European Journal of Social Psychology, Vol. 18, 1–16 (1988)

The 'Black Sheep Effect': Extremity of judgments towards ingroup members as a function of group identification*

JOSÉ M. MARQUES, VINCENT Y. YZERBYT and JACQUES-PHILIPPE LEYENS
Catholic University of Louvain,
Louvain-la-Neuve,
Belgium

Abstract

The present study proposes an extension to the phenomenon of ingroup favouritism, based on the hypothesis that judgments about ingroup members may be more positive or more negative than judgments about similar outgroup members. It contrasts predictions issued from the complexity–extremity hypothesis (Linville, 1982; Linville and Jones, 1980), from the ingroup favouritism hypothesis (Tajfel, 1982) and from Tesser's (1978; Millar and Tesser, 1986) attitude polarization model. Our main prediction, based on Social Identity Theory, is that judgments about both likeable and unlikeable ingroup members are more extreme than judgments about outgroup members. This phenomenon, coined the Black Sheep Effect, is viewed as due to the relevance that ingroup members' behaviour, as compared to that of outgroup members, has for the subjects' social identity. Three experiments supported our predictions. Experiment 1 additionally showed that inter-trait correlations were stronger for the ingroup than for the outgroup. Experiment 2 showed that the black sheep effect occurs only when the judgmental cues are relevant for the subjects' social identity, and Experiment 3 showed that levels of information about the target of the judgment were ineffective in generating judgmental extremity. Results are discussed in light of a cognitive–motivational alternative explanation to a purely cognitive interpretation of outgroup homogeneity.

INTRODUCTION

Social scientists have long attempted to explain the way individuals stereotype ingroup and outgroup members (e.g. Ashmore and Del Boca, 1981; Miller, 1982), either in

This paper is based on the first author's doctoral dissertation under the supervision of the third author. Requests for reprints should be sent to Vincent Yzerbyt, University of Louvain, Department of Psychology (EXSO), 20 Voie de Roman Pays, B-1348 Louvain-la-Neuve, Belgium or to the following BITNET/EARN address: YZERBYT@BUCLLN11 José M. Marques is at the Air Force Academy. Portugal.

0046–2772/88/010001–16$08.00
© 1988 by John Wiley & Sons, Ltd.
Received 11 May 1987
Revised 9 September 1987

cognitive–motivational terms (e.g. Allport, 1954; Tajfel, 1982), or in terms of cognitive factors alone (e.g. Hamilton, 1981; Taylor, 1981; Rothbart, 1981).

One of the most important versions of the cognitive–motivational orientation is Tajfel and colleagues' Social Identity Theory (*cf.* Tajfel, 1982; Turner 1975). Ingroup favouritism is a core-concept of this theory. It has been described as any tendency to favour ingroup members over outgroup members in perceptual, attitudinal or behavioural domains (Turner, 1981). It has been further conceived of as a strategy of self-enhancement through the social component of the individual's self-concept, that is, *social identity* (Tajfel and Turner, 1979). Social identity has been defined as 'the part of the individuals' self-concept which derives from their knowledge of their membership of a social group together with the value and emotional significance of that membership' (Tajfel, 1982, p. 24). According to Social Identity Theory, ingroup favouritism arises as a means of preserving a positive differentiation of the ingroup as compared to an outgroup along relevant comparison dimensions (Mummendey and Schreiber, 1984; Rijsman, 1982; Tajfel, 1978; Turner, 1975; Turner, Brown and Tajfel, 1979). Ingroup favouritism seems to be highly pervasive (Brewer, 1979; Brewer and Kramer, 1985), and to arise regardless of interpersonal similarity between ingroup and outgroup members (e.g. Allen and Wilder, 1975; Billig and Tajfel, 1973), of previous acquaintance among subjects (e.g. Tajfel, Billig, Bundy and Flament, 1971), of familiarity with the experimental setting (e.g. Tajfel and Billig, 1974), or of experimenter's explicit of implicit demands (e.g. Billig, 1973; St Claire and Turner, 1982).

Intergroup contact and the complexity–extremity hypothesis

The above standpoint has been recently challenged by authors who view stereotyping as an outcome of lack of intergroup contact (*cf.* Miller and Brewer, 1984; Rose, 1981; Stephan, 1985; Stephan and Rosenfield, 1982). These authors have contended that accuracy in intergroup perception is a direct function of the frequency of interpersonal interactions between members of different groups. Repeated contacts with a stimulus-domain should increase the likelihood of encountering additional information, be it consistent with, inconsistent with, or irrelevant to prestored schemata representing that stimulus-domain (e.g. Hastie, 1981; Higgins, Kuiper and Olson, 1981; Taylor and Crocker, 1981; Weber and Crocker, 1983). As a result, schemata about ingroups should be more complex than schemata about outgroups, because individuals interact more often with ingroup members than with outgroup members.

Following the reasoning summarized above, Linville (1982; Linville and Jones, 1980) consistently found that the more complex the schema for a category, the less extreme are the subsequent judgments about an instance of that category. These experiments are known under the heading of *complexity–extremity* hypothesis. In the social domain, the important result of the Linville's studies for our concerns is that likeable outgroup members were judged more positively than likeable ingroup members, and that unlikeable outgroup members were judged more negatively than unlikeable ingroup members. Thus, at least with respect to likeable targets, the results obtained by Linville and by Linville and Jones are the opposite of what should be expected from the simplest version of the ingroup favouritism hypothesis. This hypothesis would predict that likeable ingroup members should be more positively evaluated than likeable outgroup members. It is true that Linville (1982) described

how differences in extremity can coexist with a group main effect: In this case, she argued, one would obtain a higher mean for ingroup evaluations than for outgroup evaluations with subjects pooled across the likeability conditions, along with the larger difference between likeable and unlikeable outgroup targets than between likeable and unlikeable ingroup targets. However, to our knowledge no data whatsoever came in support of this speculation.

Complexity–extremity and attitude polarization

Not only is the complexity–extremity hypothesis at odds with Social Identity Theory, but it is also apparently in contradiction to the attitude polarization model of Tesser (1978; see also Fiske and Taylor, 1984). According to Tesser (1978), when provided with the opportunity to think about one stimulus, subjects' evaluations of that stimulus will be more extreme if they have a well-developed schema for it than if they have a simpler one.

