AUTHORS’ RESPONSE

Authors’ Response to Commentaries

John Sabini and Michael Siepmann
Department of Psychology
University of Pennsylvania

Julia Stein
Department of Psychology
Bryn Mawr College

Mea Culpas

First the mea culpas. Hogan (this issue) rightly claims that we claim that the notion of dispositions as employed by attribution theorists cannot do all that the idea is needed to do. But, he claims, we gave the idea of situations a free ride. Mea culpa. He is quite right. (Funder [this issue], too, takes us to task about this.) For example, Gilbert and Malone (1995) treated the “situation” member of the disposition–situation couple as everything on the outside of the epidermis. Will that do?

Let’s take a lecture hall with a class in it. Is this the same situation for the lecturer and for the students? Is it the same situation for the student there to learn the material and for the one there to flirt with a neighboring student? What about for the painters waiting to get in to paint the ceiling? What about for the person in the booth trying to get the slides to run in synchrony with the lecture? What about for the person from another culture who has no idea what is going on? The point is that any description of the situation that has any hope of explaining behavior at all will have to make reference to the goals, abilities, and beliefs of the various people in the physical location. A Skinnerian description of the situation from which all such psychological material has been bleached is not likely to predict anything at all.

Underweighing Embarrassment

Another mea culpa we owe is for not making it clearer that our account of people’s underweighing embarrassment was not intended to explain all of social psychology or to cover all of the interesting counterintuitive facts that social psychologists have discovered. We do, however, believe that the experiments we mentioned are, as we said, on everybody’s list of the classics; but we also believe, as we also said, that they do not exhaust anybody’s list.

Ego-syntonic–Ego-dystonic

Funder, Fein (this issue), and Gilovich and Eibach (this issue) take us to task over the distinction between ego-dystonic and ego-syntonic. Well, mea maxima culpa here. Sabini and Silver (1987) raised the question of whether the internal–external distinction is a dimension or a congeries. That is, is there a single distinction that underlies ordinary people’s use of this pair of terms, or is there a heap of different distinctions that underlie it? They concluded the latter, and they suggested that the various distinctions had no more in common than that they could all be evoked (in the right context) by the internal–external metaphor. They believed that ego-dystonic versus ego-syntonic is one of the distinctions in that heap. They did not believe, and we do not believe, that it covers all of the cases that the commonsense actor would use the internal–external terms for, but we do believe it has a use.

The use it has, we suggested, is to explain why people have the sense that certain behaviors are strongly externally controlled, although those behaviors are a product of the actor’s choices. Drug addicts who really want to give up their drugs, for example, might see their drug taking as externally rather than internally controlled. But in what sense? Surely not because something outside the skin rather than something inside the skin is in control. Surely both the drugs and the addict’s urge for the drugs matter. Surely their behavior is not under external control because everyone takes drugs, as the statistical conception of the
dispositional–situational distinction would have it. Well, in what sense then? We suggested that the drug addict who wants to quit finds that his taking the drug is beyond his will, control, or self. And it is in that sense that he sees his behavior as externally controlled. We did not intend this ego-syntonic–ego-dystonic distinction to cover all cases in which people partition causes into internal and external; rather, we just meant to articulate the internal–external intuition in cases like addiction, and, as it turns out, cases like the social influence experiments, where one finds that one cannot do what one wants to do. In any event, insofar as we made it seem as if we were offering a single analysis meant to cover all cases of the internal–external distinction, we apologize.

Embarrassment

Miller’s (this issue) critique takes us to task on two grounds. First, we did not place our work in sufficient historical context. Mea culpa. The second is that we were reluctant to take a stand on just what embarrassment is. We were indeed reluctant to do so, mostly out of fear of overburdening the reader. We are grateful to Miller for filling this gap with his view of the current state of the literature and, especially, of his own important contributions to it. The reader interested in our view of what embarrassment is and our recent writing on the topic might have a look at Milgram and Sabini (1978); Sabini, Garvey, and Hall (2001); Sabini, Cosmas, Siepmann, and Stein (1999); Sabini, Siepmann, Stein, and Meyerowitz, (2000); and Sabini and Silver (1998a).

