

Attributional Mediation of In-Group Bias

Celina M. Chatman

University of Michigan

and

William von Hippel

Ohio State University

Received August 30, 1999; revised August 8, 2000; accepted August 10, 2000; published online April 3, 2001

The goal of this research was to examine the mediational role of attributional processes in in-group biases among Blacks and Whites. We conducted an experiment in which participants were asked to evaluate a Black or White job applicant and to make attributions for the fact that the applicant was fired or laid off from a previous job. Both Blacks and Whites demonstrated in-group-favoring biases in attributions and evaluations. Importantly, attributional biases mediated in-group-favoring evaluations of the applicant. These results highlight the importance of focusing on mechanisms other than out-group prejudice and hostility to explain people's consistent preferential treatment of in-group members. © 2001 Academic Press

In-group biases are ubiquitous. Even when people are arbitrarily divided into generic groups based on trivial criteria such as painting preferences, they tend to favor their own group (see Hogg & Abrams, 1993). When groups are based on naturally occurring social category memberships, such as race, ethnicity, and sex, a variety of implicit and explicit biases emerge that favor in-groups over out-groups (see Hilton & von Hippel, 1996). One mechanism through which this favoritism manifests itself is attributional bias. By attributing positive in-group behaviors and negative out-group behaviors to dispositional causes, but negative in-group and positive out-group behaviors to situational causes, people can maintain their in-group favoritism even in the face of inconsistent information. Pettigrew (1979) labeled this pattern of attributional bias the "ultimate attribution error" and suggested that it plays an important role in

the maintenance of stereotyping and prejudice (for a review, see Hewstone, 1990). Here, we examine the possibility that attributional biases play a mediational role between group membership and evaluative biases favoring the in-group. Specifically, we propose that attributional biases lead to different evaluations of the same behavior by in-group and out-group targets, thereby enabling people to express in-group favoritism.

Although there is ample evidence that in-group-favoring attributional biases occur (e.g., Hewstone & Ward, 1985; Taylor & Jaggi, 1974; Yarkin, Town, & Wallston, 1982), few studies have examined the mediating role of such biases in evaluations of members of in-groups and out-groups. To our knowledge, only one study (Jackson, Sullivan, & Hodge, 1993) has directly examined the links among group membership, attributions for a target's behavior, and evaluations of that target, and this study did not test for mediation directly. Consistent with the ultimate attribution error, Jackson et al. (1993) found that white participants made more dispositional than situational attributions for a high-achieving White and a low-achieving Black target, and more situational than dispositional attributions for a low-achieving White and a high-achieving Black target. No evidence emerged in this study, however, for in-group biases in overall evaluations of the target. Thus, although

Preparation of this manuscript was facilitated by funding from the Committee on Institutional Cooperation's (CIC) Summer Research Opportunities Program (SROP) at the Ohio State University. Some of the findings reported here were presented at the Annual Meeting of the Midwestern Psychological Association in Chicago, Illinois, in March, 1993. We thank Lee Jussim for helpful comments on earlier versions of this article.

Address correspondence and reprint requests to Celina Chatman, Achievement Research Program, Institute for Research on Women and Gender, University of Michigan, 204 South State Street, Ann Arbor, MI 48109. E-mail: cchatman@umich.edu.

participants' attributions for the target's level of achievement mediated their evaluations, there was no possibility for the attributions to mediate in-group biases. Moreover, although the attributional biases observed did favor the in-group, it is unclear whether the same targets would have led Black participants to show related patterns of attributions and evaluations.

Thus, there is some evidence that attributions mediate in-group biases, but there is no direct support for this hypothesis. Accordingly, the goal of the present research was to examine attributional biases among Blacks and Whites and to see if those biases mediate in-group-favoring evaluations. To accomplish this goal, we asked Black and White participants to evaluate a Black or White job applicant who was either fired or laid off from a previous position. We also assessed the extent to which participants made dispositional or situational attributions for the applicant's job problems.

