

# The “Fundamental Attribution Error” is rational in an uncertain world.

Drew E. Walker (dehoffma@ucsd.edu), Kevin A. Smith (k2smith@ucsd.edu), Edward Vul (evul@ucsd.edu)  
Department of Psychology, University of California San Diego, La Jolla, CA 92093 USA

## Abstract

Others’ internal qualities (e.g. dispositions, attitudes) are not directly observable so we must infer them from behavior. Classic attribution theories agree that we consider both internal qualities and situational pressure when making these judgments. However, one of the most well known ideas in psychology is that social judgments are biased, and we tend to underestimate the pressure that situations exert and overestimate the influence of disposition (known as the *Fundamental Attribution Error*). We propose that the social judgments made in classic studies of attribution have been interpreted as biased only because they have been compared to an inappropriate benchmark of rationality predicated on the assumption of deterministic dispositions and situations. We show that these results are actually consistent with the behavior of a simple ideal Bayesian observer who must reason about uncertain and probabilistic influences of situations and dispositions.

**Keywords:** Social Inference; Bayesian Inference; Attribution Theory; Fundamental Attribution Error

## Introduction

To navigate our social world we must predict how others will behave, and how we should act around them. Since we cannot directly observe the internal qualities that motivate others, we must infer them from their behavior. Imagine you see someone drop money in a donation jar when entering a donation-funded museum. Do you conclude that she is a generous person or just succumbing to the pressure imposed by a watchful docent? In all such cases, it is not only internal qualities (generosity) – but also external circumstances (docent’s attention) – that influence behavior; thus attributing a behavior to a situational or a dispositional cause is an underdetermined problem.

An extensive literature suggests that there are systematic discrepancies between the social inferences people *should* make, and the inferences they *do* make. A considerable number of behavioral experiments using rich social situations have concluded that we have a tendency to disregard circumstantial pressures, and instead overestimate the role of disposition. That is, when we witness someone drop money in the museum donation jar we are prone to think that she is a generous person, and not properly consider the external pressure the docent is exerting. This phenomenon, known as the *Fundamental Attribution Error (FAE)*, is one of the most prominent concepts in social psychology, has spawned numerous theoretical explanations (for review see Gilbert & Malone, 1995), and referenced in popular culture (e.g. Gladwell, 2000). In the classic demonstration of the FAE, university students read an essay,

ostensibly written by a classmate, which either favored or opposed Fidel Castro (Jones & Harris, 1967). Even when told that the opinion expressed in the essay had been randomly assigned by a course instructor, readers still thought that the author held the view expressed in the essay. A large number of studies have since produced similar results, yielding a net assessment in the literature that people are “lay dispositionalists” (Ross & Nisbett, 1991) wired to neglect the influence of situations and instead attribute actions to stable internal qualities.

The inferences observers make in this and similar experimental paradigms typically are considered to be logically unwarranted. The classic normative models reason that when you witness an outcome, and then learn that it was caused by one event, it is inappropriate to also attribute it to a second event (e.g. Kelly, 1973; Jones & Davis, 1965). For example, if you plug your cell phone in and the battery doesn’t charge, there are two likely explanations: there is something wrong with your device, or the outlet doesn’t work. If you learn that the outlet is dead there is now no reason to worry that your phone is broken. According to this deterministic logic, if a course instructor tells someone they must write an essay in favor of Castro, it is irrational to take a pro-Castro essay as evidence that the author really favors Castro.

Most isolated situations, however, are not so powerful that they completely constrain behavior. In daily life people rarely encounter situations that are so extreme that everyone acts uniformly within them. Even when society takes great care to make behavior as constrained by the situation as possible (e.g. locking someone in a jail cell), these situations are still not *totally* deterministic (people still escape from jail). Outside of such extremes, situational pressures are far from deterministic, but rather act in conjunction with internal qualities to produce behavior. Therefore, even when a situation is presumed to be influential, it is not reasonable for people to assume that its influence is so extreme that it is sufficient to determine behavior. Situations akin to those used in FAE tasks are certainly far from deterministic. For example, Sherman (1980) asked university students to write an essay supporting a controversial school policy and less than 70 percent of students complied. Since situations are rarely (if ever) entirely constraining, we should not expect subjects in the classic FAE studies to attribute cause as if they are. How can we reason about peoples’ dispositions in non-constraining situations? For this we turn to the formalism of causal attribution in probabilistic inference.

