

Actor–observer Bias

Authored by
mohammad looti

November 19, 2022

RECOMMENDED CITATION

mohammad looti (2022). *Actor–observer Bias*. PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=39394>

The hypothesis of an "actor-observer asymmetry" was first proposed by social psychologists, Jones and Nisbett in 1971. They hypothesized that "actors tend to attribute the causes of their behavior to stimuli inherent in the situation, while observers tend to attribute behavior to stable dispositions of the actor" (Jones & Nisbett, 1971, p. 93). Jones and Nisbett's (1971) hypothesis implies that a student who studies hard for an exam is likely to explain her own (the "actor"'s) intensive studying by referring to the upcoming difficult exam whereas other people (the "observers") are likely to explain her studying by referring to her dispositions such as being hardworking or ambitious.

The background to this hypothesis was social psychology's increasing interest in the cognitive mechanisms by which people make sense of their own and other people's behavior. This interest was instigated by Fritz Heider's book, *The Psychology of Interpersonal Relations*, and the research in its wake has become known as "attribution research" or "attribution theory."

Early Evidence and Reception

Soon after the publication of the actor-observer hypothesis, numerous research studies tested its validity, most notably the first such test by Nisbett, Caputo, Legant, and Marecek (1973). The authors found initial evidence for the hypothesis, and so did Storms (1973), who also examined one possible explanation of the hypothesis: that actors explain their behaviors by reference to the situation because they attend to the situation (not to their own behaviors) whereas observers explain the actor's behavior by reference to the actor's dispositions because they attend to the actor's behavior (not to the situation). Based largely on this initial supporting evidence, the confidence in the hypothesis became uniformly high. The asymmetry was described as "robust and quite general" (Jones, 1976, p. 304), "firmly established" (Watson, 1982, p. 698), and "an entrenched part of scientific psychology" (Robins, Spranca, & Mendelsohn, 1996, p. 376). Likewise, evidence for the asymmetry was considered to be "plentiful" (Fiske & Taylor, 1991, p. 73) and "pervasive" (Aronson, 2002, p. 168).

Recent Evidence

Despite the widespread belief in the actor-observer asymmetry, over 170 studies have been published since 1971 in which the hypothesis was put to further tests. Some of these studies examined the generality of the asymmetry, many conducted a secondary analysis of data gathered for somewhat different purposes. This entire literature was re-examined in a meta-analysis, which is a most robust way of documenting consistent patterns of evidence regarding a given hypothesis. The result of this analysis was stunning: across the 170 studies, the asymmetry practically did not exist. (The average effect sizes, computed in several accepted ways, ranged from $d = -0.016$ to $d = 0.095$; corrected for publication bias, the average effect size was 0.) Under circumscribed

conditions, it could sometimes be found, but under other conditions, the opposite was found. The conclusion was that the widely held assumption of an actor-observer asymmetry in attribution was false.

Theoretical Reformulation

The result of the meta-analysis implied that actors and observers explain behaviors exactly the same way. But all the tests of the classic hypothesis presupposed that people explain behavior by referring to "dispositional" vs. "situational" causes. This assumption turned out to be incorrect for the predominant class of events that people explain in real life (Malle & Knobe, 1997a): intentional behaviors (e.g., buying a new car, making a mean comment). People explain unintentional events in ways that the traditional disposition-situation framework can capture, but they explain intentional events by using very different concepts (Hilton, 2006?). An empirical theory of how people explain behavior was proposed and tested by Malle (1999, 2004), centering on the postulate that intentional behaviors are explained by reasons--the mental states (typically beliefs and desires) in light of which agents consider to act and on the grounds of which they do in fact decide to act (a postulate long discussed in the philosophy of action). The theory developed in more detail the psychological significance of (a) giving reason explanations (rather than its chief alternative, a "causal history of reason explanation"), (b) choosing either belief reasons or desire reasons, and (c) linguistically marking a belief reason with its mental state verb (e.g., "She thought that..."; "He assumes that..."). Empirical studies have so far supported this theoretical framework (CITE).

Within this framework, the actor-observer asymmetry was then reformulated as in fact consisting of three asymmetries: that actors offer more reason explanations (relative to causal history of reason explanations) than observers do; that actors offer more belief reasons (relative to desire reasons) than observers do; and that actors use fewer belief reason markers than observers do (CITE). Malle, Knobe, and Nelson (2007) tested these asymmetries across 9 studies and found consistent support for them. In the same studies they also tested the classic disposition-situation hypothesis and consistently found no support for them.

Thus, people do seem to explain their own actions differently from how they explain other people's actions. But these differences do not lie in a predominance of using "dispositional" vs. "situational" causes. Only when people's explanations are theoretically framed as reasons (and their specific features) do the differences emerge.

Implications

The choices of different explanations for intentional behavior (reasons, belief reasons, etc.) indicate particular psychological functions. Reasons, for example, appear to reflect (among other things) psychological distance. People increase reason explanations when they want to portray the

agent in a positive light (Malle et al., 2007) and reason use reflects ownership and affection for nonhuman agents (e.g., a pet fish) (Kiesler et al., 2007). Conversely, people decrease their reason explanations when explaining whole groups' behaviors, presumably because they can't be easily perceived as having "a mind" (O'Laughlin & Malle, 2002). Actor-observer asymmetries can therefore be seen as expressing a greater psychological distance people have to others' minds than to their own minds.

Related but Distinct Concepts

Actor-Observer "Bias"

Instead of speaking of a hypothesis of an actor-observer asymmetry, some textbooks and published research articles speak of an "actor-observer bias." The term "bias" is typically used to imply that one of the explainers--either the actor or the observer--is biased or incorrect in their causal attributions (within the framework of dispositional vs. situation causes). Interestingly, which one is incorrect is not clear from the literature. On the one hand, Ross's hypothesis of a fundamental attribution error suggests that observers are incorrect, because they show a general tendency to overemphasize dispositional explanations and underemphasize situational ones. On the other hand, Nisbett and Wilson (1975) argued that actors don't really know the true causes of their actions and often merely invent plausible explanations. Jones and Nisbett (1971) themselves did not commit to calling the hypothesized actor-observer asymmetry a bias or an error. Malle et al. (2007) also don't see actor-observer asymmetries as biases. Instead, both traditional and more recent theoretical positions consider asymmetries to be a result of multiple perceptual, attentional, and cognitive differences that fundamentally exist between actors and observers.

Self-Serving Bias

The actor-observer asymmetry is often confounded with the hypothesis of a self-serving bias in attribution -- the claim that people choose explanations in a strategic way so as to make themselves appear in a more positive light than others. The important difference between the two hypotheses is that the assumed actor-observer asymmetry is expected to hold for all events and behaviors (whether they are positive or negative); by contrast, the hypothesized self-serving bias is assumed to represent a complete reversal in actors' and observers' explanation tendencies as a function of positive vs. negative events. In traditional attribution terms, this means that for positive events (e.g., getting an A on an exam), actors will select explanations that refer to their own dispositions (e.g., "I am smart") whereas observers will select explanations that refer to the actor's situation (e.g., "The test was easy"); for negative events (e.g., receiving an F on the exam), however, actors will select explanations that refer to the situation (e.g., "The test was impossibly hard") whereas observers will select explanations that refer to the actor's dispositions (e.g., "She is

not smart enough").

ARABPSYCHOLOGY.COM