Given the ubiquity of in-group biases, we predicted that Blacks and Whites would demonstrate in-group-favoring biases in their attributions for the target's job termination and in their evaluations of that target. We also predicted that greater intergroup differences would emerge in the evaluations of and attributions for a target who was fired rather than laid off. Because being fired is usually more likely to occur as a result of an individual's work-related transgressions or deficiencies, it is viewed more negatively than is being laid off. Thus, there is greater latitude for attributional biases to improve the interpretation of being fired. Finally, and most importantly, we predicted that in-group biases in evaluations would be mediated by in-group-favoring biases in the situational versus dispositional nature of the attributions.

METHOD

Participants

Participants were 80 Black and 80 White students (40 males and 40 females of each race) who were approached in the Ohio State University Student Union and asked to participate in the experiment.

Materials

Employment applications were created to represent a female applicant. All applications included identical information concerning the applicant's name, birthdate, current address and telephone number, previous work experience, citizenship, and current source of income. In addition, the application requested a statement from the applicant explaining her reason for leaving her last job. In response to this question, the applicant indicated that she had been either laid off or fired for absenteeism. A photograph of either a Black or White woman was attached to each appli-

cation. The race of each applicant was orthogonal to her termination status.

Procedure

Two female undergraduates, one Black and one White, served as experimenters for this research. Prospective participants were approached in the Student Union and asked to participate in the study by a same-race experimenter. The experimenter explained that she was a student conducting a project to examine the decision-making criteria used by people in personnel management. Participants were randomly assigned to read an application depicting either a White or Black applicant who lost her last job by being fired or laid off. The experimenter then read a series of questions to the participant. The first question asked why the applicant was absent so often and consequently fired or why she was laid off from her job. The experimenter then asked participants if they believed the reason that they had just provided was more due to external circumstances not within the applicant's personal control or more due to her personality and/or circumstances within her control. Responses to this question were given on a bipolar scale from 1 to 10, with "1" indicating a dispositional attribution and "10" indicating a situational attribution. The experimenter then asked participants how hardworking they believed the applicant would be and how likely they would be to hire her. Responses to these questions were made on 10-point rating scales with endpoints labeled *not at all* and *very much*. After completing the interview, participants were debriefed and thanked for their participation in the study. Although some participants inquired as to whether the application was really authentic, none indicated any suspicions relevant to our specific hypotheses.

RESULTS

Preliminary analyses revealed that participants' beliefs about how hardworking the applicant was and their likelihood of hiring her were strongly correlated ($r = .63, p < .001$). Accordingly, these two variables were collapsed into a single index that represented participants' evaluations of the target. Preliminary analyses also failed to reveal any effects for sex of the participant on the dependent measures, so this factor was dropped from the analyses. Finally, although a 2 (participant race) \times 2 (applicant race) \times 2 (termination status) ANOVA revealed an unexpected main effect for race, with White participants evaluating the applicants more favorably ($M = 5.69$) than Black participants ($M = 4.94$), $F(1, 152) = 7.38, p < .01$ ($\eta^2 = .05$), there were no significant differences in the in-group biases exhibited by Blacks and Whites. The participant race by applicant race interaction was thus collapsed into an in-group vs out-group main effect, facilitating the mediational analyses.

Manipulation Checks

A 2 (in-group vs out-group) \times 2 (fired vs laid off) analysis of variance (ANOVA) for evaluations of the target revealed a main effect for termination status, indicating that participants evaluated the applicant who was laid off more favorably ($M = 6.07$) than the applicant who was fired ($M = 4.56$), $F(1, 152) = 30.77$, $p < .001$ ($\eta^2 = .17$). Moreover, a one-sample t test indicated that the mean for the applicant who was fired was significantly lower than the midpoint (5.5) of the rating scale, $t(79) = -4.23$, $p < .001$, whereas the mean for the applicant who was laid off was significantly higher than the scale's midpoint, $t(79) = 3.20$, $p < .01$.

A main effect also emerged for participants' attributions for the applicant's job termination, with participants providing more situational attributions for the applicant who was laid off ($M = 6.91$) than the applicant who was fired ($M = 4.89$), $F(1, 152) = 31.41$, $p < .001$ ($\eta^2 = .17$). t Tests indicated that the mean for the applicant who was fired was significantly lower than the midpoint of the rating scale, $t(79) = -2.19$, $p < .05$ (suggesting that participants provided dispositional attributions for being fired), and the mean for the applicant who was laid off was significantly higher than the scale's midpoint, $t(79) = 5.92$, $p < .001$ (suggesting that participants provided situational attributions for being laid off). These results support the use of job termination status as a manipulation of negative behavior, thereby facilitating a test of our predictions regarding the greater impact of in-group biases under such conditions.

