU supporter and EU opponent perspectives taken as own, imagined in-group, and imagined out-group perspectives. Higher score indicates higher agreement.

In line with our second prediction, there was a significant main effect of mode of perspective taking, $F(2; 125) = 17.98$, $p < .001$, partial $\eta^2 = .13$, caused by the tendency that the participants agreed more with the explanations provided by explainers expressing their own opinion rather than the explanations given from the imagined group perspectives. Pairwise comparisons (Tukey LSD) showed that the estimated marginal mean for own explanations, $M = 4.82$ ($SE = 0.06$), was significantly higher than the mean for simulated in-group explanations, $M = 4.41$ ($SE = .06$), and the mean for simulated out-group explanations, $M = 4.55$ ($SE = .07$). Both group perspective means, in turn, did not differ significantly from each other. As Figure 1 shows, the agreement was higher when both perspectives were taken as own perspectives rather than as imagined in-group or out-group perspectives. The only exception to this pattern is that the opponent participants agreed equally strongly to the explanations

provided by opponents from their own perspective and to the simulated explanations by supporters taking the opponents' perspective.

Study 2

The aim of Study 2 was to expand the focus of Study 1 beyond simple agreement or disagreement with explanations. For example, it can be argued that a person may agree with a behavioral explanation even when knowing that the explanation has been provided by an out-group member (for example, if the person finds the explanation well substantiated or reasonable). In Study 2, we wanted to check if the participants would be able to guess the attitude behind each explanation when informed that the explanations might not reflect the true opinions of the explainers. Our prediction was that the participants would be able to determine only the perspective from which the explanation has been provided, but not the explainer attitude when that perspective had been taken as an out-group perspective. (Note that the explainer attitude and perspective coincide when a perspective is taken as an in-group perspective.)

Another aim of Study 2, related to the false polarization effect, was to examine in more detail what differences the participants would discern between the explanations provided from both perspectives when these perspectives are taken as own perspective, imagined in-group perspectives, or imagined out-group perspectives. Previous studies have shown that rationality is seen as a favorable cause of attitudes and attitude consistent behavior, more often attributed to in-group members than to out-group members (Kenworthy & Miller, 2002; Dimdins et al., 2003a). Dimdins et al. (2003a) found that attitude-consistent behavior was evaluated more positively from the actors' in-group perspective than from out-group perspective. In line with these findings, we predicted that rationality and positivity of the depicted behavior might influence the perception of differences between explanations. We also hypothesized that if an explanation has its origins in the shared reality, it should be seen as typical rather than idiosyncratic. Finally, we deemed it plausible to include a control variable of how substantiated the participants thought the explanations were, because it could arguably influence their agreement. However, these measurements were of exploratory nature, and no specific predictions were formulated regarding them.

Method

Sample. In this study, 89 students at University of Latvia (72 women [81%]) participated. All participants were students of education and languages. The mean age of participants was 19.5 years, and 77 participants (87%) were Latvians.

Questionnaire. The participants filled in a questionnaire, consisting of four parts. First, as in Study 1, the participants were asked to indicate their age,

gender, ethnicity, as well as their general attitude towards Latvia's membership into the European Union (for versus against).

In the second part of the questionnaire, the participants were presented with descriptions of two opposite behaviors and a number of explanations of these behaviors. The descriptions of target behaviors and the presented explanations were the same as in Study 1. The only difference was that for each target behavior there were only six explanations, one from each of the six combinations of perspective and mode of perspective taking. Six different versions of the questionnaire were created. The explanations came from the same pool of 72 explanations as in Study 1, but they were shuffled among the versions in such a way that those explanations, which appeared together in Study 1, were in different versions of questionnaires in Study 2.

After each explanation, the participants were asked five questions. First, the participants were asked to indicate their agreement/disagreement with each explanation (the same as in Study 1). Second, the participants were asked to rate how positive/negative the explanation was towards the target group (people whose behavior was being explained). The third question was to rate to what extent the target behavior was depicted as rational/irrational in the explanation. Fourth, the participants were asked how typical (common in Latvian society) the explanation was. Finally, the participants were asked to rate how well substantiated the explanation was.

