Attribution as a Gateway to Social Cognition

Glenn D. Reeder

Abstract

Attribution is concerned with how we make sense of our world. We often wonder why people say and do the things they do. This chapter summarizes the main contributions of attribution research, particularly as it helped to shape the development of social cognition. Beginning with the person perception insights of Fritz Heider (1944, 1958), classic attribution theories were proposed (Bem, 1967; Jones & Davis, 1965; Kelley, 1967) and applied to almost every area of social psychology. Around 1980, cross-fertilization began between attribution and social cognition models of dispositional inference which incorporated automaticity (Gilbert, Pelham, & Krull, 1988; Trope, 1986). The chapter also covers biases in the attribution process such as naïve realism and recent advances in mindreading (which concerns inferences about mental states such as beliefs, intentions, and motives) as well as the application of neuroscience to attribution.

Key Words: attribution, social cognition, person perception, dispositional inference, automaticity, naïve realism, intention, mindreading, neuroscience

Attribution as a Gateway to Social Cognition

A new idea can be exciting. In 1971, for many graduate students in social psychology, the new ideas came fast and furious from the attribution perspective. Like underground music, a series of preprinted chapters on this topic were making their way into graduate seminars. The authors of these chapters included well-known figures such as Ned Jones and Harold Kelley, but also newcomers such as Bernard Weiner and Richard Nisbett. A year later, these chapters were bound together in an orange book entitled, Attribution: Perceiving the Causes of Behavior (Jones, Kanouse, Kelley, Nisbett, Valins, & Weiner (1972). As of March 2013, Google Scholar listed 2,786 citations of the Jones and Nisbett (1972) chapter, which proposed actor versus observer differences in attribution. Two additional chapters by Kelley (1972a, 1972b) and another by Weiner, Frieze, Kukla, Reed, Rest, and Rosenbaum (1972) were each cited more than 1100 times. What can account for the immediate splash and undeniably lasting impact of this work (Weiner, 2008)?

Attribution is concerned with how we make sense of our world (Kelley, 1967). It addresses questions of subjective validity: How do we know what is true and what is not? What causes people to do the things they do? These questions were implicit in much of what social psychologists did before the 1960s, but dimly realized. Asch’s (1946) studies of impression formation provided participants with lists of traits to be integrated, Attribution: Perceiving the Causes of Behavior (Jones, Kanouse, Kelley, Nisbett, Valins, & Weiner (1972). As of March 2013, Google Scholar listed 2,786 citations of the Jones and Nisbett (1972) chapter, which proposed actor versus observer differences in attribution. Two additional chapters by Kelley (1972a, 1972b) and another by Weiner, Frieze, Kukla, Reed, Rest, and Rosenbaum (1972) were each cited more than 1100 times. What can account for the immediate splash and undeniably lasting impact of this work (Weiner, 2008)?
Attribution as a Gateway to Social Cognition

models of dispositional inference (Gilbert, Pelham, & Krull, 1988; Trope, 1986). These models linked the information-processing approach of cognitive psychology (Neisser, 1967) and the nascent literature on automaticity (e.g., Bargh & Pietromonaco, 1982; Schneider & Shiffrin, 1977) with more traditional attribution concepts such as causal discounting (Kelley, 1972a). The final section of the chapter focuses on attribution of mind, which concerns inferences about mental states such as beliefs, goals, and motives (Epley & Waytz, 2010; Malle & Hodges, 2005). Recent contributions of neuroscience to attribution are also noted (Lieberman, 2010; Mason & Morris, 2010). The developments in these areas once again illustrate the interconnections between attribution and social cognition as they co-evolve.

This chapter summarizes some of the main contributions of attribution research, particularly as it helped to shape the development of social cognition. As shown in Figure 6.1, the organization is chronological to some extent. The first section describes the inspiration and pioneering ideas—primarily from Fritz Heider (1944, 1958)—that laid the groundwork for what I will call “classic attribution” research. The classic attribution period began around 1960 and peaked before 1980. As described in this second section, the early insights of attribution led to landmark empirical advances (e.g., Bem, 1967; Jones & Harris, 1967), systematic theories of person perception (e.g., Jones & Davis, 1965; Kelley, 1967), and applications to almost every area of social psychology (Harvey, Ickes, & Kidd, 1976, 1978, 1981). Around 1980, cross-fertilization began between attribution and early social cognition. Accordingly, the third section describes stage models of dispositional inference (Gilbert, Pelham, & Krull, 1988; Trope, 1986). These models linked the information-processing approach of cognitive psychology (Neisser, 1967) and the nascent literature on automaticity (e.g., Bargh & Pietromonaco, 1982; Schneider & Shiffrin, 1977) with more traditional attribution concepts such as causal discounting (Kelley, 1972a). The final section of the chapter focuses on attribution of mind, which concerns inferences about mental states such as beliefs, goals, and motives (Epley & Waytz, 2010; Malle & Hodges, 2005). Recent contributions of neuroscience to attribution are also noted (Lieberman, 2010; Mason & Morris, 2010). The developments in these areas once again illustrate the interconnections between attribution and social cognition as they co-evolve.

The focus of the chapter is necessarily selective, emphasizing cognitive process, especially as it relates to later developments in social cognition. Where appropriate, I attempt to link classic ideas and findings with references to more current work on the same issues. Readers interested in more in-depth coverage of causal attribution, per se, and the application of attribution concepts to a variety of social psychological topics are referred to other sources (Cheng, 1997; Forsterling, 2001; Gilbert, 1998; Harvey et al., 1976, 1978, 1981; Hewstone, 1989; Kelley & Michaela, 1980; Ross & Fletcher, 1985; Weiner, 1995).

Figure 6.1 Timeline of evolution for attribution and social cognition.
At the beginning of this chapter, I posed the question of why attribution held such wide appeal to social psychologists in the 1970s. The sudden surge of interest in attribution can be traced to several factors (Aronson, 1997). First, although social psychology was always cognitive (Asch, 1946; Heider, 1944, 1958; Fiske & Taylor, 1991), systematic theoretical advances in the 1960s demonstrated the breadth and persuasiveness of an attributional analysis (Jones & Davis, 1965; Kelley, 1967). For instance, a topic such as source credibility in persuasion took on new significance when viewed through the attributional lens: Did a communicator advocate position A because of a personal bias or because of the soundness of that position? Second, the popularity of cognitive psychology contributed to a zeitgeist whereby motivational explanations (e.g., cognitive dissonance) were considered passé. Methodological issues played a role as well. In the wake of Milgram’s (1974) controversial studies of obedience and the sometimes elaborate laboratory hoaxes needed to investigate cognitive dissonance, many social psychologists became wary of deceiving research participants. Attribution studies could be done simply, without deception. Most important, the new approach quickly yielded a plethora of important new findings (Harvey et al., 1976; Jones et al., 1972; Jones, 1979; Kelley & Michela, 1980; Nisbett & Ross, 1980; Weiner, 1979).

**Inspiration and Pioneering Ideas from Heider**

Where attribution is concerned, it is no exaggeration to say that it all began with Fritz Heider (1944, 1958). Although others such as Gustav Ichheiser (1970) and Albert Michotte (1963) also made lasting contributions, Heider (1958) summarized the most important of these in his magnum opus, *The Psychology of Interpersonal Relations*. Heider’s insight was deep and penetrating. More than 50 years after his book appeared, it is possible to trace most of the important advances in attribution to elements in that book. It is well known that many of the early theoretical statements of attribution drew heavily from Heider (Jones & Davis, 1965; Jones & McGillis, 1976; Kelley, 1967, 1973; Reeder & Brewer, 1979; Weiner, 1979). Perhaps less acknowledged is that Heider hinted about additional topics that later appeared as empirical findings in the literature. A partial list of these topics would include covariation and discounting of causes (Kelley, 1967; 1973), correspondence bias (Gilbert & Malone, 1995; Jones, 1979), attribution of achievement (Weiner, 1979), the actor versus observer bias (Jones & Nisbett, 1972; Malle, 2006), false consensus (Ross, Greene, & House, 1977), attribution of responsibility (Fincham & Jaspers, 1980), and self-serving attribution biases (Taylor & Brown, 1988). Heider’s contributions follow four themes: (1) naïve psychology matters, (2) perceivers draw causal inferences, (3) gestalt principles organize impressions, and (4) attributions are biased by perceiver motivations.