At face value, thus, the attitude polarization model predicts a pattern of results contrary to the complexity–extremity hypothesis. However, according to Linville (1982), the contradiction between her results and those obtained by Tesser (1978; Tesser and Leone, 1977) is only apparent. Linville's argument is that, because her subjects were asked to make one judgment at one point in time, whereas Tesser's subjects were asked to make evaluations at two different points in time, the two predictions are not comparable to a single pattern of results. Millar and Tesser (1986), on the other hand, have argued against Linville's explanation. According to these authors, Linville's subjects did engage in a thought process while reading the information they were provided with. However, they went on, judgmental polarization occurs for simple schemata only when the subject has no commitment to his/her previous evaluation. Concomitantly, complex schemata yield polarization when the subject is committed to a previous judgment about the same stimulus. Therefore, they concluded, the results obtained by Linville (1982; Linville and Jones 1980) were possible because subjects have had no previous commitment to their judgments either about ingroup or about outgroup members.

In light of our point, then, Millar and Tesser's (1986) argument may be phrased as follows: The stimulus-materials employed by Linville and by Linville and Jones were, probably, irrelevant as far as subjects' ingroup identification is concerned. This assumption is further strengthened by examining the kind of stimulus-materials which have been employed in the tests of the complexity–extremity hypothesis: Providing subjects with rich descriptions of each stimulus person, as has been done, may well create the optimal conditions for subjects to discard stereotype-driven information, or subjective base-rates, and to apply to individualized information. Such 'dilution effect' (e.g. Locksley, Borgida, Brekke and Hepburn, 1980; Nisbett, Zukier and Lemley, 1981) may as well account for results on the extremity effect. That is, most likely, Linville's (1982) and Linville and Jones's (1980) argument is more relevant to an interpersonal than to an intergroup level of judgment (*cf.* also Tajfel, 1978). However, it is worth noticing that Tesser's (1978) attitude polarization model yields predictions which conflict to ingroup favouritism. As a matter of fact, according to the model, unlikeable outgroup members should be evaluated less negatively (i.e. less extremely) than unlikeable ingroup members.

The black sheep hypothesis

The *black sheep hypothesis* states that judgments about likeable and unlikeable ingroup members should yield more extreme positive and negative evaluations than judgments about similarly likeable and unlikeable outgroup members[1]. Taken at face value, thus, our hypothesis might seem identical to Tesser's (1978). But, we see the black sheep effect as due to identification with one's group rather than to heightened complexity of its representation.

Also, we must admit that if the prediction of favouritism towards likeable ingroupers is obvious in light of Social Identity Theory, the prediction of derogation of unlikeable ingroupers does not go without problems. Still, some studies have highlighted phenomena akin to our point of view. For instance, it has been shown that affective states induce grouping of more stimuli together and leniency in judging the criteriality for stimulus inclusion in categories, thus inducing stronger judgmental extremity than 'cold' states (Isen and Daubman, 1984). More related to the group domain, authors such as Jones and DeCharms (1957), Moreland and Levine (1982), Schachter (1951), and Sherif and Sherif (1979) have shown that behaviour of ingroup members is judged in light of subjects' interests. For example, in one of his 'communication networks' studies, Schachter (1951) has shown that a member of a discussion group who is consistently at odds with the other members of that group ends up by being marginalized after a series of essays on the part of the majority to bring him back to the mainstream. Sherif and Sherif (1979) reported that group leaders who are unable or unwilling to meet the group's desires are ultimately overthrown. Even closer to our point, Jones and DeCharms (1957) have shown that in situations of interdependence for rewards, group members who contribute negatively to the common group-outcome are more unfavourably evaluated than similar persons when the rewards depend exclusively on one's own performance. Assuming that commitment to certain group values and reward outcomes have a functional status similar to the enhancement of self-esteem through membership of a group, these results indicate that, because of their relevance to the subjects, unlikeable ingroup members may be judged more negatively than unlikeable outgroup members.

The above idea is indirectly supported by two recent studies. Meindl and Lerner (1984), for instance, showed that initial attitudinal and behavioural dispositions towards the outgroup are exacerbated when the subject's self-concept (*viz.* social identity) is threatened. These authors analyzed situations in which threat comes from outside the group, and their manipulation yielded polarization of negative attitudes towards the outgroup. However, their data allows to suppose that negative attitudes would be polarized towards threatening ingroup members as well. Also, Mackie and Cooper (1984) presented evidence that subjects involved with their group membership perceive the ingroup's norms as being significantly more extreme than do uninvolved subjects. One may thus suppose that severity towards ingroup members who do not comply with the norm would increase as a function of social identification. In the whole, the studies described above allow to suppose that negative attitudes will be exacerbated toward threatening ingroup members.

[1] The term 'black sheep effect' refers to a simultaneous emergence of ingroup bias and outgroup bias. The two phenomena are complementary, namely because the term black sheep implies that those to which it refers be judged as ingroup members.

The black sheep hypothesis is thus in accordance with Social Identity Theory: The under-evaluation of dislikeable ingroup members may be an acceptable psychological strategy for preserving the group's *overall* positivity. Therefore, the black sheep effect should be considered a 'sophisticated' form of ingroup favouritism.

Overview of the experiments

Experiment 1 was aimed at providing a first test of the black sheep effect. In Experiment 2, we attempted to validate the hypothesis that the black sheep effect will emerge only for issues which have normative relevance specific to the group with which the subject is currently identifying. Experiment 3 was aimed at attenuating a 'pure' perceptual–cognitive explanation of judgmental extremity for ingroups. If ingroup judgmental extremity were due exclusively to subjects' heightened familiarity with ingroup members as compared to outgroup members, then, the black sheep hypothesis would run unsupported. Rather Tesser's (1978) attitude polarization model would receive support. Indeed, in stating our hypothesis, we assume that judgmental extremity is a function of ingroup identification and, thus, that it depends on cognitive–motivational factors rather than on the quantity of information possessed about the judgmental target. Consequently, it was predicted that group identification rather than familiarity with the judgmental domain, will trigger extremity of negative judgments.

EXPERIMENT 1

In this study, we manipulated two variables, one pertaining to group membership (Belgian versus North African)[2] and the other to a value dimension (likeable, neutral, and unlikeable students). Our prediction is that the ingroup likeable target will be judged more positively than the outgroup likeable target, whereas the ingroup unlikeable target will be judged more negatively than the outgroup unlikeable target. This corresponds to the black sheep effect. Additionally, we measured the relative strength of intercorrelations between ingroup judgments and between outgroup judgments.