Caring what others think of you. Both Miller’s and Leary’s (this issue) critiques urge us to go deeper than claiming that the social influence studies depend on a fear of embarrassment. They urge on us the claim that these studies reveal how much we care about other people’s opinions of us. After all, they claim, is not a concern with what others think of us at the heart of embarrassment? We are, really, quite sympathetic to this move.

Emotions, we agree, rest, by and large, on cognitions and desires; indeed, it might be argued that emotions are nothing but cognitions and desires (see Sabini and Silver, 1998b, for just such an argument). So why attribute behavior in the classic social influence studies to fear of the emotion of embarrassment rather than to the motive at embarrassment’s heart: the desire to have others think well of you? Well, we had two reasons.

First, we are not as convinced as Miller and Leary are that embarrassment always has to do with a concern with others’ opinions of you. Silver, Sabini, and Parrott (1987) developed a conception of embarrassment, a dramaturgic theory derived from Goffman’s (1955), that did not implicate a concern with others’ evaluations at the heart of embarrassment. Thus, we were not willing to take a stand implicating concerns for others’ evaluations in all tokens of embarrassment. But there is a perhaps more interesting reason why we did not claim that it was a concern for others’ evaluations that lies at the heart of these experiments.

To be sure, we all care a lot about what some people think of us; it is not uncommon, for example, for graduate students to be very concerned with their advisors’ opinions of them. And perhaps we care a little about what everybody thinks of us. But we know these facts about ourselves, and participants’ reactions in the “classic” studies are really out of proportion to the degree to which the participants care about their witnesses’ opinions of them. So these studies seem to show that we care a lot more than we think we do about strangers’ views of us. But do they?!

Imagine that in the Milgram (1974) obedience experiments, participants were whisked out of the experimental room and offered the following choice: You can have the experiment continue with your delivering another 450 volt shock to the learner, or you can stop now. If you choose the latter, the experimenter will think ill of you, of course, but you will not see him again—you can leave by this back door—and the experiment will stop. What do you suppose would happen? We suggest that obedience would be dramatically reduced, as it is in the experimenter-absent version of the study, in which the experimenter delivers his orders over the phone. This suggests that it is not the negative evaluation by the experimenter that the subjects fear, it is the embarrassment that such an evaluation would produce in face-to-face interaction. More generally, we propose, if you find a way to detach evaluation by strangers from the embarrassment the evaluation provokes, you will find that people care very little about the evaluation and a lot about the embarrassment. This, we suggest, is less true of cases in which people care (and know they care) about another’s opinion of them, for example, their mentors’.

1 An evolutionary speculation: Perhaps embarrassment evolved because those prone to it avoided doing things that triggered it. Perhaps avoiding doing things that trigger embarrassment is in our best interest in the long run. But perhaps this emotional system evolved in a social world without strangers, that is, people with whom one neither has nor expects to have a relationship. Because there were no strangers in this social world, it did not pay for the embarrassment system to distinguish strangers from people whose opinions we care about; embarrassment was triggered regardless of the audience. But now urbanites, at least, live in a world of stranger—people’s whose opinions they do not care much about, such as Milgram’s (1974) experimenter—with an embarrassment system evolved for a world without strangers.
What We Do Disagree About: Difficult Empirical Questions and Conceptual Confusions

Gilovich and Eibach (this issue) have this to say about the person–situation distinction:

It is not possible, clearly, for an external stimulus to influence a person’s chosen actions if it does not resonate with some faculty or disposition of the person. Thus, a precise accounting of how much a given action stems from the impinging stimulus rather than from the faculty or disposition with which it makes contact is wrenchingly difficult and often impossible.

We agree, of course, with the first sentence. But with their second sentence we do not agree. We insist that it is not wrenchingly difficult to give an accounting—precise or otherwise—of how much John’s going out with Sue stems from her beauty rather than from his love of beautiful women. The problem with this question is not that it is difficult to answer; it is that it is conceptually incoherent.2 It is incoherent because it construes two classes of accounts that are in fact complementary as if they were competing. The heart of our argument is that one must take this point seriously. All behavior is jointly a product of environmental stimuli and dispositions.