In-Group Bias in Attributions and Evaluations

We predicted that in-group biases would emerge in participants' attributions for the applicant being terminated from a previous job and in their overall evaluations of the applicant. In support of these hypotheses, an ANOVA revealed a main effect of in-group status for both attributions, $F(1, 152) = 6.91$, $p < .01$ ($\eta^2 = .04$) and evaluations, $F(1, 152) = 5.75$, $p < .02$ ($\eta^2 = .04$). Participants made more situational attributions for the in-group applicant's termination ($M = 6.38$) than for the out-group applicant's termination ($M = 5.43$) and evaluated the in-group applicant more positively ($M = 5.64$) than the out-group applicant ($M = 4.99$). t Tests revealed that attributions for in-group applicants were significantly different from the rating scale's midpoint, $t(79) = 2.99$, $p < .01$, as were evaluations of out-group members, $t(79) = -2.16$, $p < .05$; but attributions for out-group applicants and evaluations of the in-group member did not differ significantly from the rating scale's midpoint ($ts < |1.0|$, *ns*). Attributions and evaluations were correlated, $r = .47$, $p < .001$.

We also predicted that in-group biases would be enhanced for the applicant who was fired because that behav-

TABLE 1
Participants' Evaluations of the Applicant as a Function of the Applicant's Group Membership and Termination Status

Applicant's group membership	Applicant's termination status	
	Fired	Laid off
In-group	5.14 _a	6.15 _c
Out-group	3.98 _b	6.00 _c

Note. $N = 160$. Higher numbers correspond to more favorable evaluations. Values with different subscripts are significantly different at $p < .05$.

ior provides greater latitude for the impact of such biases. For participants' attributions, however, the in-group status by termination status interaction did not reach significance, $F(1, 152) = 1.38$, $p > .10$ ($\eta^2 = .01$). The two-way interaction for participants' evaluations of the applicant, however, did reach marginal significance, $F(1, 152) = 3.42$, $p = .07$ ($\eta^2 = .02$). As can be seen in Table 1, evaluations of an in-group member who was fired were more favorable than those of an out-group member who was fired, but no differences emerged in the evaluations of in-group and out-group members who were laid off.

Attributional Mediation of Bias

In order to assess whether participants' attributions for the applicant's job termination mediated their evaluations of the applicant, we conducted a series of regression analyses. According to Baron and Kenny (1986), the following three conditions must occur for mediation to be evident: (1) The independent variable must significantly affect the mediator, (2) the mediator must significantly affect the dependent variable, and (3) the significant relation between the independent variable and the dependent variable must disappear or be significantly weakened when the mediator is included in the equation. Because the ANOVA revealed a main effect of in-group status on attributions, the first criterion was met. In order to assess the second and third criteria, we compared two regression models—one including the mediator (attributions) as a predictor variable and one not including it—with in-group status, participant race, and termination status serving as predictors and evaluations as the outcome (see Table 2).

As can be seen in Table 2, the in-group attributional bias appears to have mediated in-group biases in evaluations. When attributions are not included in the model (Model 1), in-group status has an effect on evaluations. But when attributions are included in the model (Model 2), in-group status no longer predicts evaluations. Moreover, consistent with the ANOVA results, in-group status predicted the attributional bias ($\beta = .26$, $p = .01$), and a t test revealed that the pathway from in-group status, through attributions,

TABLE 2
Regression Analyses Testing for Evaluative Bias and Attributional Mediation

Variable	Model 1 (mediator not in)		Model 2 (with mediator in)		R^2	F
	B	t	B	t		
Termination status	1.52	5.48***	.99	3.45***	—	—
Participant race	.74	2.68**	.70	2.67**	—	—
In-group status ^a	.66	2.37*	.41	1.53	.22	14.28***
Attributions	—	—	.26	4.54***	.31	17.21***

Note. $N = 160$. Change in R^2 for Model 2 = .09, $F_{\text{change}}(1, 155) = 20.62^{***}$.