Method

Subjects

Ninety-six male and 88 female Belgian students from the University of Louvain-la-Neuve, aged 18 to 24, volunteered to participate in this study[3]. They were approached

[2] In order to make sure that relevant ingroup and outgroup labels were chosen, we asked one group of students to make sociometric choices aimed at pinpointing which nationalities they would favour or reject for their room-mates. Clearly, the North African category was the one which corresponded the most to an outgroup. This result might be explained by the fact that North African immigrants, coming from Maghrebian countries, seem to have a widespread negative image in Belgium.

[3] In order to make sure that the trait-descriptors used in the analysis actually had positive or negative valences, 46 subjects out of these 184 were asked to rate one of two stimuli — 'likeable students' and 'unlikeable students' — on the 62 traits. They were excluded from further analyses.

at random in public places on campus. A Belgian interviewer asked them whether they were students at the University of Louvain and, if the answer was affirmative, whether they would agree to fill in a questionnaire.

Procedure

The questionnaire presented a stimulus-target followed by 62 trait-descriptors[4]. Subjects were asked to rate the stimulus according to the trait-descriptors on a 7-point scale ranging from 7(=agree) to 1(=disagree). Only 20 positive and 18 negative trait-descriptors were analyzed[5]. Subjects were divided into six groups (n=23), each group being presented with one stimulus: 'Belgian students', 'North African students', 'Unlikeable Belgian students', 'Unlikeable North African students', 'Likeable Belgian students' and 'Likeable North African students'.

Results and discussion

The black sheep hypothesis

We averaged the scores of each subject on the 38 traits after reversing the scores on the 18 negative traits (Cronbach's alpha for the 38 items is 0.95). A 2 (Belgian versus North African) × 3 (likeable versus neutral versus unlikeable) ANOVA was performed. A significant main effect was found for likeability $F(2, 132) = 33.67$, $p < 0.001$. No main effect was found for target's group membership, $F(1, 132) = 1.79$, $p = 0.18$. However, a highly significant interaction emerged between likeability and target's group membership, $F(2, 132) = 9.97$, $p < 0.001$. A Newman–Keuls multiple comparison test on the cell means showed, as predicted, that the ingroup likeable target was more positively evaluated than the outgroup likeable target (*viz.* ingroup favouritism), and, the ingroup unlikeable target was judged more negatively than the outgroup unlikeable target (*viz.* ingroup derogation). The co-occurrence of favouritism to and derogation of likeable and unlikeable ingroup members, respectively, supports the black sheep hypothesis (see Table 1).

Cognitive complexity of ingroup and outgroup categories

The raw scores of the subjects on the 38 items were submitted to an iterated principal factors analysis. Six factors were extracted for the total amount of variance. The

[4] The 62 trait-descriptors were the attributes most frequently assigned to the ingroup and the outgroup on a free-association task by 117 participants in another study. A correspondence factor analysis computed on these traits yielded a first factor which was clearly interpretable as a likeability dimension. This is why we presented our subjects with 'likeable' and 'unlikeable' target depictions. Importantly, the second factor clearly opposed the 'Belgian' to the 'North African' category so that there is a near-perfect orthogonal relationship between the four categories. This precludes any influence of trait typicality on ingroup and outgroup evaluations.

[5] Only items which differed significantly between the 'likeable' and 'unlikeable' stimuli were used in the analysis of ingroup and outgroup evaluations. The 20 positive trait-descriptors (here translated from the French) were: pleasant, sociable, communicative, cheerful, welcoming, kind, good, helpful, jovial, cool, stand-together, honest, enthusiastic, self-willed, motivated, alive, intelligent, polite, cultured, and athletic; and 18 negative trait-descriptors were: unpleasant, conceited, snob, violent, ill-natured, irrespectful, male-chauvinistic, profiteer, cold, shallow-minded, sectarian, apathetic, xenophobic, non-approachable, hung-up, depressed, non-active, and pessimistic.

Table 1. Favourability ratings as a function of likeability and target's group membership

Target	Likeability Likeable	Neutral	Unlikeable
Ingroup			
Mean*†	5.20 d	4.26 cb	2.97 a
S.D.	0.81	0.89	0.86
Outgroup			
Mean	4.55 c	4.56 c	3.91 b
S.D.	0.74	1.05	0.73

* Ratings on a 7-point scale ranging from 7 (=applies) to 1 (=doesn't apply). The scores of negative traits were reversed. Higher scores mean more positive ratings.
† Means containing the same letter are not different at a 0.05 level of significance as given by Newman–Keuls test.

arithmetical averages of factor scores of subjects on each factor provided the centroids of the stimulus-categories (see Figure 1). Only the first two factors, accounting for 76.2 per cent of the variance, were taken into further consideration.

As can be seen in Figure 1, the outgroup categories are considerably nearer to each other than the ingroup categories. This fact allows to derive that inter-trait correlations are more strongly negative between the likeable and unlikeable ingroup categories than between their outgroup counterparts. This comes as a validation of the first analysis and shows even more clearly that subjects endorsed a 'manichean' strategy of evaluations of the ingroup.

One comment should be made bearing to the methodology we developed in this study. One might argue that, to some extent, the procedure we used is complementary to that employed by Linville (1982; Linville and Jones, 1980). Whereas Linville has provided her subjects with value-laden descriptions and asked them for an evaluation, we provided ours with ready-made evaluations and asked them to describe the targets using evaluative traits. However, this distinction is less important than it might appear at face value. In both cases, subjects were given a value-laden information about a person and were to make evaluative judgments about this person. The basic point as to