**Naïve Psychology Matters**

All of us are amateur or naïve psychologists, gathering information about the people around us—and often struggling to understand it. From a distance, we wonder why political figures act as they do, while closer to home we try to figure out the nature of our spouse or lover. Heider’s great insight was to take these perceptions seriously. He assumed not only that social perceptions were of interest in their own right but also that such perceptions underlie and determine all our social interactions with other people. Our unique (and sometimes biased) construal of events is what matters (Ross & Nisbett, 1991). For instance, if we think a candidate for political office is dishonest, we are unlikely to vote for or provide financial support for that candidate. Likewise, if we come to view a spouse or lover as cruel and heartless, we are more likely to end the relationship. Because of its central role, Heider believed that the study of naïve psychology (or social cognition, as we now know it), should take precedence over other areas of social psychology (Ostrom, 1984).

But what are the rules of naïve psychology? Heider noted that these rules are implicit, rather than explicit: We act in accord with these principles without necessarily being consciously aware of them (Nisbett & Wilson, 1977). Modern research on implicit social cognition and the automatic influence of attitudes certainly validates Heider’s position (Fazio & Olson, 2003; Greenwald & Banaji, 1995). Heider also believed that language and communication styles need to be explicated so that we may understand their subtle influences (Hilton, 1995; Semin & Fiedler, 1988). For example, concepts such as desire, ought, and revenge were analyzed in detail (Heider, 1958).

**Perceivers Draw Causal Inferences**

People see events in both the physical and social world as causally connected. Just as a dam or levee may break because of heavy storm waters, a person’s angry outburst may be seen as caused by a stressful day at the office. Heider suggested that perceivers put
the various causes of human behavior into two categories: causes within the person (e.g., dispositional characteristics such as having a “quick temper”) and causes within the environment or surrounding situation (e.g., looming deadlines in the workplace). Three noteworthy themes emerge in the discussion of these causal forces. First, perceivers expect causes to interact in systematic ways. Second, perceivers prefer certain stable, invariant causes (e.g., traits) over other types of explanations. Finally, Heider noted some differences between personal causality (i.e., intentional acts) and impersonal causality (i.e., unintentional acts).

EXPECTATIONS ABOUT CAUSAL INTERACTION

Heider’s (1958) rather simplistic distinction between dispositional and situational causality became the bedrock assumption of early attribution theories. If behavior was seen as due to the actor, perceivers could feel free to attribute corresponding traits and dispositions to the target person, whereas if the behavior was seen as caused by situational forces, little could be learned about the target person (Bem, 1967; Jones & Davis, 1965; Kelley, 1967). In other words, dispositional and situational causality were placed at odds with one another—in a hydraulic relationship—where strength in one implied weakness in the other. Although the hydraulic idea may be intuitively appealing, recent findings suggest that perceivers often see both dispositional and situational factors as necessary to produce behavior (Gawronski, 2004; McClure, 1998; Reeder, Vonk, Ronk, Ham, & Lawrence, 2004). For instance, situational forces (such as an incentive to steal) tend to express a given trait (such as dishonesty). Thus, rather than discounting the role of a disposition, sometimes the situation can explain why the disposition came to be expressed.

Heider (1958) provided a more elaborate description of causal interaction in his “naïve analysis of action.” For instance, imagine a person who attempted to row a boat across a lake. Progress toward the goal will be seen as a joint function of forces within the person (such as effort and ability) and situational forces (such as a wind that either hinders or facilitates the boat’s progress). Heider described the relation between effort and ability as multiplicative: When a person with high ability exerts high effort on a task, the resulting effect should be much more dramatic than when a person of low ability exerts the same effort (Anderson & Butz, 1974; Reeder, Hesson-McInnis, Kroshe, & Scialabba, 2001). For instance, if a person is weak and uncoordinated, the person’s efforts to concentrate on rowing will have minimal impact.

PERCEIVERS PREFER INVARIANT CAUSES

When looking for an explanation, perceivers want to know the underlying cause. Heider (1958) provided an example in which he found sand on his desk, traced it to a crack on the ceiling and, ultimately, to a weakness in the walls. So too, in Heider’s view of social perception, perceivers are less interested in the changing circumstances surrounding behavior than in stable factors that may reliably determine actions in the future. If perceivers are searching for such invariance, they often find it in a target person’s traits and dispositions. In making such dispositional attributions, Heider (1958) implied that perceivers are prone to underestimating situational forces. At least two tendencies may account for such a bias. First, target persons are seen as “action centers,” sufficient causes of their behavior. Second, a target’s behavior stands out (relative to the situation) and tends to capture our attention: “…behavior in particular has such salient properties it tends to engulf the total field . . .” (p. 54). The general notion that perception of causality follows causal salience proved influential (Pryor & Kriss, 1977; Taylor & Fiske, 1978). Nevertheless, Heider’s observations about “behavior engulfing the field” led to controversy. On the one hand, his emphasis of dispositional attribution set the stage for later research on the correspondence bias (Gilbert & Malone, 1995; Jones, 1979) and the so-called fundamental attribution error (Ross, 1977). On the other hand, much research indicates that perceivers can be highly attentive to situational information (Gawronski, 2004; Malle, Knobe, & Nelson, 2007), even when their goal is to form trait judgments (Gosling, Ko, Mannarelli, & Morris, 2002; Reeder, Monroe, & Pryor, 2008).

PERSONAL CAUSALITY VERSUS IMPERSONAL CAUSALITY

Most researchers in attribution took their cue from Heider’s distinction between dispositional and situational causes, a distinction concerned with the locus of causality. Researchers gave less attention to a related distinction Heider (1958) drew between personal and impersonal causality (see also Malle, 2004, 2008). Personal causality refers to instances in which action is produced intentionally—with a purpose behind it. In contrast, impersonal causality can refer to environmentally produced events.
(as when rocks fall from a mountain) or a person's unintentional actions. The reader may rightly ask how the personal causality distinction (personal vs. impersonal) differs from the locus of causality distinction (dispositional vs. situational). The difference can be clarified in the case of unintended behavior. Suppose a person is seen to fall on an icy sidewalk. The behavior is clearly unintentional, an instance of impersonal causality. But locus of causality is still at issue: Was the fall due to something about the person (e.g., a lack of coordination) or due to the icy conditions? In more recent discussions of this issue, researchers suggest that intentional actions (such as donating to a charity) are typically explained in terms of a person's goals, motives, or reasons (Malle, 1999; Reeder, 2009a; White, 1991). Such reason explanations often recognize the role of situational factors (e.g., the recipients of the donation were very needy). In contrast to intentional behavior, unintentional acts are explained in terms of causal locus: Was the act caused by something in the person or something in the situation?

Given the goal directed nature of intentional actions, Heider (1958) noted at least two implications. First, perceivers assume that intentional actions are characterized by equifinality. By this, Heider meant that when a person has a goal in mind, we expect that a variety of strategies or means may be employed to achieve that goal, depending on the circumstances. For instance, a graduate student who has the goal of becoming a college professor may find herself switching research topics or changing faculty mentors if new job opportunities require a different research emphasis. In other words, given a constant goal, the means to achieve the goal may show much variability.

A second implication is that perceivers tend to hold people more responsible for their intentional actions than their unintentional actions (Alicke, 2000; Hart, 1968; Malle, 2004; Shaver, 1985; Weiner, 1995). For instance, if a person threw a stone with the intention of hitting someone, we would judge the person more harshly than if the stone had gone off course and produced accidental harm.

But Heider (1958) acknowledged that matters of responsibility can be more complicated. He hinted at a model of responsibility with five levels (Shaw & Sulzer, 1964). First, based on mere associations, people may be held at least partly responsible for outcomes not of their own doing (Collins & Loftus, 1975; Gawronski & Bodenhausen, 2006; Skowronski, Carlton, Mae, & Crawford, 1998). For instance, a tourist visiting in a foreign country may be condemned for outcomes of a war (involving the tourist's home country) that took place centuries ago. At the second level, people may be held responsible for outcomes they caused unintentionally. Thus, a person who was driving within the speed limit may be held partly responsible for hitting a child who suddenly darted out between two parked cars. The third level of responsibility involves cases in which the outcome could have been foreseen, but was not intended. Thus, a person may be held more responsible for an auto accident when speeding was involved. Presumably, the driver should have realized the increased dangers that come with speed. In making such a judgment, the perceiver may consider alternative scenarios in which a more thoughtful or moral person would not have behaved in this way (Hilton & Slugoski, 1986; Roese, 1997). The fourth level concerns cases in which a person's actions appear at least partly justified by the situation. For instance, if a supervisor lost her temper in response to the dishonest behavior of a subordinate, we might say she was "provoked." Responsibility, in this case, would be discounted relative to a situation that involved lesser provocation (Kelley, 1972a; McClure, 1998). Finally, at the fifth level, people are held fully responsible for their intentional acts. Thus, a wealthy person who swindled small investors might be seen in a particularly negative light.