^a In-group = 1; Out-group = 0.

* $p < .05$.

** $p < .01$.

*** $p < .001$.

to evaluations was significant ($p < .05$). The regression analyses also revealed that attributions partially mediated the effect of termination status on evaluations because the pathway from termination status through attributions to evaluations was also significant ($p < .05$).

DISCUSSION

The findings reported here provide empirical support for the mediating role of attributional biases in evaluations of in-group and out-group members. Participants gave more favorable evaluations to in-group than out-group members and made more situational attributions for an in-group member's termination from a previous job than that of an out-group member. In addition, the degree to which people showed an in-group bias in their evaluations was mediated by the degree to which they showed an in-group bias in their attributions for the target's job termination. The implication of these findings is that because people generate different explanations for the same behavior when it is exhibited by an in-group rather than an out-group member, biased evaluations are not only likely to follow, but may also seem justified to the perceiver. That is, if someone was fired (or laid off) for situational reasons, then it is perfectly appropriate to give her a second chance at another job. But if someone was released from a job for dispositional reasons, then it is inappropriate to give her a second chance.

The research on in-group bias is marked by a lack of consensus regarding its nature. For example, there has been continual debate as to whether in-group biases are due more to enhancement of the in-group or derogation of the out-group. The results of our research support the former position. For example, mean attributions for the in-group applicant's termination, but not for the out-group applicant, were significantly different from the midpoint of the situational versus dispositional scale. This suggests that biases favoring the in-group, rather than against the out-group, were respon-

sible for the observed differences in attributions for the in-group and out-group applicants. Thus, as Brewer (1999) has noted, inequities are likely to emerge in evaluation and treatment of group members, even in the absence of any prejudicial attitudes held toward out-groups.

Given that positive differentiation from out-groups is contingent on appraisals of in-group members, negatively characterized in-group members may pose a threat to the maintenance of a positive social identity (Brewer, 1979; Tajfel, 1978). Marques (1990; Marques & Yzerbyt, 1988; Marques, Yzerbyt, & Leyens, 1988) has postulated that, in order to protect social identity in such instances, individuals judge negative in-group members as extreme cases. In other words, negative in-group members are seen as outliers who are atypical from the in-group as a whole. Consequently, negative in-group members are often evaluated even less favorably than similar out-group members. Such a pattern, which is referred to as the "black sheep effect," would yield outcomes in direct contrast to the findings in our research.

The black sheep effect, however, occurs only under conditions in which the evaluative dimension is a defining feature of in-group norms (Marques, 1990; Marques, Robalo, & Rocha, 1992; Marques et al., 1988; Marques & Yzerbyt, 1988) and thus relevant to social identity. This was not the case in our research. Given the ubiquity of research on racial stereotypes (e.g., Devine, 1989; Devine & Elliot, 1995; Katz & Braly, 1933) and, to a much lesser degree, on autostereotypes (e.g., Clark, 1985; Ryan, 1996), we can confidently declare that being fired as opposed to laid off from a job is benign to in-group/out-group differentiations for Blacks and Whites. Further, research showing evidence for the black sheep effect (Marques, 1990; Marques et al., 1992; Marques et al., 1988; Marques & Yzerbyt, 1988) has tended to emphasize likable and unlikable personality traits for which, by definition, situational attributions cannot be made. This is an important point in reconciling the contrast between our findings and those providing evidence for the

black sheep effect, as we argue that attributions are in part responsible for observed biases in evaluations. Thus, in the absence of attributional biases, evaluative biases are less likely to occur.

We also recognize yet another potential explanation for our findings. Given that we used only female targets in our research, gender biases may have played a role in the observed results. Clearly, gender as a category is just as accessible as race or ethnicity (Abrams, Thomas, & Hogg, 1990; Brewer, 1988; McCall & Simmons, 1966; McGuire, McGuire, & Winton, 1979), particularly in the context of workplace inequities (Trentham & Larwood, 1998). We found no differences, however, in evaluations or attributions as a function of participants' gender. This lack of an effect should not be given too much weight, however, because we used only female experimenters. Consequently, male participants were always in the out-group relative to the experimenter, and so the research design did not allow for a full test of gender-based biases. The goal of our study, of course, was to examine racial biases in attributions and evaluations.