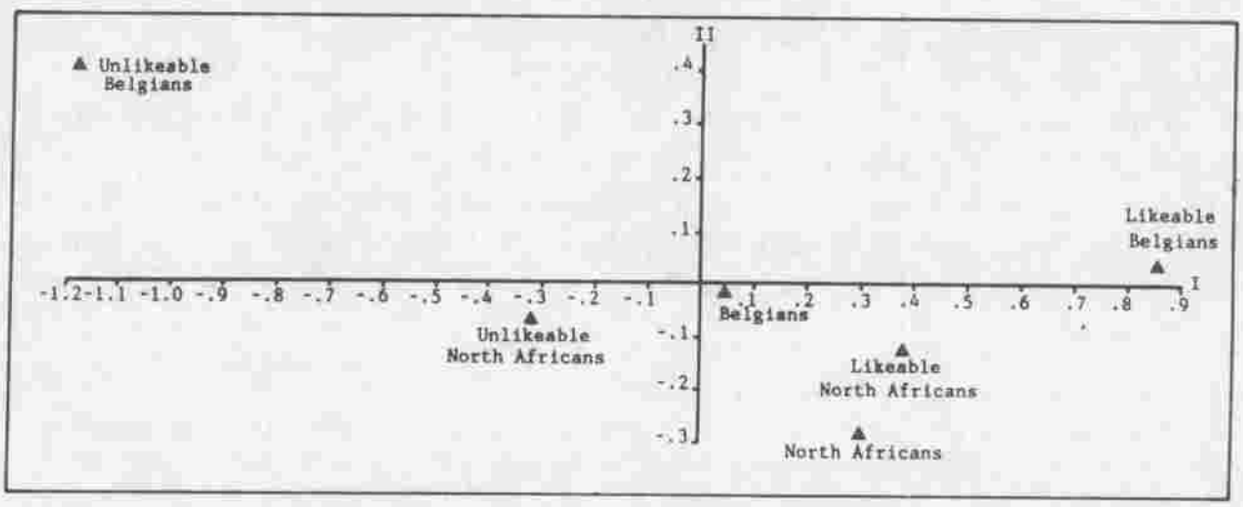

Figure 1. Stimulus-categories plotted on the first two factors of the iterated principal factor analysis

our results is that the way subjects did 'reproduce' our *a priori* descriptions strengthens even more our assumptions: Instead of merely restating the objectively unequivocal information they received, the subjects significantly biased such information on the basis of knowledge about the targets' membership groups alone.

However, this result still might be explained in terms of Tesser's (1978) attitude polarization model. That is, it might be that the subjects merely polarized their judgments for targets about whom they held more knowledge rather than with whom they identified. The following experiments aim at ruling out this alternative interpretation.

EXPERIMENT 2

In Experiment 2, subjects were provided with a description of ingroup or outgroup members who complied with versus opposed to a norm. In addition, whereas in one condition this norm applied to a general category including the ingroup and the outgroup ('students'), in the other, it applied exclusively to the ingroup ('Belgian students'). It was predicted that no differential judgmental extremity would emerge for the general norm. However, ingroup judgments would be more extreme than outgroup judgments for the exclusive norm.

Method

Subjects

Fourty-three male and 48 female undergraduates of Belgian nationality, aged 18 to 23, volunteered to participate in a study about 'certain aspects of the students' life on campus'. The number of subjects in each condition varied between 8 and 14. Subjects were run in a single session.

Procedure

The study was conducted by means of questionnaire. Subjects were informed that 'a preliminary survey had shown that students of the university campus saw a series of behaviours displayed by certain groups of people as being important and frequent', and were asked to rate 'these behaviours and people'. In the likeable–general norm condition, subjects were asked to judge 'Belgian (North African) students who always lend their lecture-notes to their fellow students'. In the unlikeable–general norm condition subjects were asked to judge 'Belgian (North African) students who never lend their lecture-notes to their fellow students'. In the likeable–exclusive norm condition subjects were asked to judge 'Belgian (North African) students who put studying behind amusement'. Finally, in the unlikeable–exclusive norm condition, subjects were asked to judge 'Belgian (North African) students who put amusement behind studying'[6].

[6] The norms were first selected on the basis of a pilot study aimed at determining the normative and counter-normative strength of several behaviours. Surprisingly, this pilot-study showed that, for our subjects, those 'who put studying behind amusement' were judged more favourably than those 'who put amusement behind study'. The choice of the general and exclusive norms was finally made by means of a group discussion among Belgian students.

Dependent measures

Judgments were obtained through five positive trait-descriptors (pleasant, sociable, welcoming, cheerful and communicative) issued from Experiment 1. This was done in order to insure that the selected traits were susceptible of being discriminated between the categories. Answers were given by means of a 7-point scale ranging from 1(=disagree) to 7(=agree).

Results and discussion

Evaluations were submitted to a 2 (ingroup versus outgroup) × 2 (likeable versus unlikeable) × 2 (exclusive norm versus general norm) ANOVA. No significant effects were found for target's group membership, $F(1, 83) = 0.52$, $p = 0.47$, for norms, $F(1, 83) = 1.09$, $p = 0.30$, and for the target's group membership × norms interaction, $F(1 83) = 0.17$, $p = 0.68$. The likeability × norms interaction showed moderate significance, $F(1, 83) = 3.01$, $p < 0.09$). Most important, significant effects were found for likeability, $F(1, 83) = 38.86$, $p < 0.001$, for the target's group membership × likeability interaction, $F(1, 83) = 4.76$, $p < 0.04$) and for the target's group membership × norms × likeability interaction, $F(1, 83) = 4.59$, $p < 0.04$. The latter interaction entirely supports our hypothesis. As can be noticed in Figure 2, virtually no difference exists between judgments of likeable and unlikeable ingroup and outgroup targets under the general norm ($M = 5.60$ and 5.49 respectively for likeable in- and outgroupers; $M = 3.04$ and 2.95 respectively for unlikeable in- and outgroupers). Thus the effect found through the likeability × target's group membership interaction seems to be primarily due to responses in the exclusive norm condition, where judgmental extremity emerged for ingroup but not for outgroup stimuli ($M = 5.53$ and 3.78 respectively for