Because we take that point seriously, what are we to make of Gilovich and Eibach’s claim that the fundamental attributional error (FAE) is “the contention that the layperson’s intuitions give more weight to dispositions and less weight to situational influence than what psychologists have learned—however uncertainly—is warranted”? We reject it as part of the same confusion of complementary with competing accounts. It is true, and people know it is true, that behavior is, as Lewin (1935) said, a joint function of the person and the environment. It matters not how one writes the equation: each and every behavior is a joint product of a person and a situation.

But have not social psychologists shown that details of situations are more important than laypeople think? Well, as we said in our article, yes and no. They have shown that some details of some situations are more important than people think relative to other details of those situations. For example, in the Milgram (1974) experiments on obedience, they show that the moral aspect of the situation is vastly less important than laypeople think. Because it is conceptually true that the degree to which an aspect of the situation influences behavior is exactly the same as the degree to which the disposition with which the aspect fits (the faculty or disposition with which it makes contact) influences behavior, it is conceptually impossible to show that situations (or aspects of situations) are more important than dispositions. We suggest that what social psychologists have actually found out is not the broad and general fact that situations are more important than dispositions, but that there is a not-all-that-tidy collection of aspects of situations that are more important than we think they are (and that, by the way, morality and rationality are, regrettably, less potent than we think they are).

Ross’s Challenge

Let us now face Ross’s (this issue) challenge to us squarely. First, we reject as confused any attempt to pit situation and disposition against one another as causal influences. We repeat that the only way to pit one against the other coherently (the statistical way) makes the classic experiments we have discussed into dispositionalist, not situationalist, classics, as Funder also noted. We challenge Ross to provide a coherent analysis of the distinction between situational accounts and dispositional ones in which these are situationalist classics.

Finding Causes Versus Making Predictions

As Gilovich and Eibach note, attribution theorists long ago turned their attention from how commonsense actors understand the causes of behavior to how commonsense actors predict other people’s behavior. As Ross suggests, this is in part because it is easier to determine whether predictions are correct than whether causal accounts are correct. But could examining predictions rather than causes be an example of looking where the light is better than where the keys are? We think it just might be. We also think Gilovich and Eibach will help us decide, but before we get to that issue, we must address a claim Ross makes about causal accounts.

Is it ever possible to assess a causal account? Ross reminded us that he and Nisbett “noted that the accuracy of the causal account offered for any particular behavior can never be evaluated.” We see no reason for such blanket despair. For example, we have no trouble evaluating an account of people’s deafness as being a consequence of their transgressions in a previous life. It is true, of course, that we cannot know with “perfect certainty” that such an account is wrong, but the previous century taught us that we cannot know “anything” with perfect certainty. But we can surely be as confi--

---

1 Wittgenstein would claim that this is a conceptual confusion masquerading as a difficult empirical question.
dent about the falsity of that claim as we can about almost anything people lay claim to knowledge about.

And we note that Nisbett himself (Nisbett & Wilson, 1977) has been willing to judge the accuracy of people’s causal attributions. In this famous experiment, Nisbett and Wilson asked participants to select pantyhose. Unbeknownst to the pantyhose pickers, the pantyhose were identical. As it happened, as a byproduct of the procedure, the participants tended to select the pair on one end. But when asked why they picked the pair they picked, they said it was because of the quality of the hose, not because the pair was on the end. Nisbett and Wilson dismissed their accounts as inaccurate. If this causal account can be tested for accuracy, we see no principled reasons why others cannot be.

People Are Interested in Causes

Although we agree with Ross and with Gilovich and Eibach that actors are sometimes concerned with drawing inferences and with predicting future behavior, we believe that they often want to know which of the actor’s dispositions (and abilities, and so on) the actor’s current behavior revealed. This, after all, is the question with which Jones and Davis (1965) began.

