Study A1

Bertram F. Malle and Sarah E. Nelson

In this study we tested a relatively simple account of actor-observer asymmetries in behavior explanations—that observers and actors form explanations with different purposes in mind. Specifically, observers might be primarily interested in predicting the actor's future behavior (e.g., Heider, 1958), whereas actors might be primarily interested in making their own actions understood (Buss, 1978). As a result, observers try to find generally guiding information about the agent that explains not only the current behavior but also related behaviors in the future. When explaining intentional behavior, a search for CHR explanations and, among reasons, a preference for desire reasons, would serve this purpose best; when explaining unintentional behavior, a search for trait causes would serve it best. By contrast, an actor's attempt to make herself understood would favor reason explanations, especially belief reasons that reveal idiosyncratic information and thus render the action intelligible. If this logic is valid, then the actor-observer asymmetries should disappear (or drastically decrease) once both actors and observers provide explanations with the same purpose in mind. Study A1 tested this possibility.

Methods

Participants and procedure. Fifty-nine undergraduate students completed the experiment in exchange for credit toward an introductory psychology course requirement. Each participant was asked to explain up to six behaviors, roughly half from the actor perspective and half from the observer perspective. The behaviors were selected idiographically from a life event story that the participant had recorded on audiotape at the outset of the study. While the participant completed an unrelated filler task, the experimenter listened to the tape and attempted to select six intentional behaviors mentioned in the story, half of them performed by the participant and half by other people in the story. Then the behaviors were turned into why-questions (e.g., "Why did you get up at midnight?", "Why did your step mom tell you to come home?") and posed to the participant in a brief interview. The experimenter emphasized that the participant's answers would be tape-recorded and later played back to another person (the "audience").

Purpose manipulation. Instructions that manipulated the purpose of the explanations were randomly assigned across participants. In the *understand* condition (N = 14), the experimenter emphasized that "with your answers about each event, you need to make the

audience understand why the people involved in the event (you or someone else) acted or felt a certain way. So the audience will hear the question about a particular event in the story, then they'll hear your answer, and then they'll indicate how well they understood each relevant person in the story. That may sometimes be difficult for them, but your goal is to help them understand." In the *predict* condition ($N = 12$), the experimenter emphasized that "with your answers about each event, you need to help the audience predict what happens next in the story right after the event. So they will hear the question about a particular event in the story, then they'll hear your answer, and then they'll have to predict what happens next in the story. That may sometimes be difficult for them, but your goal is to help them predict, without giving it away in your words, what happens next in the story."

Two additional conditions were added after some participants reported that it was difficult for them to follow the instructions. In both new conditions, participants received a very brief instruction before each why-question. In the *past* condition ($N = 14$), they were asked: "You [or other person] did [behavior]. Think about that behavior. Keeping it in mind, now explain why you [other person] did X." In the *next* condition ($N = 15$), they were asked: "You [or other person] did [behavior]. Think about what you [other person] did next. Keeping this in mind, now explain why you [other person] did X." We assumed that the *past* and *next* instructions would be easily comprehensible approximations to the *understand* and *predict* conditions, respectively.

Coding. All explanations were transcribed from the audiotapes and classified for codability. We excluded 7% of behaviors that had ambiguous agents (primarily "we" statements) and 1% of answers that were claim backings. A small number of unintentional behaviors were eliminated from analyses because they would not have constituted a useful sample for analysis. The remaining codable explanations were classified according to the F.Ex coding scheme as in all other studies.

Analysis. Instruction was a four-level between-subject factor, broken down into three contrasts: complex vs. simple instructions (understand/predict vs. past/next), understand vs. predict, and past vs. next. Each participant provided both actor explanations and observer explanations, and the multiple behaviors that participants explained made it possible to analyze perspective as a within-subject factor. Degrees of freedom varied across analyses because some participants had missing values for certain explanation parameters (e.g., no intentional behavior explanations from the observer perspective).

Results

The results were based on a total of 841 explanations for 295 behaviors (276 intentional, 19 unintentional). Actor explanations comprised 55% of all explanations.

Folk-conceptual hypotheses. The Reason hypothesis was supported, as actors provided 2.43 reasons and 0.60 CHRs whereas observers provided 1.66 reasons and 1.10 CHRs per intention behavior, $F(1, 100) = 25.4$, $p < .001$, $d = 1.0$. However, this effect did not interact with any of the three instruction contrasts (ps > .30).

For behaviors explained by reasons, actors provided 1.74 belief reasons and 0.63 desire reasons whereas observers provided 1.10 belief reasons and 0.57 desire reasons, $F(1, 98) = 5.3$, $p < .05$, $d = 0.46$, supporting the belief hypothesis. Once more, the predicted effect did not interact with any of the three instruction contrasts (ps > .70).

For behaviors explained by belief reasons, actors offered 1.57 unmarked beliefs and 0.56 marked beliefs whereas observers offered 1.05 unmarked beliefs and 0.52 marked beliefs, $F(1, 87) = 2.8$, $p < .10$, $d = 0.37$, supporting the marker hypothesis. Again, the effect did not interact with any of the instruction contrasts (ps > .36).