All ratings were given on a 7-point Likert type scale. The sequence of explanations, questions, and target behaviors was counterbalanced across the questionnaires.

In the third part, the participants were informed that the explanations they had just rated had been provided by supporters and opponents of EU membership from their own perspective and imagined group perspectives. The participants were then asked to determine if the explanation had been initially provided by an EU supporter or an opponent. The explanations for each target behavior were presented again, and the participants were asked to indicate their opinion on a 7-point Likert scale: 1 was defined as *"I am sure that the explanation has been provided by an EU opponent"*; 7 was defined as *"I am sure that the explanation has been provided by an EU supporter"*; 4 was defined as *"I cannot say"*.

In the fourth part of the questionnaire, the participants were asked to indicate their agreement/disagreement with the seven items (provided in Table 2) that were designed to measure the strength of support for Latvia's membership into the European Union on a 5-point Likert scale.

Results and Discussion

An index measuring the strength of support for Latvia's membership of the European Union was calculated for each participant by summing the

responses to the seven items displayed in Table 2. As in Study 1, the index values could vary between 5 and 35; higher scores indicated stronger support.

Sixty participants (67%) identified themselves as supporters of Latvia's EU membership, and 28 (32%) identified themselves as opponents. Both groups differed significantly in their support to EU membership as measured by the mean support index (for supporter participants $M = 27.47$, $SD = 3.72$, and for opponents $M = 17.21$, $SD = 3.28$, $t(85) = 12.47$, $p < .001$).

STUDY 2
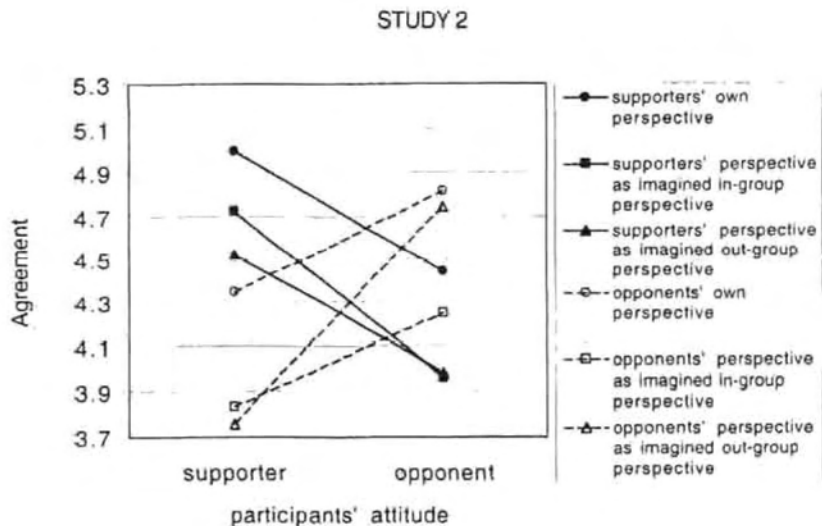
Figure 2. Mean agreement with explanations provided from EU supporter and EU opponent perspectives taken as own, imagined in-group, and imagined out-group perspectives. Higher score indicates higher agreement.

The general pattern of means for agreement scores largely resembled Study 1 (Figure 2). A 2 (participant attitude: supporter vs. opponent) X 2 (perspective: supporter vs. opponent) X 3 (mode of perspective taking: own perspective vs. imagined in-group vs. imagined out-group) mixed ANOVA with repeated measures on the last two factors was carried out on the agreement scores. The significant effects replicated those of Study 1, albeit with a smaller magnitude. There was a strong Perspective X Participant Attitude interaction, $F(1; 83) = 27.14$, $p < .001$, partial $\eta^2 = .25$, and a main effect of mode of perspective taking, $F(2; 83) = 9.05$, $p < .001$, partial $\eta^2 = .10$. The estimated marginal means were in the same direction as in Study 1: $M = 4.68$ ($SE = 0.10$) for own explanations, $M = 4.18$ ($SE = .11$) for simulated in-group explanations, and $M = 4.25$ ($SE = .12$) for simulated out-group explanations. The mean for own explanations was significantly higher than both means of simulated group

13

perspectives, which, in turn, did not differ significantly from each other. Again the results were similar to the results in Study 1, and none of the Participant Attitude X Mode of Perspective Taking interactions were significant.