Limitations and Directions for Future Research

Although the findings were consistent with our predictions, there are several limitations to our research. First, our conclusions are based on the findings from a single study, and so different findings might emerge with other groups or different scenarios. Likewise, we used only female job applicants, so it is possible that different patterns of bias (or no bias at all) might emerge for male applicants. In support of the generalizability of this research, however, it should be noted that our findings are consistent with previous research (Hewstone, 1990; Pettigrew, 1979). For example, the Jackson et al. (1993) studies revealed a similar pattern of in-group-favoring biases in White participants' attributions concerning in-group and out-group achievement.

Another potential limitation of our research is the use of a single bipolar scale to assess internal and external attributions. It has been argued (and indeed demonstrated) that the two are not necessarily opposing extremes of a single continuum (Solomon, 1978). For example, people could generate explanations for being laid off from a job that include both external (e.g., declining economy) and internal (e.g., less productivity relative to coworkers) components. We allowed for this possibility in our research by first asking participants to explain the applicant's termination. Thus, we did not preclude the generation of bidimensional attributions, which presumably would have led to a rating near the midpoint of the dispositional/situational scale (see Cacioppo & Berntson, 1994). Because attributions for being fired vs laid off were significantly different from that midpoint, participants seem to have given more weight to one type of explanation than the other. Perhaps they considered both situational and dispositional causes, but then indicated

whether the target's behavior was *better* explained by dispositional vs situational factors.

Likewise, there may be alternative explanations for why attributions and evaluations were so strongly related. The most likely contenders involve the methodology of our research. First, having participants (and not more objective judges) rate their own explanations may have introduced subjective biases in attributions. We presume, however, that managers faced with decisions about job applicants in real-world situations are likely to take into consideration whether an applicant's behavior was due more to personal or situational factors. Thus, it seemed more realistic to have participants make such decisions themselves in our research. Asking participants to evaluate the applicant just after making attributions for her behavior also may have led them to base their evaluations on their attributions. However, we conducted a preliminary study with similar materials, participants, and procedures in which attributions were not assessed (Chatman & von Hippel, 1993). This study revealed in-group biases similar to those reported here, suggesting that the attributional measure did not influence the evaluations.

In conclusion, we have shown that Blacks and Whites are subject to in-group biases in their attributions for the behavior of other Blacks and Whites and these attributions at least partially account for biased evaluations of in-group and out-group members. These findings contribute to an understanding of why people consistently favor members of the in-group over the out-group by providing an explanation other than out-group prejudice or hostility. Rather, our findings provide another piece of evidence for the argument that in-group enhancement, as opposed to out-group derogation, is responsible for the persistence of in-group-favoring biases (Brewer, 1979, 1999). At an applied level, these findings suggest that providing explicit explanations for the behavior of in-group and out-group members might reduce the potential for inequities to emerge.

REFERENCES

- Abrams, D., Thomas, J., & Hogg, M. A. (1990). Numerical distinctiveness, social identity and gender salience. *British Journal of Social Psychology*, *29*(1), 87-92.
- Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, *51*, 1173-1182.
- Brewer, M. B. (1979). In-group bias in the minimal intergroup situation: A cognitive-motivational analysis. *Psychological Bulletin*, *86*, 307-324.
- Brewer, M. B. (1988). A dual process model of impression formation. In R. S. Wyer, Jr., & T. K. Srull (Eds.), *Handbook of social cognition* (Vol. 1, pp. 1-36). Hillsdale, NJ: Erlbaum.
- Brewer, M. B. (1999). The psychology of prejudice: Ingroup love or outgroup hate? *Journal of Social Issues*, *55*, 429-444.
- Cacioppo, J. T., & Berntson, G. G. (1994). Relationship between attitudes and evaluative space: A critical review, with emphasis on the separa-