**Gestalt Principles Organize Impressions**

Like other pioneers of social cognition (Asch, 1946; Festinger, 1957), Heider stressed consistency and organization in social perception (Koffka, 1935). First, perceivers search for the meaning behind action. In an early demonstration, Heider and Simmel (1944) created a film with black geometric figures moving against a white background. Perceivers who viewed the film engaged in a bit of anthropomorphism (Epley, Waytz, & Cacioppo, 2007), endowing the figures with motives in order to make sense of what they saw. Thus, when a large triangle moved toward a circle, perceivers saw the triangle as a "bully" who was chasing the circle. Similarly, when we read a book or watch a movie, our attribution of motives to the characters allows us to comprehend the meaning of events and better understand how the different aspects of an individual's personality fit together (Read & Miller, 1993, 2005; Reeder & Trafimow, 2005).

A second indicator of the search for consistency is that persons and their acts tend to be seen as one (Heider, 1944). Although Heider believed that actual behavior reflected an interaction between
both the environment and the person, he suggested that naïve perceivers stress the role of personal causes in this equation (Jones, 1979; Ross, 1977). Heider (1944, 1958) also stressed certain types of content in the impression, particularly competence (or ability, as indicated by success) and morality (or warmth), which may organize our overall reaction to the person (Fiske, Cuddy, & Glick, 2006; Reeder & Brewer, 1979; Skowronski & Carlston, 1989; Wojciszke, 2005).

Third, consistency can lead perceivers to try to "match" a cause with its effect. The linking of cause with effect is sometimes as simple as assuming that a bad act must have been caused by a bad person. But Heider (1958, p. 51) also described more complex matching strategies that promote economy in the impression. For instance, when perceivers are confronted with a series of acts from a target person, they try to identify a cause or motive that integrates or reconciles the different behaviors. Suppose a young worker, who appeared to worship a company executive, was promoted by that executive to a position of authority. But once entrusted with the trade secrets of the company, the worker quit the job and took a lucrative position with a competing company. In order to account for the full pattern of behavior, perceivers may cite a single motive—ambition or perhaps avarice—that links the different actions together.

**Attribution Can Be Biased**

Finally, Heider’s (1944) perceiver also engaged in a good deal of wishful thinking: “The tendency to keep the ego level high is a good example” (p. 368). When seeking an explanation for an outcome, perceivers are likely to choose a cause or reason that (1) fits their own wishes and (2) fits with the available data. Two manifestations of such bias relate to the self-serving bias in attribution (Bradley, 1978; Taylor & Brown, 1988) and the tendency to perceive events egocentrically (Inhelder & Piaget, 1958; L. Ross & Ward, 1996; M. Ross & Sicoly, 1979). In line with a self-serving bias, Heider suggested that people can be quick to deny personal responsibility for a failure or frustrating event, while eager to place the blame on a scapegoat (Fein & Spencer, 1997; Haslam, 2006).

In discussing egocentrism, Heider’s view of the perceiver was quite similar to what we now call the naïve realist (Ross & Ward, 1996)—one who believes that he or she sees objects and events in the world objectively (as they really are). Being unaware of subjectivity has several consequences. First, a person may not realize the influence that he or she has on other persons (Heider, 1958, p. 55; Gilbert & Jones, 1986). For instance, a professor with a hilarious sense of humor may create a humorous classroom atmosphere, leading the professor to mistakenly conclude that students typically joke-around in the classroom. Second, perceivers may not realize that the situation they see is not the one that others see. When Marie Antoinette was informed that the people of France were hungry because they had no bread, she is reputed to have asked why they did not eat cake instead. Finally, perceivers tend to engage in social projection, believing that others hold similar attitudes to their own (Robbins & Krueger, 2005; Ross, Greene, & House, 1977).

Despite the multitude of topics covered here, this review barely scratches the surface of what Heider (1944, 1958) offered to social cognition. The difficulty in summarizing his work relates, in part, to the countless insights he offered. Many of his writings also lack the systematic organization that one expects from scientific theory. Nevertheless, future researchers are likely to continue to mine his writings for inspiration and new ideas to study. Forty years ago, when graduate students were looking for an interesting thesis topic, they were advised to "read Heider!" For the serious student of social cognition, the advice remains as true today as it was then.

**Classic Attribution: Systematic Theory, Findings, and Applications**

Two theoretical papers in the 1960s are largely responsible for the rapid growth of attribution in the decade to follow. In 1965, Jones and Davis proposed the first systematic model of dispositional inference. Two years later, Kelley (1967) published a more general theory of causal thinking in social situations. Kelley’s analysis applied not only to attributions about persons but also to attributions about situational forces. These two papers, along with the volume by Jones et al. (1972), had immediate impact. A number of important empirical findings followed quickly, and the attribution approach was applied widely in social psychology and elsewhere.

**The Jones and Davis (1965) Theory of Correspondent Inferences**

We see other people behave in contexts in which a variety of behavioral choices are available to them. Job candidates decide how to present themselves in an interview, depending on the nature of the job. High school seniors select a college to attend among those schools to which they were accepted. When
we witness these choices being made, how do we come to infer the stable characteristics (i.e., traits and attitudes) of the actors involved? Jones and Davis (1965) addressed this question in their theory of correspondent inferences.

Following Heider (1958), Jones and Davis began with the assumption that perceivers seek to understand the social interactions they witness and to attribute stable characteristics to those involved. Any behavior could potentially be analyzed in terms of an endless causal regress, whereby immediate action (e.g., U.S. President Obama criticized oil company executives) can be traced to a prior event (an oil rig exploded), which in turn is explainable by earlier causes (lax safety regulations). Yet perceivers mainly seek a sufficient cause, rather than a complete explanation. According to Jones and Davis, the sufficient cause typically comes in the form of an intention. By intention, Jones and Davis meant the reason or motive underlying the behavior (Malle, 2004; Reeder, 2009a). Knowing a person’s intention, then, allows the perceiver to make a trait inference.

It is important to note that Jones and Davis limited their attributional analysis to intentional action, apparently in the belief that a target person’s unintentional acts carry less relevant information about the target. Assumptions about the target person’s knowledge and ability were seen as playing a crucial role in determining intentionality (Malle & Knobe, 1997). If a person is assumed to have lacked foreknowledge of the consequences of his or her behavior, or lacked the ability to bring about those consequences, those consequences cannot be seen as the reasons for the behavior. For example, imagine that a small child tripped while carrying a 45-caliber pistol, leading the gun to discharge and kill someone. We would be unlikely to think the child intended harm if (1) the child thought the gun was unloaded and (2) the child lacked the strength to willfully pull the trigger.

**INFERRING INTENTIONS**

When the assumptions of knowledge and ability are satisfied, how do perceivers infer a person’s specific intention? In considering this question, Jones and Davis focused on cases in which target persons made choices between available alternatives. The unique effects associated with the chosen alternative are then viewed as revealing of the target’s intention. For instance, imagine that a talented high school senior named Emma has been admitted to two prestigious universities, State U and Private U. Both offer quality programs, but State U has lower tuition, and is closer to home, and some of Emma’s friends will enroll there. The main attraction of Private U, on the other hand, is its location in a city with many cultural opportunities. But Emma is aware that the city also has a high crime rate. Some of the effects of attending each university are shared or common. For instance, both schools are prestigious and have quality programs. According to Jones and Davis, these common effects are relatively uninformative: If Emma selected State U over Private U, we would be unlikely to say the quality of the school was the main reason for her selection.

In contrast, effects that are unique to the chosen alternative are more revealing of a target person’s intentions. In the example above, if Emma chose to attend State U, we would explain her choice as due to some combination of its low cost, its proximity to home, and the fact that her friends are going there. On the other hand, if Emma selected Private U, the uniqueness of its cultural opportunities would likely be cited as the reason for her choice. Jones and Davis acknowledged additional complexities in that unique effects that are socially desirable are more likely to be seen as reasons for the choice (e.g., Because the cultural opportunities of Private U are socially desirable, whereas its crime rate is not so desirable, perceivers should see cultural opportunities as motivating Emma’s choice). In addition, the number of unique effects is important as well. Perceivers should be more certain of the target’s intention when the chosen alternative has only one such socially desirable effect, as opposed to when there are many such effects (Jones, Davis, & Gergen, 1961; Newton, 1974). Therefore, perceivers should be more certain of Emma’s intention if she selects Private U over State U (because State U is associated with many socially desirable effects).

In summary, only effects that the target person knew about and had the ability to produce are informative about a person’s intentions. The perceiver considers the different alternatives from which the target chose and notes the effects that are unique to each alternative. If the chosen alternative has only one socially desirable and unique effect, the perceiver will be confident that the target had the intention of bringing about that effect.