The similar pattern of results in Studies 1 and 2 supports the reliability of the findings. It was repeatedly found that agreement with explanations was determined primarily by the perspective from which the explanations had been provided and not by the explainer attitude. Furthermore, in both studies we found that the participants tended to agree more with the explanations provided from own perspective than to the simulated explanations from the imagined group perspectives.

Of greater interest in the results of Study 2 are the participants' estimations of the attitude of providers of explanations. A 2 (participant attitude: supporter vs. opponent) X 2 (perspective: supporter vs. opponent) X 3 (mode of perspective taking: own perspective vs. imagined in-group vs. imagined out-group) mixed ANOVA with repeated measures on the last two factors was carried out on the attitude-guessing scores. There was a highly significant main effect of perspective, $F(1; 83) = 132.18$, $p < .001$, partial $\eta^2 = .61$. As can be seen in Figure 3, this effect corresponds to the tendency that the estimated explanation-provider attitude corresponded to the perspective from which the explanation had been provided, and it was independent of the actual provider attitude (involved in mode of perspective taking) when the explanations had been provided from an out-group perspective. In other words, the participants were able to "guess" the perspective of the explanation but not the explainer attitude when this attitude was not congruent with the perspective. These results are consistent with the initial predictions.

A result that breaks the overall pattern in Figure 3 is that the explanations provided by opponents from their own perspective in fact received an average score of $M = 4.15$ ($SE = .17$) – very close to the scale midpoint – from the supporter judges. This indicates that the supporter participants had difficulties determining the attitude of the explanation provider, and at least in some cases they ascribed the explanations to a supporter. The pattern results in significant Participant Attitude X Perspective X Mode of Perspective Taking, $F(2; 83) = 3.28$, $p < .05$, partial $\eta^2 = .04$, and Perspective X Mode of Perspective Taking, $F(2; 83) = 5.56$, $p < .01$, partial $\eta^2 = .06$, interactions, as well as in a main effect of mode of perspective taking, $F(2; 83) = 4.94$, $p < .01$, partial $\eta^2 = .06$, and a main effect of participant attitude, $F(1; 83) = 6.58$, $p < .05$, partial $\eta^2 = .07$. Perhaps something in the contents of (a number of) these explanations made them more appealing to the supporter participants and disguised the fact that they had been provided by an out-group member.
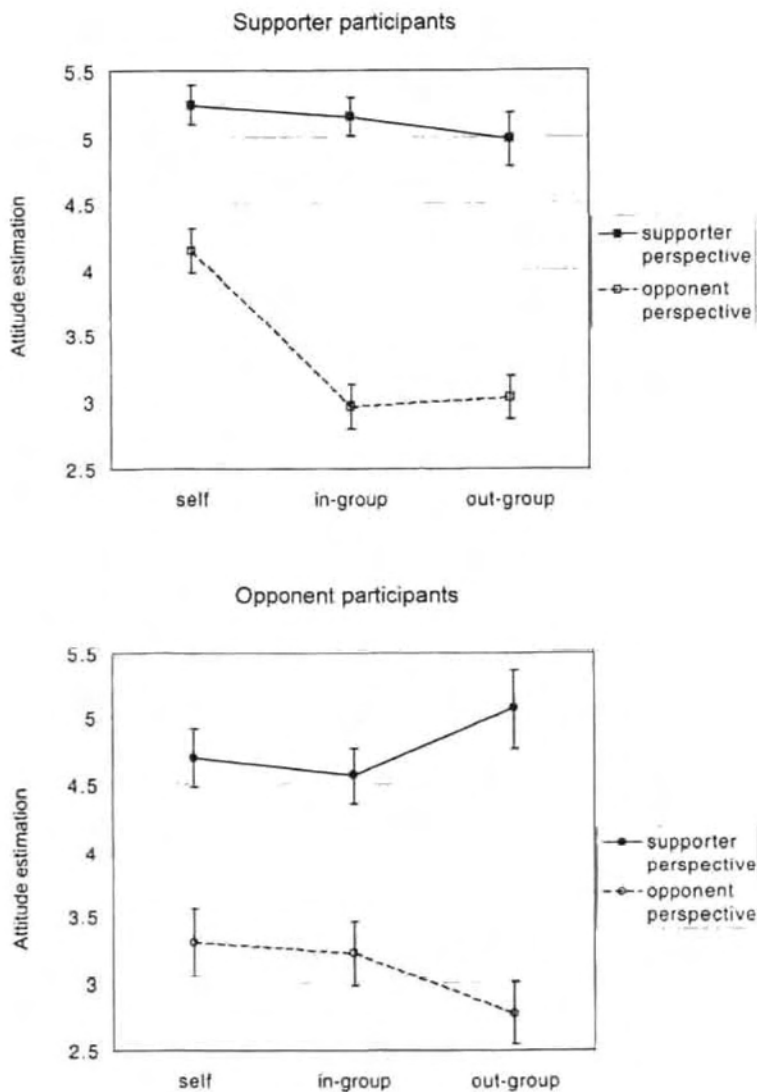
Supporter participants