- bility of positive and negative substrates. *Psychological Bulletin*, **115**, 401–423.
- Chatman, C., & von Hippel, W. (1993). *Ingroup biases among Blacks and Whites: Evidence for the ultimate attribution error*. Paper presented at the Midwestern Psychological Association, Chicago, IL.
- Clark, M. L. (1985). Social stereotypes and self-concept in Black and White college students. *Journal of Social Psychology*, **125**(6), 753–760.
- Devine, P. G. (1989). Stereotypes and prejudice: Their automatic and controlled components. *Journal of Personality and Social Psychology*, **56**, 5–18.
- Devine, P. G., & Elliot, A. J. (1995). Are racial stereotypes really fading? The Princeton trilogy revisited. *Personality and Social Psychology Bulletin*, **21**(11), 1139–1150.
- Hewstone, M. (1990). The “ultimate attribution error”? A review of the literature on intergroup causal attribution. *European Journal of Social Psychology*, **20**, 311–335.
- Hilton, J. L., & von Hippel, W. (1996). Stereotypes. In J. T. Spence, J. M. Darley, & D. J. Foss (Eds.), *Annual review of psychology* (Vol. 47, pp. 237–271). Palo Alto: Annual Reviews.
- Hogg, M. A., & Abrams, D. (1993). *Group motivation: Social psychological perspectives*. London: Harvester Wheatsheaf.
- Jackson, L. A., Sullivan, L. A., & Hodge, C. N. (1993). Stereotype effects on attributions, predictions, and evaluations: No two social judgments are quite alike. *Journal of Personality and Social Psychology*, **65**, 69–84.
- Katz, D., & Braly, K. (1933). Racial stereotypes in one hundred college students. *Journal of Abnormal and Social Psychology*, **28**, 280–290.
- Marques, J. M. (1990). The black sheep effect: Outgroup homogeneity in social comparison settings. In D. Abrams & M. Hogg (Eds.), *Social identity theory: Constructive and critical advances*. London: Harvester Wheatsheaf.
- Marques, J. M., & Yzerbyt, V. Y. (1988). The black sheep effect: Judgmental extremity towards ingroup members in inter- and intra-group situations. *European Journal of Social Psychology*, **18**, 287–292.
- Marques, J. M., Yzerbyt, V. Y., & Leyens, J.-P. (1988). The black sheep effect: Judgmental extremity towards ingroup members as a function of group identification. *European Journal of Social Psychology*, **18**, 1–16.
- Marques, J. M., Robalo, E. M., & Rocha, S. A. (1992). Ingroup bias and the ‘black sheep’ effect: Assessing the impact of social identification and perceived variability on group judgments. *European Journal of Social Psychology*, **22**, 331–352.
- McCall, G. J., & Simmons, J. L. (1966). *Identities and interactions*. New York: Free Press.
- McGuire, W. J., McGuire, C. V., & Winton, W. (1979). Effects of household sex composition on the salience of one’s gender in the spontaneous self-concept. *Journal of Experimental Social Psychology*, **15**, 77–90.
- Pettigrew, T. F. (1979). The ultimate attribution error: Extending Allport’s cognitive analysis of prejudice. *Personality and Social Psychology Bulletin*, **5**, 461–476.
- Ryan, C. S. (1996). Accuracy of black and white college students’ in-group and out-group stereotypes. *Personality and Social Psychology Bulletin*, **22**(11), 1114–1127.
- Solomon, S. (1978). Measuring dispositional and situational attributions. *Personality and Social Psychology Bulletin*, **4**(4), 589–594.
- Tajfel, H. (1978). *Differentiation between social groups: Studies in the social psychology of intergroup relations*. London: Academic Press.
- Taylor, D. M., & Jaggi, V. (1974). Ethnocentrism and causal attribution in a South Indian context. *Journal of Cross-Cultural Psychology*, **5**, 162–172.
- Trentham, S., & Larwood, L. (1998). Gender discrimination and the workplace: An examination of rational bias theory. *Sex Roles*, **38**(1/2), 1–28.
- Yarkin, K. L., Town, J. P., & Wallston, B. S. (1982). Blacks and women must try harder: Stimulus person’s race and sex attributions of causality. *Personality and Social Psychology Bulletin*, **8**, 21–24.