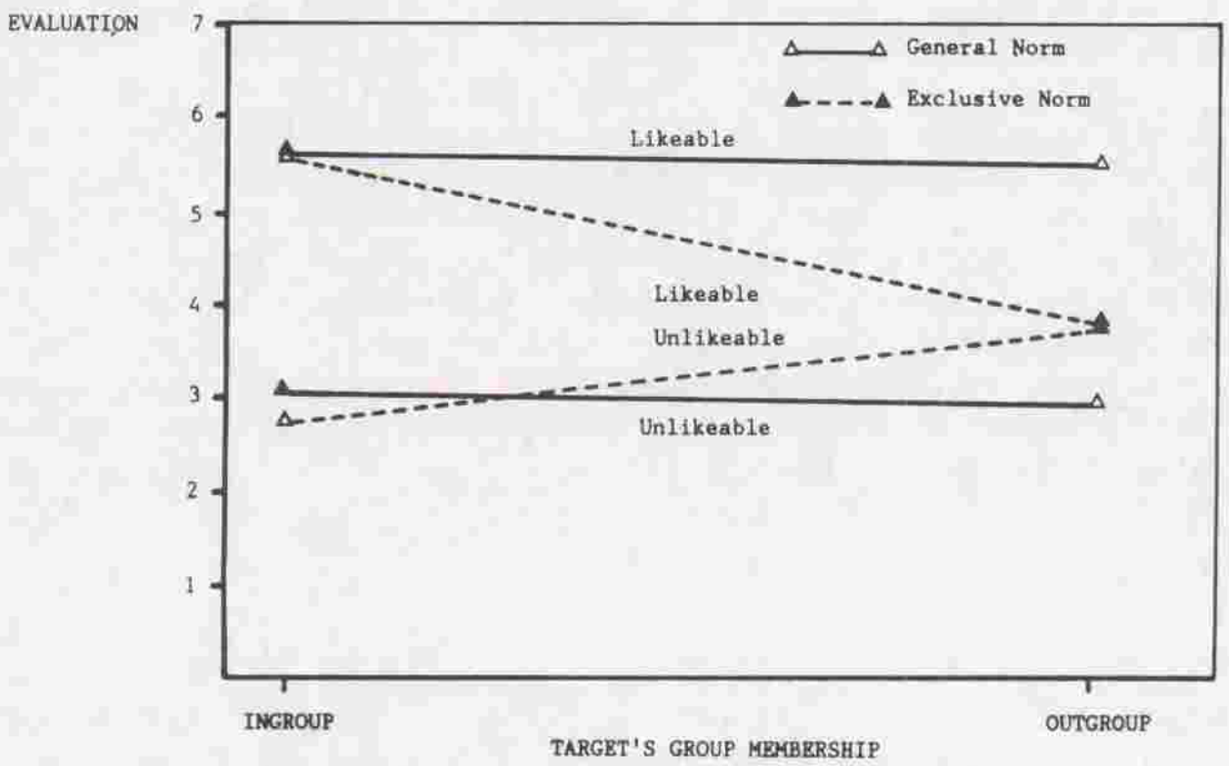

Figure 2. Judgmental extremity as a function of target's group membership, norms and likeability

likeable in- and outgroupers; M = 2.71 and 3.73 respectively for unlikeable in- and outgroupers).

Planned-comparisons were performed to see whether, for the exclusive norm, the ingroup and outgroup targets had been differently evaluated. Both the ingroup unlikeable and the ingroup likeable targets were juded more extremely than their outgroup correlatives ($F(1, 83)= 3.13$, $p < 0.05$, one-tailed, and, $F(1, 83)= 8.07$, $p < 0.01$, one-tailed, respectively). These tests stand for the predicted black sheep effect. In addition, the examination of average within-group inter-item correlations again fails in supporting the predictions of the complexity–extremity hypothesis. These correlations were higher in the ingroup condition (r = 0.71) than in the outgroup condition (r = 0.65). In brief, results again confirmed the black sheep hypothesis. A natural extension of these results would be the emergence of the black sheep effect for the general norm if an outgroup were contrasted to the general category of 'students'.

EXPERIMENT 3

The results obtained in Experiments 1 and 2 suggest that unlikeable ingroup members are more negatively evaluated than unlikeable outgroup members when their behaviours are relevant for the ingroup's social identity. Conversely, likeable ingroup members are judged more favourably then likeable outgroup members. This seems to be due to group identification rather than to schematic complexity. Experiment 3 is a check for this assumption. It was inspired by the events at the Heysel Stadium in Brussels, Belgium, in May 1985 when, following a riot between supporters of British and Italian soccer teams, 39 people were killed and many others were severely injured. The impact of these events on Belgian public opinion was exacerbated by its media coverage: the riot was broadcasted live, for, at least, four hours, on major European TV channels, including the Belgian ones.

We asked a group of Belgian students to imagine that German versus Belgian supporters had triggered those events, and to evaluate such fictive supporters on a series of trait-descriptors. Subjects were also asked to answer a series of questions aimed at determining their level of familiarity with the domain of soccer and with soccer team supporters. Only one prediction was made: Belgian (ingroup) supporters would be judged more negatively than German (outgroup) supporters, regardless of familiarity. Because we assume that, in this case, judgmental extremity is due to the target's relevance for the subjects' social identity, we expect that familiarity with soccer will be of little effect: both familiar and unfamiliar subjects would strongly identify with the Belgian category, therefore, strongly rejecting unlikeable Belgian supporters.

Method

Subjects

Two days after the Heysel events, 40 male Belgian undergraduates, aged 19 to 26 volunteered to answer to a questionnaire on 'soccer and violence'. Nineteen subjects were presented with the outgroup target while 21 were presented with the ingroup target.

Procedure

Subjects were presented with a set of 26 personality and attitude descriptors, issued from a content analysis of the Belgian French-speaking newspapers published the morning after the Heysel incident. Each item was judged on a 7-point scale ranging from 7(=agree) to 1(=disagree). Five questions were aimed at classifying subjects in low- versus high-familiarity groups: (1) 'To what extent are you interested in soccer?'; (2) 'To what extent would you consider yourself as a soccer team supporter?'; (3) 'Do you regularly participate as a player in soccer matches?' (4) 'How close did you follow the Heysel events, on the radio, in the press, on televsion, and so on?'; (5) 'How many soccer matches do you attend on the average per year?' Except for the last question, which asked for a number, the other questions were rated on 7-point scales.

Results and discussion

Familiarity with the stimulus-domain

Subjects were classified in the high- or the low-familiarity groups by means of a median-split on the sum of the five information items (Cronbach's alpha = 0.81). This step allowed us to construct four conditions: For the outgroup condition, nine subjects were classified as non-familiar and 10 as familiar; for the ingroup condition, 11 subjects were classified as non-familiar and 10 as familiar.

Evaluative judgments

Subjects' average scores on the 26 items were submitted to a two-way ANOVA for target's group membership (ingroup versus outgroup) and familiarity (high versus low). Results are presented in Table 2. As predicted, a significant main effect emerged for target's group membership, $F\,(1, 36) = 6.16$, $p < 0.02$, but no effect was found either for familiarity, $F\,(1, 36) < 1$, n.s. or for the target's group membership × familiarity interaction, $F\,(1, 36) < 1$, n.s.[7].

Table 2. Unlikeability judgments as a function of target's group membership and information

	Target	
Information	Outgroup	Ingroup
Low		
*M**	4.13	5.09
S.D.	0.66	0.69
High		
M	4.57	5.07
S.D.	1.23	1.01

*Ratings on a 7-point scale ranging from 7(=applies) to 1(=doesn't apply).

[7] It is important to note that testing the familiarity effect amounts to testing the null hypothesis. However, the alpha value of 0.47 for the information main effect, allows to be quite confident as to the validity of the conclusions being drawn.