Opponent participants

Figure 3. Mean scores for guessing the attitude of providers of explanations. Higher score indicates ascribing the explanations to EU supporters; lower score indicates ascribing the explanations to EU opponents; 4 stands for *"I cannot say"*. Vertical lines depict standard errors of the means.

ANOVAs were run also on the results of the four other questions asked about each explanation. The results were highly similar for all four scores, and they are summarized below according to the most often found effects. It should be noted that we did not predict any relations between these variables.

Therefore we carried out separate ANOVAs for each of the questions rather than including all of them in one analysis.

Main effect of mode of perspective taking. A main effect of mode of perspective taking was found in all four analyses, and in all cases the means were in the same direction. For *positivity* ratings, $F(2; 85) = 7.57$, $p < .01$, partial $\eta^2 = .08$, own explanations were rated as more positive towards the target group than the simulated in-group explanations (explanations given from the imagined in-group perspective), or the simulated out-group explanations (explanations given from the imagined out-group perspective). The corresponding estimated marginal means were $M = 4.58$ ($SE = 0.10$) for own explanations, $M = 4.14$ ($SE = .11$) for simulated in-group explanations, and $M = 4.22$ ($SE = .11$) for simulated out-group explanations.

For *rationality* ratings, $F(2; 82) = 8.08$, $p < .001$, partial $\eta^2 = .09$, own explanations ($M = 4.36$, $SE = 0.08$) were rated as depicting the target behavior as more rational than the simulated in-group explanations ($M = 3.94$, $SE = .10$) or simulated out-group explanations ($M = 4.05$, $SE = .10$).

The same result was found for *argumentation* ratings, $F(2; 82) = 4.68$, $p < .05$, partial $\eta^2 = .05$, own explanations ($M = 4.46$, $SE = 0.10$) being rated as better substantiated than the simulated in-group ($M = 4.12$, $SE = .11$) or out-group ($M = 4.18$, $SE = .12$) explanations.

For the ratings of *typicality*, $F(2; 80) = 4.30$, $p < .05$, partial $\eta^2 = .05$, the mean differences were smaller, although in the same direction. Own explanations ($M = 4.98$, $SE = 0.10$) were rated as more typical than the simulated out-group explanations ($M = 4.63$, $SE = .11$), but not significantly different from the simulated in-group explanations ($M = 4.75$, $SE = .11$).

Although the effect sizes are relatively small, it should be noted that the same significant effect was found for all dependent variables, and the means were always in the predicted direction. This is a clear indication that the participants could see differences between the explanations provided from own perspective and the imagined group perspectives.

Main effect of perspective. The main effect of perspective was the second most common effect found. Again, in all cases the means were in the same direction. Regarding the *positivity* ratings, $F(1; 85) = 23.63$, $p < .001$, partial $\eta^2 = .22$, the explanations provided from supporter perspective ($M = 4.54$, $SE = .08$) were rated as more positive than the explanations provided from opponent perspective ($M = 4.09$, $SE = .10$). The same effect, although of smaller magnitude, was present for ratings of *rationality*, $F(1; 82) = 7.91$, $p < .01$, partial $\eta^2 = .09$, $M_{sup} = 4.23$ ($SE = .08$), $M_{opp} = 4.00$ ($SE = .08$), and *typicality*, $F(1; 80) = 4.89$, $p < .05$, partial $\eta^2 = .06$, $M_{sup} = 4.90$ ($SE = .09$), $M_{opp} = 4.68$ ($SE = .10$).