**THE DETERMINANTS OF CORRESPONDENT INFERENCES ABOUT DISPOSITIONAL CHARACTERISTICS**

Once an intention is isolated, perceivers are in a position to infer the dispositional characteristics of the target person. Jones and Davis introduced the
term correspondence in this context. Correspondence refers to the extent to which a target person's behavior—and the intention on which it is based—are similar to the underlying characteristics that are attributed to the target. Consider the example in which Emma chose to attend Private U because she desired cultural opportunities. A correspondent dispositional inference would occur to the extent the perceiver views Emma as a cultured, sophisticated person. In other words, high correspondence is present if the perceiver infers a disposition that is congruent with (or corresponds to) the observed behavior.

Yet perceivers may not always attach much significance to knowing that a person acted with a given intention or wanted a particular outcome. What, then, are the determinants of correspondent inferences? Jones and Davis (1965) suggest that correspondence is high for an act that is (1) socially undesirable, (2) freely chosen, and (3) done for a clear reason (i.e., there is only one unique effect). The role of these three factors can be illustrated by the unflattering dispositional characteristics we would attribute to a bank teller who stole $10,000. Our negative reaction reflects our beliefs that the act of stealing was socially undesirable, unforced, and clearly done for the purpose of obtaining ill-gotten gain.

**The Role of Perceiver Motivations**

Jones and Davis (1965) also followed Heider by assuming that attributions can be biased. The main biasing factors are hedonic relevance and the related concept of personalism—both of which increase the extent of correspondence between actions and attributed dispositional inferences. Hedonic relevance is present to the extent that a target person's actions produce consequences that either “…promote or undermine the perceiver's values…” (p. 237). For instance, if the owner of a company made bad business decisions that put the company workers' pensions in jeopardy, company workers would likely make harsh judgments about the owner (Jones & deCharm, 1957). As the social cognition perspective developed, the presence of such “outcome dependency” was found to also influence attentional processes related to seeking accuracy and greater individuation in the impression (Fiske & Neuberg, 1990; Neuberg & Fiske, 1987).

In the example above in which the company owner made bad decisions, there is no reason for the workers to assume negative intent on the owner's part. In contrast, personalism represents a special case of hedonic relevance in which the perceiver believes the actor intended to alter the outcomes of the perceiver. For instance, if the company owner was seen as deliberately slashing worker pension benefits in order to increase his or her own wealth, the example would qualify as an instance of personalism: Not only were benefits decreased, but pensioners could also assume that the owner was well aware of this effect. Thus, personalism represents an extreme case of hedonic relevance that is revealing of the actor's motives and can lead to especially extreme reactions to the target person (Dodge & Crick, 1990).

In summary, Jones and Davis (1965) proposed the first systematic model of dispositional inference, and many of their ideas were incorporated into the social cognition literature, including intentional-ity (Malle & Knobe, 1997), mental states such as motives (Ames, 2004; Dodge & Crick, 1990; Read & Miller, 1993; Kammrath, Mendoza-Denton, & Mischel, 2005; Reeder, 2009a), chosen and non-chosen alternatives (Roese, 1997), and motivational bias (Fiske & Neuberg, 1990).

**Kelley's Covariation Approach to Causal Attribution**

More so than Jones and Davis (1965), Harold Kelley's work (1967, 1973) addressed the central question of causal attribution: How do perceivers decide if a given effect (e.g., a target person's behavior) was caused by something about the person or by environmental factors surrounding the target's behavior? In addition, whereas Jones and Davis focused on intentions and dispositions internal to the target person (attempting to rule out situational determinants), Kelley aimed to illuminate external attributions about the situation or environment. Following Heider's (1958) reliance on J. S. Mills' method of difference, Kelley proposed covariation logic as the crucial tool of attribution: “The effect is attributed to that condition which is present when the effect is present and which is absent when the effect is absent.” For instance, if a child typically misbehaves only after ingesting large quantities of sugar, the misbehavior will be viewed as caused by the sugar. In this example, causality will be assigned externally (to the sugar), rather than internally (to the child).

Kelley (1967) further elaborated his model by identifying three important sources of information that are used in the covariation analysis: (1) The distinctiveness of the behavior across similar stimuli, (2) the agreement or consensus between the target person's reaction to the stimulus and the reactions.
of other people, and (3) the consistency of the target person's behavior across time and modality (i.e., ways of being exposed to the stimulus). To apply the model, consider the example of the child who misbehaved after eating sugar. The perceiver will be highly confident of an external attribution under the following pattern of information: Distinctiveness is high (the child misbehaves only after eating sugary foods), consensus is high (other children also tend to misbehave after eating sweets), and the child's misbehavior after eating sweets is relatively constant across time and modality (at school, home, and on the ball field). On the other hand, the perceiver would be more confident of an internal attribution given a pattern of this sort: Distinctiveness is low (the child misbehaves after eating a wide variety of foods), consensus is low (other children typically do not misbehave after eating sugar), and the child's (negative) reaction to sweets is relatively consistent (McArthur, 1972).

Notice that the covariation model (Kelley, 1967) assumes the perceiver was able to accumulate quite a bit of information about the target person's behavior over time, all the while keeping track of the behavior's covariation with other types of stimuli (distinctiveness), the reactions of others (consensus), and different locales (modality). Perhaps Kelley overestimated the capabilities of the perceiver. For instance, subsequent social cognition research on illusory correlation suggested that perceivers have considerable difficulty assessing covariation between social groups and the positive versus negative behaviors of the group members (Hamilton & Gifford, 1976; see also Alloy & Tabachnik, 1984).

In later papers, Kelley (1972b, 1973) simplified matters by acknowledging that perceivers often make attributions after only a single observation of behavior. With restricted information, perceivers can still react to the configuration of causes, based on their expectations or schemata about how causes typically interact. According to Kelley's discounting principle, the role of a given cause will be discounted if other possible causes are also present. Thus, if an interviewee claimed to be an extravert while applying for a job that required a person who could work long hours in solitude, this situation, the interviewee should be seen as especially extraverted.

The discounting and augmentation principles make essentially the same predictions as the more elaborate Jones and Davis (1965) analysis of unique effects. Consequently, over the next 20 years, Kelley's more parsimonious account tended to be incorporated into models of dispositional inference (Gilbert et al., 1988; Reeder & Brewer, 1979; Trop, 1986). For the time being, at least, Jones and Davis' (1965) analysis of intentionality, and its concern with perceived intentions, fell by the wayside.

SELF-ATTRIBUTION
Kelley (1967) viewed his covariation model as applying to judgments about the self as well as to judgments about others. Consider the strength with which a person holds a given attitude toward an object (Petty & Krosnick, 1995). Kelley suggested that a person would be more committed to the attitude to the extent that he or she typically feels the same way about that object (consistency) and notices that other people feel the same (consensus).

Kelley (1973) also applied the discounting principle to self-attribution (Bem, 1972; Lepper, Greene, & Nisbett, 1973; Miller & Ross, 1975; Nisbett & Wilson, 1977; Taylor & Brown, 1988). When Bem (1967) proposed his theory of self-perception as an alternative to cognitive dissonance theory (Festinger, 1957), he drew a parallel between attributions to self and attributions to others. When making either type of attribution, the perceiver compares the relative strength of internal and external forces operating on the actor's behavior. For example, consider the forced-compliance studies of attitude change first popularized by cognitive dissonance theorists. In the most famous of these studies (Festinger & Carlsmith, 1959), an experimenter offered participants either $1 or $20 to tell another person that a tedious peg-turning task was "interesting." Those who complied for the low incentive typically came to believe in the attitudinal position they advocated more so than those who complied for a higher incentive. Bem (1967) demonstrated the exact same pattern of inference among "outside" observers, who were merely told that the subject complied for either the low or high incentive. Kelley (1973) interpreted this evidence as support for his discounting principle: Both the original actors and their observers apparently concluded that compliance in the face of a large, external incentive ($20)
was relatively unrevealing about the (internal) attitudes of the actor.

**BIAS IN ATTRIBUTION**

Kelley (1967) pointed to some important sources of bias in attribution, including a tendency toward pluralistic ignorance (Prentice & Miller, 1996), whereby people assume that they, unlike other people, are immune to social influence and conformity pressures [see also Pronin's (2007) concept of the bias blind spot]. Kelley (1967) also gave insightful descriptions of situations that can mislead people into making biased attributions. For example, recall that in forced-compliance studies (Bem, 1967; Festinger & Carlsmith, 1959), both the original actors and their observers attributed stronger attitudes (in line with the actor's behavior) when the incentive for complying was small (e.g., $1), as opposed to large ($20). This pattern implies that perceivers assumed different degrees of consensus for compliance in the two incentive conditions, such that almost everyone was thought to comply with the experimenter's request in the $20 condition, but not in the $1 condition. Given this assumption, perceivers could feel justified in attributing an extreme attitude for "low consensus" behavior in the $1 condition. But the assumption of differential compliance is mistaken—compliance rates in both conditions are near 95%. The error represents an illusion of freedom, whereby perceivers overlook the fact that the actor in the low incentive condition complied because of subtle social pressure applied by the experimenter.