Thus, results again support our hypothesis about the role of group membership, and do not verify the causal effect of familiarity and, indirectly, cognitive complexity on the extremity of ingroup judgments.

GENERAL DISCUSSION

The findings of the above experiments strongly support the black sheep hypothesis: Ingroup members were consistently evaluated in a more extreme way than outgroup members, either favourably or unfavourably.

The black sheep effect and the attitude polarization model

Experiment 2 also indicated that the black sheep effect arises only when the cues for evaluation are relevant for the subjects' social identity. This finding outlines the subjects' flexibility with regard to which aspect of their social identity is actually used in the judgmental situation (*cf.* Turner, 1981; Tajfel, 1982), a finding which is fully compatible with recent approaches to context-effects on cognitive categories (e.g. Murphy and Medin, 1985; Wright and Murphy, 1984). Probably, what happened in the general norm condition is that subjects perceived the targets' behaviour as conforming (likeable condition) or not conforming (unlikeable condition) to the normative standard of a larger scale membership category, namely the *student* group. Conversely, in the exclusive norm condition, where the target's behaviour was relevant for the ingroup positive definition, the black sheep pattern emerged quite clearly. Further research needs to be done, however, to examine whether the general norm would yield less extreme judgments with respect to an outgroup presenting the same level of generality as the student category.

Also, the fact that, in this experiment, there was no differential extremity for ingroup targets and outgroup targets when their behaviour concerned a general norm may appear contradictory with Tesser's model. Indeed, according to this model such difference should arise because of subjects' differential schematic complexity about the targets. More recent research reported by Millar and Tesser (1986) does not help solving this problem, even though one accepts that ingroup schemata are more complex than outgroup schemata and that subjects commit themselves more for ingroup judgments than for outgroup judgments. Millar and Tesser would still have to explain why did commitment arise for the outgroup in the general norm condition but not in the exclusive norm condition. Probably, these authors would have to admit that subjects shifted from one schematic level (Belgian or North African students) to a more abstract one (students in general). This assumption is extraneous to the attitude polarization model, but entirely accounted for by Social Identity Theory (Turner, 1987). This theory states that the individuals' social identities have a multitude of facets and that these facets may gain different weights in different situations. But even in this case, the authors would have to explain why should the general category of students be more complex than its Belgian or North African sub-groups. Indeed, as shown in the literature, schemata tend to become simpler as they increase in generality (e.g. Rosch, 1978; Smith and Medin, 1981; Weber and Crocker, 1983), a result that Millar and Tesser themselves reported with respect to 'group' versus 'individual' schemata. In

brief, our findings are more parsimoniously interpreted in terms of the black sheep model than in terms of the attitude polarization model.

Considered along with Experiments 1 and 2, results of Experiment 3 strongly suggest that ingroup favouritism can emerge in the form of an 'outgroup bias', possibly as a means to preserve the overall positive image of the ingroup, and that this phenomenon bears very little relationship, if at all, to the richness of information about the stimulus-domain under scrutiny. To summarize, the results allow to conceive of the black sheep effect as a consequence of the joint operation of cognitive and motivational factors, namely the ingroup and outgrup schemata, and the relevance of the norm to the ingroup.

The black sheep effect and the outgroup homogeneity hypothesis

A final comment should be addressed to schematic complexity. First, cognitive complexity has been defined as the number of orthogonal attribute dimensions which represent a category. Also, it has been accepted that in the intergroup domain, cognitive complexity increases with the amount of interpersonal contacts between ingroup and outgroup members, respectively. Consequently, cognitive complexity has been assumed to be higher for ingroup than for outgroup representations (e.g. Linville and Jones, 1980).

Second, we have provided our subjects with a single attribute dimension which was an evaluative one. At this point, it is important to dispel an apparent confusion in the literature on cognitive complexity. For some authors, complexity means the number of attribute dimensions employed by subjects in order to judge a stimulus on a given dimension. For others, it means the number of response dimensions subjects are provided with (*cf.* Judd and Lusk, 1984). We have followed the former version of the concept, a version which entirely meets Linville and Jones's (1980) definition.

Third, by theoretical implication, the emergence of ingroup favouritism — in which we were interested — stands for *one* evaluative dimension. As a result, cognitive complexity may, indeed, be higher for the ingroup than for the outgroup. However, Experiment 3 and, to a certain extent, Experiment 2 as well, stringently demonstrated that likely contact with the stimulus-domain was unrelated to judgmental extremity. In this case, outgroup homogeneity was clearly a function of the relevance of the judgmental targets for our subjects.

As a conclusion, it is interesting to note that our approach may shed new light upon the original assumptions of the outgroup homogeneity hypothesis. Instead of being uniquely due to the amount of cognitive complexity of ingroup and outgroup representations, or of cognitive factors such as the typicality of traits to be judged (e.g. Park and Rothbart, 1982), which, supposedly, are built from contact, outgroup homogeneity may be caused by the lack of relevance of outgroup targets or outgroup behaviour in light of the subject's social identity. Our results thus seem to indicate that the process-oriented (motivational–cognitive) Social Identity Theory can easily cope with the predictions of the structure-oriented (cognitive–schematic) outgroup homogeneity hypothesis[8], and that, for the moment, Social Identity Theory seems to

[8] Thanks are due to an anonymous reviewer for pointing out to us this important implication of our study in the first version of the present draft.

be the most parsimonious account for the phenomenon spotlighted by our studies. This is a further reason for considering the black sheep effect as a promising opportunity for new conceptual developments within the field of intergroup perception.

ACKNOWLEDGEMENTS

Very special thanks are due to Icek Ajzen, Robert Hogenraad, the Editor and three anonymous reviewers for thier insightful comments on previous versions of this paper.

REFERENCES

Allen, V. L. and Wilder, D. A. (1975). 'Categorization, belief similarity and intergroup discrimination', *Journal of Personality and Social Psychology*, **32**: 971–977.

Allport, G. W. (1954). *The Nature of Prejudice*, Addison-Wesley, Reading, Mass.

Ashmore, R. D. and DelBoca, F. K. (1981). 'Conceptual approaches to stereotypes and stereotyping'. In: Hamilton, D. L. (Ed.) *Cognitive Processes in Stereotyping and Intergroup Behavior*, Lawrence Erlbaum Associates, Hillsdale, NJ.

Billig, M. (1973). 'Normative communication in a minimal intergroup situation', *European Journal of Social Psychology*, **3**: 339–344.