In general, Study 2 confirmed our initial predictions. When asked to guess the attitude of the explanation providers, the participants could correctly

16

infer the perspective from which the explanation had been provided but not the explainer attitude (in those cases when the attitude was opposite to the perspective). Regarding the ratings of positivity, rationality, argumentation, and typicality, the participants were able to see some quality in the explanations provided from the explainers' own perspective, which differentiated them from the simulated explanations given from the imagined in-group and out-group perspectives. One may speculate that the same quality prompted the participants of this study, as well as those of Study 1, to agree more with explanations provided from the explainers' own perspective.

General Discussion

The studies reported in this paper tested whether group perspective taking is based on a shared reality that interacting social groups inhibit and construe. According to the shared reality hypothesis, most of the time the groups should be highly accurate when taking their opponent's perspective because the opinions of the parties involved in the debate are exchanged regularly and are freely accessible. We hypothesized that the behavioral explanations produced as a result of such perspective taking should be highly similar to the ones provided by the members of the particular group themselves. Furthermore, we hypothesized that when presented with these explanations, they should detect no difference between explanations provided by their in-group members, and those provided by out-group members from the perspective of their group.

Both studies supported the hypothesis. Our conclusion is that in general group members can recognize the perspective from which an explanation has been provided irrespective of whether the perspective has been taken by an in-group or out-group member. However, they are not able to distinguish effectively between perspectives that are congruent (in-group perspective) or incongruent (out-group perspective) with the explainer attitude. This leads us to speculate that there may be no perceivable basis for distinguishing between these two types of perspectives as long as they reflect the common knowledge about how both sides *should* explain both target behaviors. In general, our results agree with previous findings that concluded that group members were accurate in taking their opponents' perspective in various social contexts.

From a theoretical point of view, the most interesting finding in both the reported studies is the stability of mode of perspective taking effect. Regardless of the explainer attitude, the participants agreed more with the explanations provided from the explainers' own perspective than with those given from imagined in-group or out-group perspectives. Moreover, explanations provided from the explainers' own perspective, regardless of their attitude, were rated as more positive, typical, better substantiated. In addition, interpreting the target behavior was rated as more rational than the explanations provided from either of the imagined group perspectives. The general tendency is that the participants perceived own explanations more favorably than the simulated explanations.

We think that these results support the hypothesis that the false polarization effect results from overestimation of bias in other people. One may speculate that the participants' expectation of other people's bias in the Dimdins et al. (2003b) study resulted in producing lower quality (perhaps simplified, more extreme, or less reasonable) explanations from the imagined group perspectives, which, in turn, resulted in lower agreement with these explanations in the studies reported here.

On the one hand, the presented results can be seen as illustrating the efficiency of social debate in communicating the opinions between opposing parties. The explanations provided from an out-group perspective turned out to be of the same quality as those given from the in-group perspective. Explainers on both sides of the debate knew very well how the other side is supposed to explain a certain target behavior. On the other hand, the same results demonstrate the inferior quality of the simulated explanations if compared with the actual explanations from own perspective. It seems that people agree about what is the social norm in the case of explaining attitude-consistent behavior of EU supporters and opponents; however, they do not themselves completely agree with this norm. One may speculate that the false polarization in the explanations reflects a sort of pluralistic ignorance (Prentice & Miller, 1996; Miller, Monin, & Prentice, 2000). By definition, pluralistic ignorance is a widespread misperception of a group norm that leads to a situation where a majority of the group members privately disagree with the norm, but (erroneously) they assume that majority of other group members support it. Our results agree with such interpretation of the false polarization effect.