In summary, Kelley (1967, 1973) extended Heider's (1958) theorizing by focusing squarely on causal attribution as an instance of seeking subjective validity. The covariation principle and causal schemata that he proposed could be applied to both self and other perception. Although both Jones and Davis (1965) and Kelley (1967, 1973) anticipated bias in attribution, their basic models implied a highly logical attributional calculus—a rational baseline (Jones & McGillis, 1976). This idealized view of the perceiver proved to be somewhat unrealistic. As empirical findings and applications accumulated in the attribution literature, perceivers came to be portrayed in less flattering fashion—as falling prey to correspondence bias, self-enhancement bias, and naive realism/social projection. The emerging research also demonstrated the importance of considering the dimensions of causality (e.g., locus, stability, and controllability) and cultural differences in attribution.

**Important Findings and Applications**

**CORRESPONDENCE BIAS**

The correspondence bias refers to a tendency for dispositional inferences to be based on observed behavior and insufficiently adjusted for situational factors. In an influential study by Jones and Harris (1967), perceivers read about a target person who wrote an essay either defending Fidel Castro's record in Cuba or criticizing that record. In the "free choice" condition, the target could write on either side of the issue. In the "no choice" condition, however, the target was assigned to support one side of the issue, and the resulting essay complied with the assignment. In line with what Kelley (1973) later called the "discounting principle," perceivers inferred less extreme, essay-correspondent attitudes in the no choice than free choice condition. Nevertheless, discounting in the no choice condition was "incomplete." That is, perceivers still tended to infer attitudes that corresponded to the direction of the essay the target composed. The correspondence bias proved to be remarkably robust, although it could be diminished somewhat if situational constraints on the target were emphasized (Jones, 1979).

A variety of theoretical factors are apparently involved in correspondence bias (Gawronski, 2004; Gilbert & Malone, 1995; Ross, 1977). First, perceivers may fail to notice or give inadequate attention to situational forces surrounding behavior (Heider, 1958; Ross, Amabile, & Steinmetz, 1977; Taylor & Fiske, 1978). Second, situational expectations (e.g., expecting sad behavior at a funeral) can lead perceivers to make an inflated behavior categorization (e.g., identifying a neutral expression as "sad"), leading to a relatively extreme dispositional inference about the person (e.g., this woman really is depressed), as described by Trop's (1986) model of dispositional inference. Third, adjusting a dispositional inference for situational factors may require cognitive resources, which may be in short supply for perceivers involved in a social interaction (Gilbert et al., 1988; Gilbert, Krull, & Pelham, 1988). Fourth, when perceivers are suspicious of a target person's ulterior motives, they may engage in more extensive processing, thereby increasing the use of situational information—which can decrease correspondence bias (Fein, 1996; Reeder et al., 2004).

Finally, perceivers may have unrealistic expectations concerning how situations affect behavior. Reeder and Brewer (1979) elaborated on Kelley's (1973) causal schemata by describing perceivers' expectations about how traits and attitudes are
related to the expression of corresponding behaviors. For instance, perceivers expect that writers will compose essays reflecting their personal opinions and think that an experimenter’s instructions are insufficient to lead writers to express alternative opinions. Consequently, regardless of the situational constraints surrounding the creation of an essay, perceivers may see the essay as highly diagnostic of the writer’s true opinion (Gawronski, 2003; Reeder, Fletcher, & Furman, 1989). Trait-behavior expectations can also explain the tendency for behaviors demonstrating either high ability or low morality to produce highly correspondent trait inferences (Reeder, 1993; Skowronski & Carlson, 1989; Trafimow & Schneider, 1994; Trafimow & Trafimow, 1999; Vonk, 1998; Vonk & Van Knippenberg, 1994). In the case of ability, perceivers expect that a person with high ability (e.g., a professional tennis player) can perform at either a high level or a low level, depending on motivation. But perceivers expect that a person with low ability can only perform at a low level. In the case of morality, however, immoral persons are expected to emit both moral and immoral behaviors, whereas highly moral persons are expected to always behave in a moral fashion. The upshot of these asymmetrical trait-behavior expectations is that behavioral demonstrations of either high ability or immoral behavior lead to especially correspondent inferences (i.e., that the target person has high ability or is immoral, respectively). These patterns occur because persons without the corresponding disposition are assumed unlikely to have behaved in that fashion.

SELF-ENHANCEMENT BIAS

The pioneers of attribution (Heider, 1958; Jones & Davis, 1965; Kelley, 1967) proposed a self-enhancement bias in the perceiver, and subsequent research clearly validated that assumption (Bradley, 1978; Taylor & Brown, 1988). For example, perceivers are more likely to credit themselves for a successful outcome than a failure. At least in part, this self-serving bias reflects motivational factors related to maintaining high self-esteem and protecting the self from threats (Miller, 1976; Sedikides, Gaertner, & Vevea, 2005; Taylor & Brown, 1988). Still, a tendency to take relatively greater credit for success can also reflect seemingly rational tendencies—such as intending to succeed, rather than fail (Miller & Ross, 1975)—and self-enhancement biases can also be subject to cultural differences (Kitayama, Markus, Matsumoto, & Norasakkunkit, 1997).

ACTOR VERSUS OBSERVER DIFFERENCES

In an influential paper mentioned at the beginning of this chapter, Jones and Nisbett (1972) made a simple, yet provocative, suggestion: They proposed that actors tend to attribute their behavior to situational forces, whereas observers tend to attribute the actor’s behavior to dispositions within the actor. Notice that the observers’ tendency to draw dispositional inferences is reminiscent of the correspondence bias mentioned earlier. Jones and Nisbett (1972) suggested that actor versus observer differences could arise because actors typically have more background information about the situation surrounding their actions (e.g., the actor knows that her angry outburst was provoked by an earlier insult). In addition, actors tend to focus their attention on the situation at hand, whereas for the observer, the actor’s behavior draws attention because it is dynamic and appears figural against the background of the situation.

Although the Jones and Nisbett (1972) actor versus observer hypothesis seemed highly plausible to many social psychologists, subsequent empirical work was more equivocal. An early review of the literature found, for instance, that actors and observers did not differ in terms of their tendency to draw dispositional inferences for the actor’s behavior (Watson, 1982). A recent meta-analysis of 173 published studies reported further limitations (Malle, 2006). The average effect size for the traditional actor versus observer bias was quite small ($d < .10$), and moderator analysis indicated that the bias emerged under a restricted range of circumstances. For example, the traditional bias is strong for attributions about negative events (e.g., relative to observers, actors downplay their personal role in failures), but not for attributions about positive events (for which actors readily take credit). This latter pattern, of course, describes a self-serving bias (Taylor & Brown, 1988). Malle et al. (2007) suggest that there are indeed actor versus observer differences, but the real differences are not those of the traditional actor versus observer hypothesis. For instance, relative to observers, actors are more likely to explain their own behavior in terms of reasons (e.g., a situation that influenced their motivation). This difference may arise because actors are relatively more (1) aware of their own motives and (2) motivated to portray themselves as rational (which can be accomplished by citing good reasons for their actions).
ATTRIBUTION AS A GATEWAY TO SOCIAL COGNITION

NAÏVE REALISM AND SOCIAL PROJECTION

Naïve realism is the assumption that one perceives events “objectively,” as they actually are in the external world (Heider, 1958; Jones & Nisbett, 1972; Kelley, 1967; Ross & Ward, 1996). Social projection is one important manifestation of realism: Perceivers tend to think that other people see the world as they do. Thus, a person who believes in the existence of extraterrestrial beings and alien abduction is likely to think that others believe similarly (Gunther & Christen, 2002). Although this false consensus bias (Ross, Greene, & House, 1977) is remarkably robust, it is stronger for attributions about the ingroup than outgroup (Robbins & Krueger, 2005) and may be influenced by a variety of processes, including selective exposure to information, focus of attention, and both logical and motivational factors (Marks & Miller, 1987).

Another important manifestation of realism involves the perception of unfairness or bias in other people (Ross & Ward, 1996). For instance, fans of opposing teams tend to see more rule infractions on the side of the other team (Hastorf & Cantril, 1954), and political partisans are especially alert to a hostile media bias, whereby supporters of each side of a controversial issue think the media has been unfair to their side (Vallone, Ross, & Lepper, 1985). Moreover, political partisans tend to see people on the other side of the issue as ignorant of the facts and biased by both ideology and self-serving motives (Reeder, Pryor, Wohl, & Griswell, 2005; Robinson, Keltner, Ward, & Ross, 1995). In contrast, perceivers tend to see themselves—and others who agree with their opinions—as relatively free from bias (Pronin, 2007).

