

## Attribution Biases

**Attribution** refers to the way in which people explain their own behavior and that of others. An **attribution bias** occurs when someone systematically over- or underuses the available information when explaining behavior. There is evidence that when we are making judgments about the behavior of our own group (the ingroup) and that of other groups (outgroups), we show attributional **biases** that favor the ingroup. Specifically, where ingroup members are concerned, we explain positive behaviors in terms of internal characteristics (e.g., personality) and negative behaviors in terms of external factors (e.g., illness). Conversely, where outgroup members are concerned, we explain positive behaviors in terms of external characteristics and negative behaviors in terms of internal characteristics. The study of **attribution biases** is an essential aspect of group processes and intergroup relations because these **biases** can fuel negative relations between opposing groups. Understanding how and why **attribution biases** arise, however, facilitates the development of interventions to reduce them.

This entry outlines the basic theory, discusses how it applies in individual and group contexts, and describes research showing how **attribution bias** may be mitigated.

### Attribution Theory

Following the pioneering work of Fritz Heider, Harold Kelley developed a theory of causal **attribution** based on a scientific analysis of how people should explain, or attribute, their own or others' behavior by using the available information in a systematic manner. Heider and Kelley investigated the *locus of causality*, whether behavior is caused by something internal or external to the actor (the person performing the behavior). Later work, by Bernard Weiner, identified three further causal dimensions in terms of which **attributions** can be classified: *stability*, the extent to which causes are stable and permanent versus temporary and fluctuating; *controllability*, the extent to which causes can be influenced by the actor; and *globality*, whether a cause is global in nature or specific to a given situation.

Of most relevance to the issue of intergroup **attribution biases** is locus of causality. An *internal attribution* is any explanation that locates the cause as being internal to the person, such as personality, mood, abilities, attitudes, and effort. An *external attribution* is any explanation that locates the cause as being external to the person, such as the actions of others, the nature of the situation, social pressures, or luck. Thus, if people see a mother shouting at her child and decides that she is doing this because she is an aggressive person, they are making an internal **attribution**. In contrast, if they decide that she was reprimanding the child for behaving badly, they are making an external **attribution**.

### Individual Attribution Biases

Kelley's model is a rather idealized account of how people make causality judgments. Given that we normally have limited time and resources, we have a tendency to use heuristics, or shortcuts, when making social judgments, rather than taking into account all of the available information. As a result, researchers have observed a number of systematic **biases** that are made when people are assessing the causes of behavior.

There are three well-documented **attribution biases**. The *correspondence bias* refers to the fact that behavior

is often viewed as a reflection of an actor's corresponding internal disposition even when it was actually caused by situational factors. The *actor-observer bias* arises when we attribute other people's behavior to internal causes and our own behavior to external causes. Both of these effects can be explained by perceptual salience. The people being observed are the most salient aspect of the situation, as they are actually performing the action—they and their behavior appear to go together, so an internal **attribution** is made. In contrast, when making self-**attributions**, we are focused outward and the situation is salient, and thus we attribute causality for our behavior to external factors.

The *self-serving attribution bias* refers to our tendency to make internal **attributions** for our successes and external **attributions** for our failures. If students excel in an exam, for example, they are likely to think this is because they are very intelligent, but if they fail, they may attribute this to the poor quality of their teacher. In contrast to the *perceptual* processes underlying correspondence and actor-observer **biases**, the self-serving **attribution bias** has a *motivational* basis. We are motivated to view ourselves in a positive light, to have high self-esteem. Attributing success to internal causes boosts our feelings of self-worth, whereas attributing our failures to external causes protects us from feeling bad when we do not do well. Together, these processes enable us to maintain and enhance our self-esteem. Extending these findings, research has shown that as well as making **attributions** that favor the self, we are also motivated to make **attributions** that favor *groups* to which we belong over groups to which we do not.

## Intergroup Attribution Biases

*Intergroup attribution* refers to the ways in which members of different social groups explain the behavior of members of their own and other social groups. A person attributes the behavior of another person not simply to individual characteristics, but also to characteristics associated with the group to which the other person belongs. Moreover, the group membership of the perceiver, or attributor, can also affect the intergroup **attribution** process.

Social psychologists have investigated how we make **attributions** in an intergroup context. Hindus (a minority group) and Muslims (a majority group) in Bangladesh read scenarios about an individual from either their ethnoreligious group or the other group, and they were instructed to imagine that this person had behaved in either a positive or a negative way toward them (e.g., a passerby either helped or failed to help the participant when he or she had fallen off a bike). Among Muslim participants, positive behavior of a Muslim (an ingroup member) and negative behavior of a Hindu (an outgroup member) tended to be attributed to causes rated as internal, stable, uncontrollable by others, and global. In contrast, positive behavior of a Hindu and negative behavior of a Muslim were typically attributed to causes rated as external, unstable, controllable by others, and specific. Notably, Hindu participants showed considerably less intergroup **bias** in **attributions**, suggesting that these **biases** are stronger among majority groups than minority groups.