The results are especially strong because we found significant differences between how the explanations given from own perspective and the imagined in-group perspective were rated. In line with the findings on social influence and the in-group role in forming people's opinions (Turner, 1991; Abrams et al., 2001), it would have been reasonable to expect the explanations generated from own and in-group perspective to be highly similar. Our findings, however, suggest that the participants were distancing themselves (in terms of lower agreement) from the simulated explanations given from the in-group perspective as strongly as from those given from the out-group perspective. It seems that rather than representing a genuine shared social reality, the simulated explanations reflect a "shared pseudo-reality" – an agreement both between and within two opposing groups about how both groups are supposed to explain certain attitudes and attitude-consistent behaviors. However, this "reality" is measurably different from how the group members in fact explain such behaviors. Our results, along with other findings related to the false polarization effect, illustrate the complexity of the question of how shared reality should be defined and measured in the context of intergroup perception.

One result deserving a comment is that in Study 2 the explanations given from the supporter perspective were seen as more typical than those given from

the opponent perspective. For many years, all the major political forces in Latvia (including opposition parties) have been promoting EU membership as the main goal of Latvia's foreign policy, and for most of the 1990s pro-EU arguments dominated the mass media. Only during the last couple of years, the opinions of both sides have become more or less equally represented in the public debate. It is possible that this fact was reflected in the contents of the explanations. Moreover, because the pro-EU stance has represented the mainstream opinion for so many years, EU opponents might arguably be seen as partisans trying to change the status quo. Previous studies of partisanship and status quo (e.g., Kray & Robinson, 2001, Keltner & Robinson, 1997) have found that parties, which are seen as trying to change the status quo, are perceived as more extremist and evaluated less positively than parties advocating the status quo. This could also have been reflected in the contents of the explanations, making the explanations from the opponent perspective sound more extreme and from the supporter perspective sound more reasonable and moderate. These are of course post hoc speculations not related to the main predictions of the current study. However, these findings point to one possible area of further research.

Two other interesting directions emerge for the future research. First, the characteristics that make own explanations more "attractive" than the simulated group explanations seem to be context-free and independent of the perspectives from which they had been provided. Defining these characteristics and exploring their effects would improve our understanding of how biases are manifested and perceived in everyday discourse.

Another interesting question is whether the quality of arguments or behavioral explanations provided from a group perspective would increase with the individual's level of familiarity with the members of that group. In relatively small social groups, people should be familiar with the actual opinions of most, if not all, individual group members. This should reduce ungrounded assumptions of bias in others, at least regarding topics that the group members have sufficiently discussed among themselves. It is feasible to speculate that in such a case the simulated arguments or explanations provided from the imagined group perspectives would become truly indistinguishable from those provided from the explainer's own perspective. Some initial findings suggest that this speculation may be true (Kemdal & Montgomery, 2000). Further research in this direction might help finding ways to reduce the false polarization effect and thus improve the quality of intergroup communication and facilitate conflict resolution.

References

Abrams, D., Wetherell, M., Cochrane, S., Hogg, M. A., & Turner, J. C. (1991). Knowing what to think by knowing who you are: Self-categorization and the nature of norm formation, conformity and group polarization. In M. A. Hogg & D. Abrams (Eds.), *Intergroup relations: Essential readings* (pp. 270-288). Philadelphia, PA: Psychology Press.

Austers, I. (2002). Ability to take an out-group's perspective in explaining positive and negative behaviors. *Scandinavian Journal of Psychology, 43,* 407-412

Austers, I., & Montgomery, H. (2001). Explaining an out-group's political behavior from four different perspectives. *Baltic Journal of Psychology, 2,* 23-33.

Billig, M. (1991). *Ideology and opinions.* London: Sage.

Dawes, R. M., Singer, D., & Lemons, F. (1972). The contrast effect and its implications for intergroup communication and the indirect assessment of attitude. *Journal of Personality & Social Psychology, 21,* 281-295.

Dimdins, G., Montgomery, H., & Austers, I. (2003a). The false polarization effect in explanations of attitudinal behavior. *Current Research in Social Psychology, 8,* 275-301.

Dimdins, G., Montgomery, H., & Austers, I. (2003b). *False polarization when imagining others' explanations of attitudinal behavior.* Manuscript submitted for publication.

Gilovich, T., & Eibach, R. (2001). The fundamental attribution error where it really counts. *Psychological Inquiry, 12,* 23-26.