DIMENSIONS OF CAUSALITY

Perceivers can conceptualize the causes of a behavior in terms of a variety of dimensions (e.g., locus of the cause and its stability), with each dimension having a unique effect on cognitive, affective, and behavioral reactions to the behavior (Abramson, Seligman, & Teasdale, 1978; Weiner, 1985). Consider the case of a young scholar who took great pride in composing a manuscript, yet received a rejection letter after submitting it to a journal. According to Abramson et al. (1978), the scholar is likely to experience feelings of hopelessness, lowered self-esteem, and depression to the extent that he or she attributes the rejection to causes that are internal (“I am not very creative”), stable (“This is the third time this article has been rejected”), and global (“Even my love life is in the dumps”). On the other hand, by attributing to external, unstable, and specific causes, the scholar may maintain a more optimistic outlook (Seligman, 1990).

Bernard Weiner (1979, 1985; 1995) offered what is, perhaps, the most comprehensive analysis of the dimensions of causality. Weiner suggested that the stability of an outcome influences the expectancy of future success (e.g., a stable cause should lead to the reoccurrence of the outcome), whereas the locus of the cause affects self-esteem and pride (e.g., an internal attribution for success increases pride). In addition to locus and stability, Weiner stressed the controllability of outcomes. The controllability dimension influences a variety of social emotions as well as judgments of responsibility. In the case of social emotions, perceivers tend to feel anger when a target person fails because of a controllable cause (e.g., lack of effort), but pity when the cause is less controllable (e.g., lack of ability). In addition, targets are held more responsible for producing negative outcomes that were relatively more controllable (Weiner, 1995).

CULTURAL DIFFERENCES

As we are socialized in a given culture, we come to acquire implicit assumptions about causality. Individualist cultures such as Western Europe and the United States tend to emphasize personal causality for social behavior, whereas collectivist cultures such as China tend to emphasize more situational and holistic thinking about causality (Mason & Morris, 2010; Morris & Peng, 1994; Nisbett, Peng, Choi, & Norenzayan, 2001). For instance, when describing a crime such as murder, English-language newspapers tend to stress the dispositional characteristics of the perpetrator, whereas Chinese-language newspapers tend to stress situational explanations (Morris & Peng, 1994).

Social Cognition Meets Attribution: Stage Models of Dispositional Inference

The social cognition perspective emerged full-force with the nearly simultaneous publication of three volumes circa 1980 (Hastie, Ostrom, Ebbesen, Wyer, Hamilton, & Carlston, 1980; Higgins, Herman, & Zanna, 1981; Wyer, & Carlston, 1979). The impetus for these volumes appears to be two-fold. First, person perception researchers grew disenchanted with the then current algebraic (Anderson, 1974) and structural (Rosenberg & Sedlak, 1972) approaches in the field. Although Heider (1958) had hinted about a hierarchy of processing stages (p. 80–81), research before the 1980s typically shed
litt le light on underlying cognitive processes—the events that intervene between stimulus and response (Hastie & Carlston, 1980). Second, rapid advances in theory and method in cognitive psychology held the promise of better elucidating these processes. For instance, by employing measures such as response time and recall, researchers believed they could better understand the information processing sequence of (1) acquisition and encoding of information, (2) representation and retention, and (3) retrieval and decision making.

The shift toward social cognition soon began to transform models of person perception (see Figure 6.1). Dispositional inference now became a matter of anchoring/adjustment because dispositional inferences were viewed as being “adjusted” in response to perceived situational forces (Jones, 1979; Quattrone, 1982). Accordingly, initial dispositional inferences were made spontaneously, a finding called spontaneous trait inference (Winter & Uleman, 1984). On this foundation, stage models of dispositional inference could then be proposed, with the early stages of inference being relatively automatic and the later stages being of a more controlled nature (Gilbert et al., 1988; Trope, 1986).

**Anchoring/Adjustment**

The anchoring/adjustment heuristic describes a decision-making tendency whereby an initial estimate is adjusted (often insufficiently) for additional factors that may be relevant (Tversky & Kahneman, 1974). For instance, after watching a TV commercial that portrayed product X as superior to its competitors, a viewer might initially be very impressed, then realize the biased nature of the information source, yet still maintain a relatively positive view of product X (Gilbert, 1991). Quattrone (1982) proposed an explanation of the correspondence bias (Jones & Harris, 1967) based on this anchoring/adjustment idea. The innovation—straight from the playbook of social cognition—was that dispositional inferences often undergo a change as the attribution process unfolds. For instance, consider a perceiver who learned that a target person wrote a pro-marijuana essay after being assigned to write in favor of marijuana. The perceiver would begin by inferring a correspondent attitude in the target (“This writer must really love marijuana!”), but subsequently consider the situational constraints (“On the other hand, the experimenter did pressure the writer to praise marijuana”), and then make an adjustment (“Well, maybe the writer only likes marijuana”).

This view suggests that perceivers in the attitude attribution paradigm typically anchor on a correspondent attitudinal inference and then make insufficient adjustment for situational factors. In short, dispositional inference initially comes easy to the perceiver. Quattrone (1982) provided evidence for a general anchoring/adjustment explanation by demonstrating that, under special circumstances, perceivers could be led to anchor on situational factors and, ultimately, overemphasize those situational factors. Perceivers were told that a target person had either a pro or anti attitude toward marijuana legalization and were then shown an essay that the target wrote which expressed that very attitude. Perceivers also received information about the subtle ways that experimenters can influence their subjects. Notice that, despite the possibility of this situational influence, the target’s own attitude is a sufficient explanation for the direction taken in the essay. Nevertheless, in line with the anchoring/adjustment tendency, perceivers who read the pro-marijuana (rather than anti-marijuana) essay saw the situation as applying relatively more pressure to write in favor of legalizing marijuana. In other words, perceivers apparently anchored on an extreme situational attribution, and then made insufficient adjustment for the writer’s personal attitude as a cause of the behavior.

**Spontaneous Trait Inference**

The notion that dispositional inference comes easy to the perceiver received further support from research on spontaneous trait inference (STI). Winter and Uleman (1984) adapted the encoding specificity paradigm (Tulving & Thomson, 1973) by having research participants read behavioral descriptions (e.g., The secretary solves the mystery halfway through the book) that implied traits (i.e., clever). The participants were simply instructed to memorize the sentences, and traits words were never explicitly presented at study time. Nevertheless, when participants attempted to recall the sentences, the implied trait words served as powerful recall cues—even though participants were unaware of having made trait inferences at encoding. The overall pattern suggests that trait inference occurs as part of the normal comprehension process (Asch, 1946; Gilbert, 1991).

Although STI is now recognized as a robust phenomenon (Uleman, Saribay, & Gonzalez, 2008), the original Winter and Uleman (1984) method was limited because it was unclear whether the effect was due to an encoding process or a retrieval process (Carlston & Skowronski, 1994). For instance, it is...
possible that participants used the trait cue provided at recall (e.g., clever) to think of associated behaviors (e.g., correctly guessed the direction of the stock market) that could then cue recall of the sentence stimuli. Stronger evidence in favor of STI emerged when cognitive load at encoding was shown to decrease the effectiveness of traits as recall cues, implying that the locus of the effect was indeed at encoding (Uleman, Newman, & Winter, 1992). In addition, when encoding goals were manipulated, recall was similarly affected (Uleman & Moskowitz, 1994). Finally, Carlston and Skowronski employed a relearning paradigm in which participants first learned trait-implying behavioral descriptions that were accompanied by a photo of the actor (Carlston & Skowronski, 1994; Carlston, Skowronski, and Sparks, 1995). In a second task, participants were asked to learn pairings of photos and traits, such that some of the traits had been implied by the earlier behavioral descriptions. Recall for a trait was better when the trait had been implied by the (earlier presented) behavioral descriptions. This finding strongly suggests that traits were inferred as the behavioral descriptions were originally encoded, and this “learning” made it easier to relearn the association between the photo and the trait.

**Stage Models of Dispositional Inference**

As researchers began to suspect that some types of information could be processed with special ease (Quattrone, 1982; Winter & Uleman, 1984), it was just a matter of time before models of dispositional inference incorporated ideas about automaticity (Bargh & Pietromonaco 1982; Schneider & Shiffrin, 1977). Stage models of dispositional inference proposed by Trope (1986) and Gilbert (1989; Gilbert et al., 1988) were particularly influential. Each model suggested that the initial stages of processing were relatively automatic, followed by more controlled processing. These models and an important theoretical extension are described next.