Imagine that John gives money to a shabbily dressed woman in the presence of his date. Let us imagine that an observer (rightly or wrongly) believes that (a) John believed that giving money to the shabbily dressed woman would impress his date, (b) most people in John’s situation would want to impress their dates, and (c) most people would give money to the woman even if they did not want to impress their dates. Because the observer believes all of this, the observer knows, before the inquiry begins, that a situational (in Ross’s sense) attribution will be in order. But surely, even having decided that, there is another question, namely, which of John’s dispositions (or aspects of John’s situation) led to the giving of the money? That is, did John give out of generosity, or in order to impress his date, or (as Reeder [this issue] reminds us) both? An observer might be interested in this question even though both answers lead to a situational account and, therefore, no updating of the observer’s prediction of what John will do in the future relative to the typical person. (Ross’s unpacking of the situation–disposition distinction allows one to account for the behavior with a situational attribution while still tracing it back to one of the actor’s dispositions, though that is not to say that one gives it a dispositional explanation; even on Ross’s analysis of the distinction, it turns out, situations and dispositions are not in as intense competition as one might imagine.) With Heider (1958), however, we believe that observers are drawn toward trying to understand which disposition of the actor led to a particular behavior in at least some cases in which the behavior is common.

Ross is right in claiming that in some—perhaps most—contexts, to say that John gave money out of generosity is to warrant the inference that John is more generous than most. But this is not true in all contexts. When, à la Jones and Davis (1965), two motives for an act occur to us and we resolve our uncertainty in favor of one by saying, “Ah ha, it was his generosity that led to it,” we are not warranting the inference that John is more generous than most; rather, we are commenting on why John did this particular thing. And, as Fein’s important line of work reminds us (Fein, 1966; Fein, Hilton, & Miller, 1990), people do seem to think about these things when ulterior motives are in the air. Why would people be interested in deciding which of a person’s dispositions (or which aspect of the situation) resulted in an act, even when they believe the act to be common and, therefore, know that tracing the act back to its causes will not effect their predictions about the actor (compared to the typical actor)? There are several answers.

Causes and the Nature of the Act

Famously, person perception is supposed to work as follows (see, e.g., Gilbert, 1998; Gilbert & Malone, 1995): People witness an act, say, John’s giving the money. They categorize the act, say, as generous. Then they characterize the actor in the same terms, as a generous person. That is, they make a correspondent inference. But then, if they have time, energy, and motivation, they think about it some more and, perhaps, notice a gun. Then they withdraw the characterization of the person. And this is where the story usually ends. But note that if the money was given up at gunpoint, the act was no more an act of generosity than the actor was a generous person in giving the money. The observer should, if endowed with interest, time, and energy, recategorize the act, too, from a generous act to a prudent act. The point here is that at least some categorizations of acts in ordinary English (and, we suspect, but do not know, in other natural languages) depend on the mental states of the actors performing the acts. So an observer might be concerned to know about the mental states of an actor performing an act (dispositions) not in order to predict future behavior (or not just in order to predict future behavior), but because the observer, quite literally, is concerned to know what just happened (cf. Heider’s, 1958, very similar analysis). Indeed, this way of thinking leads

---

3 Although, as is typical of anchoring and adjusting, they do not adjust their attributions enough.
one to the idea that a steady-state correspondence between characterizations of acts and characterizations of actors may not be a bias at all but, rather, may be a semantic requirement of the language. The problem with this line of thought, however, is what to do about acts which flowed from generosity but of an ordinary degree? Or what to do with evil acts that are common? Should those get correspondent inferences or not? Well, what do people do?

Evidence comes from a study by Miller, Gillen, Schenker, and Radlove (1974). In this study (Experiment 2), some participants were told about (and others were not) the Milgram (1974) studies of obedience and their results. They were then asked to predict what they and others would do in the experiment, and they were asked to make trait attributions on a set of evaluative traits about people who obeyed in the Milgram experiment. The results were that knowing the Milgram results affected participants’ predictions of what people would do but not their trait attributions toward those who obeyed. These results suggest that when confronted with common but extreme behavior, people are sensitive to the predictive implications of the act’s prevalence, but they still tend toward correspondent trait inferences; that is, they think the behavior is appalling in various ways, and they think the actor is too. So, as it were, people fall into correspondence bias, even while adjusting their predictions. But is this a bias? Are they wrong in making these trait ascriptions?