Hardin, C. D., & Higgins, E. T. (1996). Shared reality: How social verification makes the subjective objective. In R. M. Sorrentino, & E. T. Higgins, (Eds.), *Handbook of motivation and cognition.* (Vol. 3, pp. 28-84). New York: The Guilford.

Hewstone, M. (1989). *Causal attribution: From cognitive processes to collective beliefs.* Oxford: Basil Blackwell.

Hewstone, M. (1990). The "ultimate attribution error"? A review of the literature on intergroup causal attribution. *European Journal of Social Psychology, 20,* 311-335.

Keltner, D., & Robinson, R. J. (1997). Defending the status quo: Power and bias in social conflict. *Personality and Social Psychology Bulletin, 23,* 1066-1077.

Kemdal, A. B., & Montgomery, H. (2000). *Seeing it from the other side: Perspectives and attitudes towards animal experimentation.* Unpublished manuscript, Stockholm University.

Kemdal, A. B., & Montgomery, H. (2001). Explaining own and others' behavior in a controversial issue: Animal experimentation. *Journal of Social Psychology, 141,* 693-713.

Kenworthy, J. B., & Miller, N. (2002). Attributional biases about the origins of attitudes: Externality, emotionality, and rationality. *Journal of Personality and Social Psychology, 82,* 693-707.

Kray, L. J., Robinson, R. J. (2001). Partisanship and the status quo. *European Journal of Social Psychology, 31,* 321-335.

Kruger, J., & Gilovich, T. (1999). "Naïve cynicism" in everyday theories of responsibility assessment: On biased assumptions of bias. *Journal of Personality and Social Psychology, 79,* 309-326.

Kunda, Z. (1990). The case for motivated reasoning. *Psychological Bulletin, 108,* 480-498.

Malle, B. F. (1999). How people explain behavior: A new theoretical framework. *Personality and Social Psychology Review, 3,* 23-48.

Malle, B. F. (2000). *F. Ex: A coding scheme for folk explanations of behavior (Version 4.0)* [On-line]. Retrieved December 17, 2002, from: http://darkwing.uoregon.edu/~interact/fex.html

Malle, B. F., Knobe, J., O'Laughlin, M. J., Pearce, G. E., & Nelson, S. E. (2000). Conceptual structure and social functions of behavior explanations: Beyond person-situation attributions. *Journal of Personality and Social Psychology, 79,* 309-326.

Miller, D. T., Monin, B., & Prentice, D. A. (2000). Pluralistic ignorance and inconsistency between private attitudes and public behaviors. In D. J. Terry & M. A. Hogg (Eds.), *Attitudes, behavior, and social context: The role of norms and group membership* (pp. 95-113). Mahwah, NJ: Erlbaum.

Prentice, D. A., & Miller, D. T. (1996). Pluralistic ignorance and the perpetuation of social norms by unwitting actors. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 28, pp. 161-209). San Diego, CA: Academic Press.

Pronin, E., Gilovich, T., & Ross, L. (in press). Objectivity in the eye of the beholder: Divergent perceptions of bias in self versus others. *Psychological Review*

Pronin, E., Lin, D. Y., & Ross, L. (2002). The bias blind spot: Perceptions of bias in self versus others. *Personality and Social Psychology Bulletin, 28,* 369-381.

Pronin, E., Puccio, C., & Ross, L. (2002). Understanding misunderstanding: Social psychological perspectives. In T. Gilovich, D. Griffin, & D. Kahnemen (Eds.), *Heuristics and biases: The psychology of intuitive judgement* (pp. 636-665). Cambridge: Cambridge University Press.

Robinson, R. J., Keltner, D., Ward, A., & Ross, L. (1995). Actual versus assumed differences in construal: "Naïve realism" in intergroup perception and conflict. *Journal of Personality and Social Psychology, 68,* 404-417.

Ross, L. (2001). Getting down to fundamentals: Lay dispositionism and the attributions of psychologists. *Psychological Inquiry, 12,* 37-40.