**TROPE’S TWO-STAGE MODEL**

Heider (1958) noted that we often experience our perceptions as a given—they just occur to us—without consciously realizing the underlying mediating mechanisms. Trope (1986; Trope & Liberman, 1993) contributed by clarifying the perceptual mechanisms by which perceivers initially interpret and identify attribution-relevant information. At the initial identification stage of the model, perceivers notice three types of information: (1) behavioral cues (e.g., the target person appears slightly anxious), (2) situational cues (e.g., an interviewer asked the target about some embarrassing experiences in her past), and (3) prior cues about the target (e.g., from past experience, the perceiver believes the target is a calm person). Information is often ambiguous, however, and its interpretation is context dependent (Asch, 1946; Higgins & Bargh, 1987). For instance, consider the example of a woman who appears vaguely anxious while answering questions posed by an interviewer. Given that the target’s behavior is ambiguous (e.g., is she anxious or just pleasantly excited?), the perceiver’s identification of her behavior will be affected by other cues that are less ambiguous (e.g., if the interview situation involves the target answering anxiety-inducing questions, her behavior may appear more anxious). At the initial stage of processing, then, the identification of any given informational cue tends to be assimilated in the direction of other cues that are clearer, or less ambiguous. This “disambiguation” occurs more of less unconsciously and automatically (Bargh & Chartrand, 1999).

At the second stage of the model—dispositional inference—the identifications of behavior, situation, and prior information are combined by more controlled processing to produce a dispositional inference. Of greatest relevance, behavior cues contribute additively to the dispositional inference. Thus, if the target is identified as behaving in a very anxious manner, the target will tend to be seen as having a very anxious disposition. In contrast, situational information at this stage is subtracted from the dispositional inference. In the example above, the target person in the interview will be attributed a less anxious disposition to the extent the anxiety-inducing situation appears to account for her display of anxiety. Trope (1986) suggested that the perceiver’s reasoning at this stage reflects the application of causal schemata associated with the discounting principle (Kelley, 1973; Reeder & Brewer, 1979; Reeder, 1997).

By clarifying the relatively automatic processes by which informational cues are identified, Trope (1986) was able to reconcile apparently conflicting findings in the literature. Some studies reported evidence supportive of the discounting principle, whereby situational factors “subtract” from dispositional inferences (e.g., Jones et al., 1961). Nevertheless, many others reported weak or “insufficient discounting,” in line with the correspondence bias (Jones, 1979; Jones & Harris, 1967), and other studies found an opposite effect whereby situational forces apparently “added” to
a dispositional inference (e.g., Snyder & Frankel, 1976). Trope’s (1986) model was able to account for such additive effects of the situation by focusing on the identification stage. For instance, if a target person displayed vague signs of anxiety during an interview, that same behavior might be identified as strong anxiety by a perceiver who believed the target was answering anxiety-inducing questions. At the dispositional inference stage, when the situation is subtracted from the trait inference, that subtraction may not overcome the effect of the earlier, extreme behavior identification. Consequently, the target in this example could ultimately be seen as more dispositionally anxious precisely because she was asked anxiety-inducing questions.

**GILBERT’S THREE-STAGE MODEL**

In contrast to Trope (1986), who focused on early perceptual processing, Gilbert (1989; Gilbert, Pelham, & Krull, 1988) investigated the more controlled, downstream stage of dispositional inference. His model contains three stages: (1) categorization of behavior, which resembles Trope’s (1986) identification stage; (2) characterization, which involves an automatic, correspondent dispositional inference; and (3) correction, whereby the dispositional inference is adjusted for the influence of situations forces. The latter two stages of characterization and correction represent an “unpacking” of dispositional inference as the endpoint of the process.

Recall that Quattrone (1982) and research within the STI paradigm (Winter & Uleman, 1984) suggested that trait inference, or characterization as Gilbert called it, could be relatively automatic. A series of clever experiments by Gilbert (1989) supported this notion and—more importantly—demonstrated that the correction stage required cognitive resources. In one of these studies, participants watched a woman who behaved anxiously while being interviewed about topics that were either anxiety inducing (e.g., sexual fantasies) or relaxation inducing (e.g., world travel). In addition, half of those in each condition were simultaneously given a memory task that consumed additional cognitive resources. These cognitively busy perceivers made dispositional inferences of anxiety that were little affected by the interview topics. Additional studies demonstrated that a variety of cognitive demands in everyday life, such as trying to make a positive impression during a social interaction, can disrupt the correction stage of dispositional inference (Gilbert, 1989; Gilbert & Malone, 1995).

**EXTENSION OF THREE-STAGE MODEL**

The three-stage process described above assumes that perceivers begin with the goal of dispositional inference. Krull (1993) extended the model by suggesting that the process could be more flexible and dependent on the perceiver’s goals. In this mixed model, perceivers can begin with the goal of assessing the potency of situational factors (e.g., Is Bob’s success due to his taking a simple test?), automatically infer situational causality at the second stage of characterization (The test must be easy), and then adjust that inference to reflect the influence of the actor’s disposition (…then again, Bob might be a competent guy).

In summary, although attribution research provided one of the foundations for the development of social cognition, the social cognition perspective also transformed the way attribution questions were studied. Concepts from cognitive psychology such as anchoring/adjustment (Tversky & Kahneman, 1974) and encoding specificity (Tulving & Thomson, 1973) were instrumental in establishing that aspects of trait inference could occur relatively automatically (Quattrone, 1982; Winter & Uleman, 1984). Building on these findings, stage models of dispositional inference were proposed that integrated the automatic and controlled stages of processing. The confluence of attribution and social cognition was now complete, setting the stage for an exciting new topic, called attribution of mind, and a new method of inquiry, called social neuroscience.

**Current Directions: Attribution of Mind and Social Neuroscience**

The common thread of attribution research is social construal—the processes by which people come to understand social reality. Perceivers strive to make sense of human action by wondering about its causes (e.g., What is the President really like, and how much of what he says and does is shaped by outside forces such as current politics, long-range goals, and even family obligations?). In seeking to address these questions, attribution researchers have traditionally collected data on two types of content
and employed two types of methods. In terms of content, the preponderance of research examined either causal attributions (including judgments of control, responsibility, and blame) or trait inferences. The methodology either relied on simple paper-and-pencil surveys or (more recently) borrowed from cognitive psychology by using computers to precisely time and manipulate the presentation of stimuli and collection of responses. Exciting new research in the field is extending both the contents and methods of attributional inquiry. As described below, attribution of mind (Epley & Waytz, 2010) examines the beliefs, desires, intentions, and experiential capacities (e.g., feelings and emotions) we attribute to others. In addition, social neuroscience (Lieberman, 2010; Van Overwalle, 2009) examines the brain to advance our understanding of both traditional topics like trait attribution and newer topics related to mind attribution.

**Attribution of Mind**

Research on mindreading represents the newest extension of the attribution perspective (Epley & Waytz, 2010; Malle & Hodges, 2005). On the one hand, work in this area can be seen as extending attribution’s long-standing concerns with perceived mental states such as beliefs and intentions (Jones & Davis, 1965; Heider, 1958). On the other hand, the study of mindreading appears somewhat revolutionary owing to the multidisciplinary contributions from developmentalists (e.g., Apperly & Butterfill, 2009), neuroscientists (e.g., Iacoboni, 2009; Mitchell, 2009), and philosophers (e.g., Goldman, 2006). New questions are being raised about the attribution of psychological capacities (Epley, Waytz, & Cacioppo, 2007; Gray & Wegner, 2009), folk psychology and perceived intentionality (Lillard, 1998; Malle, 2004), and the attribution of mental states such as goals and motives (Ames, 2004; Read & Miller, 2005; Reeder, 2009a). As in the case of stage models of dispositional inference, most of these new research areas can be viewed as branches of social cognition.

**ATTRIBUTION OF PSYCHOLOGICAL CAPACITIES**

What qualities of mind separate an infant, a woman, a man in a persistent vegetative state, a frog, a chimpanzee, a dead woman, God, and a robot? When asked this question, people’s answers tend to fall along the two dimensions of conscious experience and intentional agency (Gray & Wegner, 2009). Conscious experience involves the capacity to feel emotions such as fear and experience states such as hunger. Intentional agency, in contrast, involves self-control, planning, and morality. Normal adults tend to be perceived as high on both of these dimensions, whereas robots and God are perceived as having agency, but little capacity for experience. Additional research examined the tendency to view physical objects (e.g., a computer) as human (Epley et al., 2007). Perceivers are prone to anthropomorphize when human-related knowledge is accessible and there are strong needs to be effective or make a social connection. The other side of the coin is dehumanization (Haslam, 2006; Leyens, Cortes, Demoulin, Dovidio, Fiske, & Gaunt et al., 2003), such that outgroup members and enemies are seen as animalistic (e.g., coarse, lacking moral sensibility) and machine-like (e.g., lacking emotional responsiveness).