Recent work suggests that human causal learning and inference can be explained within a rational probabilistic inference framework across a number of domains. This work proposes that we can form a causal model of the

world, and condition on our observations to determine what might have been true of the world to yield the observations we have seen. This reasoning framework can account for a wide range of the causal inductions made by humans (Holyoak & Cheng, 2011), including social inferences, such as inferring goals from the movement of simple animated agents (e.g. Baker, Saxe, & Tenenbaum, 2009). We propose that the social inferences made in classic FAE studies can also be accounted for within this framework, and that given uncertainty about the strength of situational variables, human tendencies to make dispositional inferences in these paradigms are quite sensible, and do not reflect a bias.

In the past there have been scattered proposals that the social inference process could be accounted for in terms of Bayesian inference (Azjen & Fishbein, 1975; Morris & Larrick, 1995). However, these proposals predated modern computational statistical methods (Griffiths & Tenenbaum 2005, 2009; Kemp & Tenenbaum 2009), and could offer only verbal descriptions of conceptual possibilities<sup>1</sup> without being able to make concrete predictions about the variety of manipulations in the decades of established FAE literature.

In this paper we first explain how a rational agent should make inferences about dispositions in non-constraining situations using Bayesian inference. We then apply this framework to the results of two classic studies of the FAE – Jones & Harris (1967) and Quattrone (1982) – to demonstrate that the patterns of behavior that have been used to argue that our social inferences are flawed can instead be explained by considering what people ought to do given their uncertainty about other people and the situation.

## A Rational Model of Social Inference

Given the uncertainty inherent in reasoning about the causes of others’ behavior, we cannot expect the social inferences that people make to be errorless. The relevant question is not whether we make errors when reasoning socially, but rather if these errors consistent with the errors a rational observer would make, or if they are systematically biased? A simple causal inference account is sufficient to capture how an ideal, probabilistic observer would make dispositional and situational inferences in the classic FAE task. For instance, how would an ideal observer infer the generosity of a museum patron who makes a donation in the presence of a docent?

We will assume that the influence of the situation and the influence of the person’s disposition will combine to yield the probability of an action. This can be expressed as a simple three-node graphical model (Figure 1; Pearl, 1988): Making a donation will be a function of the situation (pressure imposed by the docent to leave a donation) and the individual’s disposition (how generous the person is).

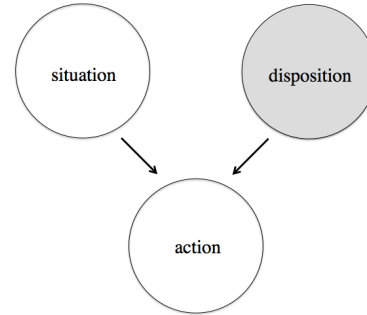


Figure 1: Graphical model of an action arising from two possible classes of causes: Situation and disposition influence the probability that an action will occur. Various attribution experiments amount to conditioning on (observing) two of the three nodes, and inquiring about the third.