Billig, M. and Tajfel, H. (1973). 'Social categorization and similarity in intergroup behaviour', *European Journal of Social Psychology*, **3**: 27–52.

Brewer, M. B. (1979). 'Ingroup bias in the minimal intergroup situation: a cognitive–motivational analysis', *Psychological Bulletin*, **86**: 307–324.

Brewer, M. B. and Kramer, R. (1985). 'The psychology of intergroup attitudes', *Annual Review of Psychology*, **36**: 219–243.

Fiske, S. T. and Taylor, S. E. (1984). *Social Cognition*, Addison-Wesley, Reading, Mass.

Hamilton, D. L. (1981). 'Stereotyping and intergroup behavior: some thoughts on the cognitive approach'. In: Hamilton, D. L. (Ed.) *Cognitive Processes in Stereotyping and Intergroup Behavior*, Lawrence Erlbaum Associates, Hillsdale, NJ.

Hastie, R. (1981). 'Principles in human memory'. In: Higgins, E. T., Herman, C. P. and Zanna, M. P. (Eds) *Social Cognition: The Ontario Symposium on Personality and Social Psychology*, Vol. 1, Lawrence Erlbaum Associates, Hillsdale, NJ.

Higgins, E. T., Kuiper, N. A. and Olson, J. M. (1981). 'Social cognition: a need to get personal'. In: Higgins, E. T., Herman, C. P. and Zanna, M. P. (Eds) *Social Cognition: The Ontario Symposium on Personality and Social Psychology*, Vol. 1, Lawrence Erlbaum Associates, Hillsdale, NJ.

Isen, A. M. and Daubman, K. A. (1984). 'The influence of affect on categorization', *Journal of Personality and Social Psychology*, **47**: 1206–1217.

Jones, E. E. and DeCharms, R. (1957). 'Changes in social perception as a function of the personal relevance of behavior', *Sociometry*, **20**: 175–185.

Judd, C. M. and Lusk, C. M. (1984). 'Knowledge structures and evaluative judgments: effects of structural variables on judgmental extremity', *Journal of Personality and Social Psychology*, **46**: 1193–1207.

Linville, P. W. (1982). 'The complexity–extremity effect and age-based stereotyping', *Journal of Personality and Social Psychology*, **42**: 193–211.

Linville, P. W. and Jones, E. E. (1980). 'Polarized appraisals of outgroup members', *Journal of Personality and Social Psychology*, **38**: 689–703.

Locksley, A., Borgida, E., Brekke, N. and Hepburn, C. (1980). 'Sex stereotypes and social judgments', *Journal of Personality and Social Psychology*, **39**: 821–831.

Mackie, D. and Cooper, J. (1984). 'Attitude polarization: effects of group membership', *Journal of Personality and Social Psychology*, **46**: 575–585.

Meindl, J. R. and Lerner, M. J. (1984). 'Exacerbation of extreme responses to an out-group', *Journal of Personality and Social Psychology*, **47**: 71–84.

Millar, M. G. and Tesser, A. (1986). 'Thought-induced attitude change: the effects of schema structure and commitment', *Journal of Personality and Social Psychology*, **51**: 259–269.

Miller, A. G. (1982). 'Historical and contemporary perspectives on sterotyping'. In: Miller, A. G. (Ed.) *In the Eye of the Beholder: Contemporary Issues in Stereotyping*, Praeger, New York.

Miller, N. and Brewer, M. B. (1984). *Groups in Contact: The Psychology of Desegregation*, Academic Press, New York.

Moreland, R. L. and Levine, J. M. (1982). 'Socialization in small groups: temporal changes in individual-group relations'. In: Berkowitz, L. (Ed.) *Advances in Experimental Social Psychology*, Vol. 15, Academic Press, New York.

Mummendey, A. and Schreiber, H.-J. (1984). '"Different" just means 'better': some obvious and some hidden pathways to ingroup favouritism', *British Journal of Social Psychology*, **23**: 363–368.

Murphy, G. L. and Medin D. L. (1985). 'The role of theories in conceptual coherence', *Psychological Review*, **92**: 289–316.

Nisbett, R., Zukier, H. and Lemley, R. (1981). 'The dilution effect: non-diagnostic information weakens the implications of diagnostic information', *Cognitive Psychology*, **13**: 248–297.

Park, B. and Rothbart, M. (1982). 'Perception of outgroup homogeneity and levels of social categorization: memory for the subordinate attributes of ingroup and outgroup members', *Journal of Personality and Social Psychology*, **42**: 1051–1068.

Rijsman, J. (1982). 'The dynamics of social competition in personal and categorical comparison-situations'. In: Doise, W. and Moscovici, S. (Eds) *Current Issues in European Social Psychology*, Vol. 1, Cambridge University Press, Cambridge.

Rosch, E. (1978). 'Principles of categorization'. In: Rosch, E. and Lloyd, B. B. (Eds) *Cognition and Categorization*, Wiley, New York.

Rose, T. L. (1981). 'Cognitive and dyadic processes in intergroup contact'. In: Hamilton, D. L. (Ed.) *Cognitive Processes in Stereotyping and Intergroup Behavior*, Lawrence Erlbaum Associates, Hillsdale, NJ.

Rothbart, M. (1981). 'Memory processes and social beliefs'. In: Hamilton, D. L. (Ed.) *Cognitive Processes in Stereotyping and Intergroup Behavior*, Lawrence Erlbaum Associates, Hillsdale, NJ.

Schachter, S. (1951). 'Deviation, rejection, and communication', *Journal of Abnormal and Social Psychology*, **46**: 190–217.

Sherif, M. and Sherif, C. W. (1979). 'Research on intergroup relations'. In: Austin, W. G. and Worchel, S. (Eds) *The Social Psychology of Intergroup Relations*, Brooks/Cole, Monterey, Ca.

Smith, E. E. and Medin, D. L. (1981). *Categories and Concepts*, Harvard University Press, Cambridge, Mass.

St Claire, L. and Turner, J. C. (1982). 'The role of demand characteristics in the social categorization paradigm', *European Journal of Social Psychology*, **12**: 307–314.

Stephan, W. G. (1985). 'Intergroup relations'. In: Lindzey, G. and Aronson, E. (Eds) *The Handbook of Social Psychology*, Vol. 2, Random House, New York.

Stephan, W. G. and Rosenfield, D. (1982). 'Racial and ethnic stereotypes'. In: Miller, A. G. (Ed.) *In the Eye of the Beholder: Contemporary Issues in Stereotyping*, Praeger, New York.