Research has also considered whether there are **biases** in **attributions** made for the historical actions of entire outgroups. (Non-German) Jewish and (non-Jewish) German participants were asked why they thought Germans mistreated Jewish people during the Second World War. Jewish participants were more likely to attribute the behavior of the Germans to internal characteristics such as German aggression than were German participants. In a further study, Dutch participants were asked to make internal or external **attributions** for behavior in two historical contexts: Dutch behavior toward Indonesians during the colonization period (negative ingroup behavior) and German behavior toward the Dutch during the Second World War (negative outgroup behavior). Participants were more likely to make internal **attributions** about negative outgroup behavior than negative ingroup behavior, and more likely to make external **attributions** about negative ingroup behavior than negative outgroup behavior.

Finally, there is evidence for *linguistic* intergroup **attribution biases**. People tend to use relatively abstract terms to describe the negative behavior of an outgroup member and the positive behavior of an ingroup

member, because this implies that the behavior is generalized to the personality of the actor. In contrast, people use relatively concrete terms to describe the negative behavior of an ingroup member and the positive behavior of an outgroup member because this implies that the behavior is specific to a particular context.

To summarize, in an intergroup context, we tend to make **attributions** regarding locus of causality that favor the ingroup over the outgroup. This is a form of self-serving **attribution bias**, but instead of enabling us to view ourselves in a positive light compared to other individuals, it enables us to view the groups to which we belong positively compared to other groups. Specifically, we tend to explain the positive behavior of ingroup members in terms of internal characteristics but the positive behavior of outgroup members in terms of external characteristics. In contrast, we tend to explain the negative behavior of ingroup members in terms of external characteristics, but the negative behavior of outgroup members in terms of internal characteristics. We also have also a tendency to make **biased** intergroup **attributions** based on linguistics, globality, stability, and controllability.

So why do we make these intergroup **attribution biases**? According to social identity theory, we tend to favor our own group over other groups to maintain a positive perception of the ingroup and therefore maintain a high level of self-esteem. We make intergroup **attribution biases** to ensure that our group is perceived in a positive light compared to other groups. Three findings support this social identity explanation. First, making group membership salient prior to completing an intergroup **attribution** task increases the extent to which participants show intergroup **attribution biases**. Second, intergroup **attribution biases** are stronger among participants who highly identify with their ingroup. Third, it has been demonstrated that making internal **attributions** about ingroup members and making global **attributions** about the negative behavior of outgroup members predicts higher self-esteem.

## **Reducing Intergroup Attribution Biases**

According to social identity theory, making our group membership salient increases intergroup **bias**, as we are motivated to maintain a positive perception of our own group relative to other groups. To reduce attributional **bias**, it is therefore necessary to change the nature of categorization. One way of doing this is *cross-categorization*, which involves crossing a dichotomous categorization with a second categorization. In the case of Hindus and Muslims in Bangladesh, for example, it is possible to introduce a second categorization, the distinction between Bangladeshi and Indian nationality. This cross-categorization creates four groups. For a Bangladeshi Muslim, the double ingroup refers to those who share both group memberships (other Bangladeshi Muslims), the partial ingroups are those who share one group membership (Bangladeshi Hindus and Indian Muslims), and the double outgroup refers to those who share neither group membership (Indian Hindus). People tend to favor double ingroup members and show the greatest discrimination against the double outgroup. Intergroup **bias** against partial ingroup members, however, is reduced compared to the double outgroup. Thus, seeing an outgroup member as being an ingroup member on a second dimension has benefits for intergroup relations. Research on intergroup **attribution biases** mirrors these findings. Bangladeshi Muslim study participants made the most positive **attributions** about a Bangladeshi Muslim protagonist and the most negative **attributions** about an Indian Hindu protagonist. **Attributions** made about Bangladeshi Hindus and Indian Muslims were, however, significantly more positive than those made about Indian Hindus.

In sum, intergroup attributional **biases** arise because of our motivation to maintain a positive social identity, and these **biases** contribute to the maintenance and exacerbation of conflict between groups. Research has shown, however, that changing our perceptions of intergroup categories through cross-categorization can lead to reductions in intergroup **attribution biases**. This research therefore makes an important contribution to our understanding of how intergroup relations can be improved.

—Rhiannon N. Turner

## Further Readings

Doosje, B. and Branscombe, N. R. **Attributions** for the negative historical actions of a group. *European Journal of Social Psychology*, vol. 33 (2003). pp. 235–248.

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

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

---

### Entry Citation:

Turner, Rhiannon N., and Miles Hewstone. "Attribution Biases." *Encyclopedia of Group Processes & Intergroup Relations*. Ed. John M. Levine and Michael A. Hogg. Thousand Oaks, CA: SAGE, 2009. 43-46. *SAGE Reference Online*. Web. 30 Jan. 2012.



© SAGE Publications, Inc.

Brought to you by: SAGE