Ross, L., & Ward, A. (1996). Naive realism in everyday life: Implications for social conflict and misunderstanding. In T. Brown, E. Reed, & E. Turiel (Eds.), *Values and knowledge* (pp. 103-135). Hillsdale, NJ: Erlbaum.

Rouhana, N. N., O'Dwyer, A., & Morrison Vaso, S. K. (1997). Cognitive biases and political party affiliation in intergroup conflict. *Journal of Applied Social Psychology, 27,* 37-57

Sabini, J., Siepmann, M., & Stein, J. (2001). The really fundamental attribution error in social psychological research. *Psychological Inquiry, 12,* 1-15.

Tajfel, H. (1978). Social categorization, social identity and social comparison. In H. Tajfel (Ed.), *Differentiation between social groups: Studies in the social psychology of intergroup relations* (pp. 61-76) London: Academic Press.

Turner, J. C. (1991). *Social influence.* Milton Keynes: Open University Press

ERRATA

Introduction, page 13, 2nd paragraph, line 2.
Instead of "...others are less susceptible..." it should be "...others are more susceptible."

Introduction, page 32, 2nd paragraph, line 3.
Instead of "...using the BEB coding scheme..." it should be "...using the FEB coding scheme."

Study 2, page 15, 1st paragraph.
Instead of "*Journal of Abnormal Psychology, 18*, 317-334" it should be "*Journal of Abnormal Psychology, 87*, 49-74*".

Study 2, page 17, 3rd paragraph.
Instead of "*Journal of Cross-Cultural Psychology, 119*, 162-171" it should be "*Journal of Cross-Cultural Psychology, 5*, 162-171"

Study 3, page 13, 3rd paragraph.
Instead of "*Baltic Journal of Psychology*" it should be "*Journal of Baltic Psychology*".

Study 4, page 20, 1st paragraph.
Instead of "Abrams, D., Wetherell, M., Cochrane, S., Hogg, M. A., & Turner, J. C. (1991)." it should be "Abrams, D., Wetherell, M., Cochrane, S., Hogg, M. A., & Turner, J. C. (2001)."

Study 4, page 20, 3rd paragraph.
Instead of "*Baltic Journal of Psychology*" it should be "*Journal of Baltic Psychology*".

Shared Reality and False Polarization in Intergroup Perception

Girts Dimdins

Girts Dimdins
ISBN 91-7265-785-5
Akademitryck, 2003

Psykologiska institutionen
Stockholms Universitet
S-106 91 Stockholm

Abstract

False polarization is an overestimation of a gap between the views of two sides that occurs because social perceivers expect their peers to be more susceptible to group norms than themselves. Shared reality is a generic term denoting social verification of information by individuals and groups. This thesis explores the relation between shared social reality and false polarization in situations where two groups take the perspectives of each other. A special emphasis in the thesis is put on intergroup attributions. In Study 1, two experiments examined the differences between evaluative and descriptive aspects of intergroup perception. In Experiment 1, Latvian and Russian students agreed on the descriptive scores of mutual ratings, but they disagreed on the evaluative ones. In Experiment 2, male and female psychology students did not show any disagreement either for evaluative or descriptive scores, but they expected their peers to be in-group biased in evaluation of both groups. In Study 2, supporters and opponents of Latvia's EU membership exhibited a false polarization effect when rating various causal explanations of attitudinal behavior of both groups. Study 3 replicated this result with free response causal explanations. Content analysis of the data showed that simulated explanations of a target behavior given from the opposite group perspectives differed more (in terms of perceiving the actors' awareness of causes of behavior and in terms of using mental state markers) than those given from the partisans' own perspective. In Study 4, two experiments demonstrated that EU supporters and opponents could not differentiate between explanations given by in-group members and out-group members who imagined the same perspective and responded more favorably to explanations given from the explainers' own perspective than from the imagined perspectives. In general, the results demonstrate the robustness of the false polarization effect across various contexts. The findings also illustrate the complexity of the term "shared social reality." Furthermore, in the context of intergroup perception, these findings suggest that the most appropriate operationalization of this concept should be shared stereotypes between groups.

Key words: Shared reality, social reality, false polarization effect, social perception, perception of bias, social categorization, in-group favoritism, intergroup perception, intergroup attribution, causal explanation.