Tajfel, H. (1978). *Differentiation Between Social Groups*, Academic Press, London.

Tajfel, H. (1982). 'Social psychology of intergroup relations', *Annual Review of Psychology*, **33**: 1–39.

Tajfel, H. and Billig, M. (1974). 'Familiarity and categorization in intergroup behaviour', *European Journal of Social Psychology*, **10**: 159–170.

Tajfel, H., Billig, M., Bundy, R. P. and Flament, C. (1971). 'Social categorization and intergroup behaviour', *European Journal of Social Psychology*, **1**: 149–178.

Tajfel, H. and Turner, J. C. (1979). 'An integrative theory of intergroup conflict'. In: Austin, W. G. and Worchel, S. (Eds) *The Social Psychology of Intergroup Relations*, Brooks/Cole, Monterey, Ca.

Taylor, S. E. (1981). 'A categorization approach to stereotyping'. In: Hamilton, D. L. (Ed.) *Cognitive Processes in Stereotyping and Intergroup Behavior*, Lawrence Erlbaum Associates, Hillsdale, NJ.

Taylor, S. E. and Crocker, J. C. (1981). 'Schematic bases of social information processing'. In: Higgins, E. T., Herman, C. P. and Zanna, M. P. (Eds) *Social Cognition: The Ontario Symposium on Personality and Social Psychology*, Vol. 1, Lawrence Erlbaum Associates, Hillsdale, NJ.

Tesser, A. (1978). 'Self-generated attitude change'. In: Berkowitz, L. (Ed.) *Advances in Experimental Social Psychology*, Vol. 11, Academic Press, New York.

Tesser, A. and Leone, C. (1977). 'Cognitive schemas and thought as determinants of attitude change', *Journal of Experimental Social Psychology*, **13**: 340–356.

Turner, J. C. (1981). 'The experimental social psychology of intergroup behaviour'. In: Turner, J. C. and Giles, H. (Eds) *Intergroup Behaviour*, Blackwell, Oxford.

Turner, J. C. (1987). *Rediscovering the social group: A self-categorization Theory*. Blackwell, Oxford.

Turner, J. C., Brown, R. J. and Tajfel, H. (1979). 'Social comparison and group interest in ingroup favouritism', *European Journal of Social Psychology*, **9**: 187–204.

Weber, R. and Crocker, J. C. (1983). 'Cognitive processes in the revision of stereotypic beliefs', *Journal of Personality and Social Psychology*, **45**: 961–967.

Wright, J. and Murphy, G. L. (1984). 'The utility of theories in intuitive statistics: the robustness of theory-based judgments', *Journal of Experimental Psychology: General*, **113**: 301–322.

ZUSAMMENFASSUNG

In der vorliegenden Untersuchung wird eine Erweiterung des Phänomens der ingroup Favorisierung vorgeschlagen, welche auf der Hypothese aufbaut, daß Mitglieder der eigenen Gruppe positiver oder negativer beurteilt werden können als ähnliche Mitglieder einer anderen Gruppe. Dies steht im Gegensatz zu Vorhersagen, die aus der Komplexitäts-Extremitäts-Hypothese (Linville 1982, Linville und Jones 1980), aus der Hypothese der ingroup Favorisierung (Tajfel 1982) und aus Tessers Modell der Einstellungspolarisierung (1978, Millar und Tessar 1986) folgen. Unsere hauptsächliche Vorhersage, die sich auf die Theorie der sozialen Identität gründet, sagt aus, daß die Urteile sowohl über sympathische als auch über unsympathische ingroup Mitglieder extremer sind als über outgroup Mitglieder. Dieses Phänomen, als "Black-Sheep-Effect" bezeichnet, wird als durch die Relevanz hervorgebracht angesehen, welche das Verhalten der ingroup Mitglieder im Vergleich mit dem der outgroup Mitglieder für die soziale Identität des Urteilers hat. Drei Experimente bestätigten unsere Vorhersage. Experiment 1 zeigte zusätzlich, daß Korrelationen zwischen Attributen für die eigene Gruppe höher als für die andere Gruppe sind. Experiment 2 zeigte, daß der Black Sheep Effekt nur dann auftritt, wenn die Grundlage des Urteils für die soziale Identität des Urteilers relevant ist. Experiment 3 wies nach, daß das Niveau der Information über die zu beurteilende Person keinen Einfluß auf die Urteilsextremität hat. Die Ergebnisse werden im Lichte einer kognitiv-motivationalen Erklärung, die eine Alternative zu der rein kognitiven Interpretation der outgroup Homogenität darstellt, diskutiert.

RÉSUMÉ

Cette étude propose un élargissement du phénomène de favoritisme à l'égard de son propre groupe; elle repose sur l'hypothèse que des jugements à l'égard des membres du groupe peuvent être plus tranchés, de façon positive ou négative, que des jugements vis-à-vis de membres similaires appartenant à d'autres groupes. L'étude confronte les prédictions issues d'une part de l'hypothèse complexité-extrêmité (Linville, 1982; Linville & Jones, 1980), d'autre part de l'hypothèse du favoritisme à l'égard du groupe propre (Tajfel, 1982) et du modèle de la polarisation d'attitude de Tesser (1978; Millar & Tesser, 1986). Notre attente principale, basée sur la théorie de l'identité sociale, consiste en ce que les jugements à l'égard des membres de l'in-groupe (qu'ils soient agréables ou désagréables) sont plus contrastés que ceux qui concernent des membres du hors-groupe. Ce phénomène, dénommé l'"Effet black sheep", serait dû à l'importance des comportements des membres du groupe propre, comparé aux membres des hors-groupes du point de vue de l'identité sociale des sujets. Trois expériences confirment nos prédictions. L'expérience l en outre indique des corrélations inter-traits plus fortes dans le groupe même que vis-à-vis de membres d'autres groupes. L'expérience 2 montre que l'"effect black sheep" se manifeste seulement lorsque les indices de jugement sont pertinents par rapport à l'identité sociale des sujets, et l'expérience 3 révèle que les niveaux d'information relatifs à la cible du jugement se sont révélés inefficaces lorsque des jugements extrêmes étaient émis. On discute des résultats à la lumière d'une explication cognitivo-motivationnelle, alternative à l'interprétation purement cognitive de l'homogénéité du hors-groupe.

Copyright of European Journal of Social Psychology is the property of John Wiley & Sons Inc. and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.