Sabini and Silver (1982) argued that there is a set of terms that characterize an actor, terms that obey the logic of “if you did the act, you earned the label.” Consider murderer, thief, adulterer, child molester, hero, and so on. These are all terms that seem to have to do with what a person has done rather than with what a person will do. As Sabini and Silver pointed out, if a person murders someone and in the process becomes a quadriplegic, that person is a murderer. But in calling him such, one is certainly not predicting future murders. To be sure, one would be inclined in general to predict that those who have murdered are more likely to do so again, so there is a coupling between characterizing a person as a murderer and predicting that the person will murder, but we believe that our example shows that having murdered is the sufficient condition for the application of the term; prediction is inessential.

Of course murderer, thief, and so on are not trait terms. Nonetheless, many trait terms are heavily saturated with moral content. And we are suggesting that to the degree that a trait term is saturated with moral content, the appropriate rule for its application might be: A person is an X if and only if the person has performed actions X. Reeder and Brewer (1979) made a similar suggestion. They suggested that some traits might be ascribed to people based on the most extreme act of the relevant type those people had ever performed, and they singled out moral traits as candidates for this sort of treatment. The prevalence of misbehavior, they suggested, might affect how wrong we thought a behavior was, but, they suggested, prevalence would not affect ascriptions to people of traits based on those behaviors. They too, in other words, argued that attributions should be correspondent. But why would people do this? Why would they label people based on single acts if such labeling does not lead to more accurate predictions? Maybe prediction is not the only thing that is important. Let us offer some suggestions to make this idea more palatable.

Control of Behavior

Predicting people’s behavior is nice, but controlling it is better. It would not be crazy, for example, for a university to punish students for cheating, despite recognizing that cheating might be common. Indeed, the more common cheating is, the more important it is to punish it. And one important punishment is the mere fact of being labeled publicly as dishonest, a cheat, irresponsible, immoral. The public labeling of someone as something also recruits the community in punishing the misbehavior. Of course, praise works this way too. True, giving money to the hungry woman out of generosity does not show that one is Mother Teresa, a more generous person than most, but it is nonetheless worthy of praise or, at least, of thanks. It is certainly more worthy of praise than is giving money to someone to impress one’s date. Finding which dispositions of an actor were revealed in the act we just saw may be important, then, even if the search does not update our predictions, because we may need to know which disposition was engaged not only in order to know what just happened but also in order to know how to react. And these evaluative reactions—praise for generous acts (common or not) and censure for cruel acts (common or not)—constitute a web of social control.

It is worth noting in this context Axelrod’s (Axelrod & Dion, 1988; Axelrod & Hamilton, 1981) simulations of the Prisoner’s Dilemma Game. The strategy that prospered in that game, “Tit for Tat,” is one that made absolutely no effort to predict its opponent’s future behavior. Other (losing) strategies did that. “Tit for Tat” remained firmly rooted in the past. Perhaps social life is a bit like Axelrod’s simulations. Perhaps we are the sort of creature that returns generosity with generosity, kindness with kindness, meanness with meanness, criticism with criticism. (Oh, heavens, that could not be true!) To be this kind of species, we need to know what happened, we need to know which dispositions have just
been revealed, but we do not need to make any predictions whatsoever.

Social Psychology’s Humanizing Message

We have been making the case that although the light is bright around predictions, the keys might not be there, or at least not all of the keys. There is one last argument in favor of the keys’ being elsewhere, one made by Gilovich and Eibach. They rightly point out that attribution theory should not be about what happens in labs but about what happens in the rest of the world. And they suggest that attributional issues are important when people inappropriately blame the victim.

First, we point out that there are reasons people blame victims that seem to have little to do with the FAE. Self-serving and defensive motives come into play here. And, indeed, the overly broad message that situations, not dispositions, cause behavior seems to erode responsibility for behavior. This message lets people off the hook for what is their fault (as well as for what is not their fault) and denies them praise for what they should be praised for. If claiming that situations are more important than dispositions lets the innocent off the hook, it does so by a blanket denial of human responsibility, and that is dehumanizing, not humanizing. But let us follow Gilovich and Eibach’s example to see more clearly how attributional issues are involved in blaming the victim.