Traditional hypotheses. The person-situation hypothesis was tested on causal history of reason explanations and reasons contents. In the first case, actors offered 1.02 person CHRs and and 0.28 situation CHRs whereas observers offered 0.91 person CHRs and 0.52 situation CHRs. This pattern was not significant, $F(1, 70) = 1.6$, $p = .21$, and went counter to the hypothesis, $d = -0.29$. In the second case, actors offered 1.12 reasons with person content and 0.86 reasons with situation content whereas observers offered 0.72 reasons with person content and 1.03 reasons with situation content. This was a significant effect counter to the hypothesis, $F(1,98) = 4.8$, $p < .05$, $d = -0.44$. No interaction with any instruction contrast emerged.

The trait hypothesis was tested on person factors in causal history explanations and was not supported, whereby actors used 0.34 traits and 0.85 nontraits wheras observers used 0.46 traits and 0.95 nontraits, $F(1, 53) < 1$, $d = -0.01$.

Study A2

Sarah E. Nelson and Bertram F. Malle

This study, a shortened and updated version of a first-year project paper submitted to the University of Oregon Psychology Department by Sarah E. Nelson in 1999, was originally designed to test the sensitivity of behavior explanations to individual differences in personal agency beliefs and manipulated agency mindsets (portrayals of human action as "free" vs. "causally determined"). These factors did not have a noteworthy impact on explanations, but the explanation task allowed actor-observer comparisons of intentional action explanations.

Method

Participants. Participants were 96 college students enrolled in psychology courses at the University of Oregon who participated as part of a course requirement.

Procedure and Materials. This study originally consisted of three parts that were treated as separate studies with separate consent forms. Only the third part is of direct relevance here. In that part, participants were asked to remember specific positive and negative intentional actions they or someone else performed and to explain those actions. (In the first part, participants filled out several attitude measures, including one about personal agency. In the second part, they were assigned to read a passage that either argued for causal determinism or free will as the fundamental principle underlying human action. Very few effects of these two variables were found in the data.)

The primary measure requested a set of free-response explanations. The questionnaire consisted of four questions that asked participants to remember times they or someone else performed specific actions and explain why those actions were performed. (See Table 1.) Each subject explained one positive and one negative action in the actor role and one positive and one negative action in the observer role. Thus, role (actor, observer) and valence (positive, negative) were within-subject factors. Order of questions and pairing of perspective with specific behaviors was counterbalanced across participants.

Coding of explanations. The explanations were coded according to the F.Ex coding scheme (Malle, 1998/2006), which reliably detects a variety of modes of explanation and types within each mode (Malle, 1999). Two coders classified the explanations. A small number of explanations that falsely construed the target actions as unintentional were eliminated, and

Table 1. *Actions that participants were asked to remember and explain*

Actor	Observer
Positive	
Remember a time you helped someone.	Remember a time someone helped you.
Remember a time you gave somebody something (not on a special occasion).	Remember a time someone gave you something (not on a special occasion).
Negative	
Remember a time you intentionally took something that wasn't yours (without the other person's permission).	Remember a time someone intentionally took something of yours (without your permission).
Remember a time you intentionally hurt someone (physically or mentally).	Remember a time someone intentionally hurt you (physically or mentally).

so were statements with no explanatory content (e.g., mere evaluations). Reliability was calculated based on the coders' initial responses, and disagreements were settled through discussion. Overall agreement for whether a behavior was codable or not was 90%. Comparison of codings revealed 93% agreement ($\kappa = .80$) on explanation types (whether an explanation was a reason or a causal history), 98% agreement ($\kappa = .94$) on reason type (desire, belief, or valuing), 98% agreement ($\kappa = .96$) on mental state markers, 94% agreement ($\kappa = .88$) on causal history type (person or situation), and 94% agreement ($\kappa = .90$) on trait vs. nontrait.

Results

In several within-subject analyses of variance, the explanation parameters of interest were examined as a function of actor-observer role. In all cases, the dependent variable was the mean number of explanations given per subject or a proportion of explanations given per subject .

Reasons and causal histories. The predicted reason asymmetry was found, $F(1,95) = 21.4$, $p < .001$, $d = 0.66$. Participants gave more reasons as actors ($M = 1.57$) than as observers ($M = 1.17$), but they gave more causal histories as observers ($M = .60$) than as actors ($M = .39$).

Beliefs and desires. Evidence for the belief asymmetry was present but weaker than in other studies, $F(1,90) = 2.1$, $p = .15$, $d = .22$. Actors offered more beliefs ($M = 1.16$) than observers did ($M = 0.91$) but not fewer desires ($Ms = 0.42$ and 0.39 for actors and observers, respectively). This asymmetry was the only one that interacted with valence such that for

positive actions, actors and observers differed substantially and significantly but for negative actions they did not. We are currently testing the replicability of this effect.

Marked and unmarked beliefs. The marker asymmetry held reliably and right at the average effect size of previous studies, $F(1,70) = 5.5$, $p < .05$, $d = 0.41$.

Traditional asymmetries. Because people explained only intentional actions, the persons-situation hypothesis was tested for causal history of reason explanations and for the content of reasons. However, neither of the two tests showed significant person-situation differences (both ds = 0.02). The trait hypothesis was tested for traits in person factors of causal history explanations and showed a strong effect, with actors offering 0.13 traits against 1.06 nontraits and observers offering 0.53 traits against 0.72 nontraits, $F(1,15) = 4.1$, $p = .06$, $d = 0.77$. It should be noted, however, that this results is based on only 16 participants and limited to causal history explanations.