**FOLK PSYCHOLOGY AND PERCEPTION OF INTENTIONALITY**

Folk psychology refers to the set of commonsense principles that laypersons employ to explain human behavior (Lillard, 1998; Malle, 1999). In line with Heider’s (1958) distinction between personal and impersonal causality, Malle (1999, 2004) suggested that intentional behavior tends to be described primarily in terms of the actor’s goals and motives (i.e., reasons). In contrast, unintentional behavior tends be explained in terms of causes (e.g., by the relatively mechanistic, internal vs. external causality distinction). Of course, the criteria that perceivers use to distinguish between intentional and unintentional behavior are crucial (Guglielmo & Malle, 2010; Malle & Knobe, 1997; Monroe & Reeder, 2011). When discussing intentionality, Aristotle (1941) described a beleaguered ship captain who threw cargo overboard during a storm in order to save the ship. According to the five criteria offered by Malle and Knobe (1997), the sea captain’s act will be judged intentional to the extent the following five criteria are met: The captain desired to save the ship, believed that throwing cargo overboard would accomplish that outcome, planned his actions, had the requisite skill or ability, and was aware of what he did. Notice, however, that employing these five criteria could involve a great deal of analytic thought. Perceivers may also infer intentionality in a more intuitive fashion, relying on simple heuristics (Morris & Mason, 2009; Reeder, 2009b). For example, jerky movements or verbal expressions of “Oops!” tend to imply unintentional behavior. Most inferences of intentionality are probably implicit (and possibly below awareness) in the
sense that perceivers simply assume, by default, that action is usually intentional (Rosset, 2008).

PERCEPTION OF MENTAL STATES SUCH AS GOALS AND MOTIVES

If intentional actions are explained by the actor’s goals and motives, how are such mental states inferred? Building on earlier insights (Jones & Davis, 1965; Heider, 1958), recent approaches to motive attribution fall into two categories: simulation and reliance on implicit theory (Ames, 2004; Epley & Waytz, 2010; Goldman, 2006; Reeder & Trafimow, 2005). When using simulation, people use themselves as a point of departure for judging others, essentially asking themselves, “What would my motives be if I were in the actor’s situation?” Immediate, intuitive reactions often follow this projection tendency, which may then be followed by a more controlled “adjustment” for additional information (Epley, Keysar, Van Boven, & Gilovich, 2004). When using implicit theory to predict motives, perceivers may rely on group stereotypes (e.g., People from the X political party have self-serving motives” or other types of abstract knowledge. But what determines which strategy is employed? It appears that when perceivers feel similar to the actor, they gravitate toward the simulation strategy (Ames, 2004; Mitchell, 2009).

Finally, perceived motives play an important role in recent models of dispositional inference (Kammrath et al., 2005; Read & Miller, 2005; Reeder, 2009a). According to the multiple inference model (MIM), perceivers process intentional behavior in the context of situational forces that give it meaning, allowing perceivers to infer one or more underlying motives (Reeder, 2009a, 2009b). In contrast to the view of the perceiver as a naïve dispositionist (Ross & Nisbett, 1991), MIM holds that situational information can be of great interest to the perceiver because it helps to identify the motives or reasons behind behavior. Perceivers then use information about motives to infer where the target stands on one or more trait dimensions.

In an early test of MIM, perceivers read about a soccer player who deliberately injured another player in response to different forms of situational encouragement (Reeder, Kumar, Hesson-McInnis, & Trafimow, 2002). When the aggression followed an earlier provocation (e.g., an insult from a member of the opposing team), perceivers inferred a revenge motive. But when the aggression was due to the exceptionally skillful play of the opposing player, perceivers inferred a selfish motive (i.e., wanting to win the game). In each case, the situation (provocation vs. skillful play by the opposing player) encouraged the aggression, yet the different situations suggested unique motives. In turn, these motives led to different sorts of trait inferences about morality. Perceivers saw the aggressor as higher in dispositional morality if he was motivated by revenge, rather than by selfishness. According to this model, perceivers can integrate a variety of types of information simultaneously—including motives and traits—to create a coherent impression (Asch, 1946; Kammrath et al., 2005; Read & Miller, 1993, 2005; Roese & Morris, 1999; Todd, Molden, Ham, & Vonk, 2011).

Application of Social Neuroscience to Attribution

Neuroscience represents what is, perhaps, the newest and most controversial tool available to attribution research—and social cognition more generally. Although recognizing its limitations, advocates of this new approach point to its unique strengths relative to traditional behavioral observation (Lieberman, 2010). For instance, multiple systems can be investigated simultaneously, without the potentially contaminating influence of instructional sets or manipulations of attentional focus. A skeptical reader might ask sarcastically, “How so, by pointing to brain regions where an attribution happens? If the anterior cingulate lights up, how does that explain the correspondence bias?!” But there is, in fact, a good case to be made for brain mapping (Iacoboni, 2009; Lieberman, 2010; Mitchell, 2009; Van Overwalle, 2009). Neuroscientists have made vast strides in isolating brain regions that perform particular mental tasks and computations. Some areas of the brain like the amygdala seem to be involved in emotional reactions, whereas areas of the prefrontal cortex are involved in complex problem solving. Armed with information about the typical cognitive tasks accomplished by a given brain area, attributions researchers can begin to test hypotheses of particular relevance to their own field. For example, if an attribution theory implies that a particular social inference task is cognitively taxing (rather than relatively automatic), the brain of perceivers who engage in that task should actually light up precisely in neural territory where other types of serious computation occur. In other words, brain outcomes can serve as another form of evidence to test a hypothesis (Amadio, 2008; Lieberman, 2010; Spunt, Satpute, & Lieberman, 2011). The discussion below will consider how neurological findings can address the question of whether automatic
processing and controlled processing rely on separate systems.

For more than 25 years, stage models of attribution have assumed there is a crucial distinction to be drawn between automatic (associative, intuitive) processing and controlled (deliberate, rule-based) processing (Gilbert et al., 1988; Trope, 1986). Yet others have challenged this assumption as unwarranted, suggesting instead that all processing is rule based, such that the same rules can underlie both types of processing (Kruglanski & Gigerenzer, 2011). So is it really necessary to think in terms of two different systems of information processing? Although a variety of perspectives should be considered, evidence from neuroscience can help answer this question.

In an important paper, Lieberman, Gaunt, Gilbert, and Trope (2002) described the neural substrates of an X-system and a C-system. The X-system uses pattern recognition to allow perceivers to form quick and easy impressions of an actor’s behavior and traits. This system is responsible for automatic social cognition and appears to involve a variety of brain areas, including the lateral temporal cortex and superior temporal sulcus. The system progressively analyzes visual information, identifying both what the stimulus is and where it is (Mason & Morris, 2010). Thus, it provides the basis for general semantic categorization and for what attribution researchers specifically refer to as behavior identification (Gilbert et al., 1988; Krull, 1993; Reeder, 1993; Trope, 1986). The neural perspective suggests that layers of analysis occur within (automatic) visual processing, implying that there are important differences in the ways visual and verbal information are handled (Morrish & Morris, 2009).

In contrast, the C-system relies on symbolic logic and reflective awareness (Lieberman et al., 2002). It reacts to the output of the X-system, particularly when there is a “lack of fit” between elements of information. For instance, the X-system might see evil and immorality at work when a participant in a psychological experiment is observed pressing a shock key to deliver a 450-volt shock to an innocent person (Milgram, 1974). But the C-system can “correct” that initial impression by recognizing that the participant was motivated to obey an experimenter who ordered the delivery of the shock (Reeder et al., 2008). The deliberate processing involved when dispositional inferences are adjusted to take account of situational factors appears localized in areas of the prefrontal cortex (Lieberman et al., 2002; Mason & Morris, 2010). This area of the brain is known to be involved in propositional reasoning and hypothesis formation. In short, separate areas of the brain produce the different computations attributed to automatic and controlled processing. These different neural substrates offer one more line of supporting evidence for stage models of attribution.

Final Thoughts

Attribution questions about social reality served as an important launching pad for social cognition in the late 1970s and early 1980s. The voluminous literature on attribution firmly established a tradition within social psychology of taking cognition seriously, allowing the field to more easily embrace a social cognition approach. Also, attributional questions, such as those concerning the automaticity of trait inferences (Winter & Uleman, 1984), inspired some of the earliest social cognition studies. In turn, the social cognition approach, with its emphasis on stages of processing, broadened the analysis of attribution—so much so, that it is now difficult to separate attribution research from social cognition. New attribution questions are likely to emerge in the future, continually expanding the borders of social cognition. It is equally clear, however, that the theory and methodology of social cognition will continue to shape the nature of such attribution research.

References


Attribution as a Gateway to Social Cognition