Attributions and blaming the victim. Gilovich and Eibach ask us to imagine a single working parent who is criticized for not working enough. We need to flesh out this story a bit to get to its attributional core. Let us assume that she works part-time in order to be there when her kids get out of school. (Perhaps this is not the case Gilovich and Eibach had in mind; we do not know. If not, we apologize.) And maybe there are people who would claim that she works only part-time out of laziness. Gilovich and Eibach argue that people who characterize her as lazy have failed to consider the situation from her point of view. They have failed to take into account that it is her need to take care of her children that causes her to work only part-time. Now, what is at issue here? What kind of dispute is this?

Hard and soft constraints. First, we must clear away a red herring. Gilovich and Eibach refer to her need to take care of her children as a “constraint,” but that blurs an important distinction. Goffman (1971; and we agree with Gilovich and Eibach that his work is insufficiently integrated into social psychology), follow-

4 There is a regrettable tendency in the literature to conflate these two kinds of constraints. If hard constraints exist, then behavior must indeed be utterly discounted—nothing about character is revealed. If soft constraints exist, then something about character is revealed. The attributional literature speaks of a debate coach’s instruction to write an essay as a constraint. And, indeed, it is—a soft one. Fein, in his commentary, argues that subjects in the original Jones and Harris (1967) study do not discount their attributions in light of what we would call the soft constraint. But he says they do not discount enough. How much should they discount? The notion of a constraint has been even more loosely used. In the famous Ross, Amabile, and Steinmetz (1977) quizmaster experiment, the actual empirical result was that the person assigned the role of answering the questions was seen as about as smart as the average Stanford student by all concerned in the experiment in each condition, but the person asking the questions was seen by the contestant and by observers (although not by themselves) as smarter than the average Stanford student. It is this inflation of how much the person asking knows that is the error—if there is one—in this experiment. Gilbert and Malone (1995) characterize these results as being about subjects’ failing to notice constraints.
ior? Are the single working mother and her critic engaged in a conflict of predictions? Predictions of, for example, whether she will work longer hours after her kids have grown up are certainly not the main focus of the mother’s and critic’s disagreement, if they are concerned with them at all.

Is the single working mother pleading a hard constraint? Is she offering an excuse for not working more? Is she denying that any aspect of her character is revealed in her behavior? We think not. Instead, she is proposing that in working only part-time so that she can be home with her kids when school is out, she is displaying her character, specifically, her devotion to her kids. Her critic, of course, agrees that her character is on display, specifically, her love of her own comfort, that is, her laziness.

We agree with Gilovich and Eibach, then, that this confrontation between the working mother and her critic is where the keys are; it is, ultimately, what attribution theory needs to articulate and explain. But we deny that it has anything directly to do with prediction. And we suggest that it is not about what most people would do—whether the single mother is uniquely or commonly lazy, or uniquely or commonly devoted to her kids. We suggest, instead, that what it is about is the question that Heider (1958) said consumed the native observer, namely, why did the actor do what she just did? It is not about whether the actor’s character was revealed in the act; it is about which aspect of the actor’s character was revealed in the act.

Summary

We once again apologize to our critics and readers for not making it clear that we did not intend “people underweigh embarrassment” as an account of all of social psychology. And we apologize for not making it clearer that we did not intend ego-syntonic versus ego-dystonic as an account of all meanings of the internal–external distinction. But we continue to insist that the notion that behavior is more (or less) a function of the situation rather than of the person is incoherent. And we also suggest that finding causes and making predictions are not the same thing. We believe that actors are often eager to find causes in order to know what just happened and in order to know how to act. They need to know causes not because (or not only because) causes lead to predictions, but because causes are sometimes needed for the correct categorization of behavior. Also, insofar as people play “Tit for Tat,” they need to know what just happened, but they do not need to predict their opponent’s future behavior.