In this scenario and the scenarios considered in this paper behavior can be treated as a simple, binary action<sup>2</sup>: the museum patron either leaves a donation, or doesn’t. In such situations the resulting probability ( $q$ ) is the chance that one of the two outcomes occurs, formalized as a draw from a Bernoulli distribution. For simplicity and convention, we will assume that the strength of the situational ( $s$ ) and dispositional ( $d$ ) influences on the action ( $a$ ) is represented as additive in log-odds. Thus, they each take on real values from negative infinity to positive infinity: positive numbers reflect influences that favor a behavior (donating) and negative numbers favor the alternative (not donating). Therefore, the log-odds of an individual donating is the sum of the situational and dispositional influences expressed in this manner. The probability ( $q$ ) of donating may be calculated by applying the logistic transformation to the log-odds of behavior:

$$q = \frac{1}{1 + e^{-(s+d)}}$$

A donation is therefore made with probability  $q$ , and is not made with probability  $1-q$ :

$$P(a | q, \theta) = \begin{cases} 1 & q \\ 0 & 1-q \end{cases}$$

We can quantify both the strength of a person’s disposition and the situation based on how we expect people to act. A person with a disposition of  $d=0$  is equally likely to take the chosen action or not in an unconstrained situation (e.g., will donate 50% of the time when not watched by the docent). People with positive disposition scores are more likely to take the chosen action in an unconstrained

<sup>1</sup> However, see Jennings (2010) for an alternate quantitative Bayesian conceptualization of the FAE and the Discussion for a comparison of his framework with the account proposed here.

<sup>2</sup> Although the situations considered here lend themselves to binary outcomes and thus make the logistic model appropriate, this framework can be extended to continuous action spaces by considering alternate linking functions between actions and the sum of situational and dispositional influences. Indeed, we have expanded the model to accommodate graded actions, but due to space constraints this extension is not discussed here.

situation, and people with negative disposition scores are less likely. The situation strength then defines how much the situation changes this probability. We define an unconstraining situation to have the situation strength  $s=0$ , and so the probability of taking an action (e.g., donating) relies only on the actor's disposition. Positive situation strengths represent conditions that encourage taking the chosen action (e.g., the docent watching), while negative situation strengths represent conditions that encouraging the alternative (e.g., you hear the museums' board is skimming money).

So, for example, you might expect 73% of people to donate to the museum with no docent, which would represent an average disposition of  $d=1$ , since  $1/(1 + e^{-(1+0)}) = 0.73$ . If 92% of people donate when the docent is watching, this would be captured by a situation strength of  $s=1.5$ , since  $1/(1 + e^{-(1+1.5)}) = 0.92$ .

This only explains how situation and disposition might combine to determine the probability that a specific action did or did not occur. But in most cases, people know the situation and observe an action, but do not know disposition. Now we consider how people might reason backwards to disposition once they observe someone take an action in a known situation.

This reasoning requires understanding how dispositions are distributed in the world – e.g., how many people are more generous or stingy than average? Given this prior distribution on dispositions, and the observation of an action  $a$  in a situation of some strength  $s$ , we can calculate the posterior probability of the generosity of the actor using Bayes rule:

$$P(d|a,s) = \frac{P(a|d,s)P(d)}{\sum_{d'} P(a|d',s)P(d')}$$

This provides a posterior distribution over the disposition that the actor might hold. We assume that when people are asked to make a point judgment about disposition afterwards, this judgment will be based on the expected value of this distribution.

### Some Situations are More Informative than Others

With this inference framework, upon observing an actor take an action, that observer should always attribute *some causal role* to the actor's disposition. However, the amount of attribution should depend on the strength of the situation.

Imagine we are now in a world with somewhat more stingy museum patrons than before, where only 50% would donate with no outside influences (here represented by the prior  $d \sim \mathcal{N}(0,1)$ ). Now you observe that a visitor donates money with no docent watching ( $s=0$ ). Based on the equation above, you should infer that the visitor is somewhat more generous than average ( $E[ds=0,a=1]=0.39$ , Figure 2 point A).

But what if there is strong pressure against donating? For instance, if a patron walks in and mentions that they are

broke and choosing between donating or eating lunch ( $s=-3$ ) yet donates anyway, you should infer even more strongly that they are generous ( $E[ds=-3,a=1]=.75$ , Figure 2 point B). If the action occurred despite pressure against the action, it must have been motivated by attitude, and strong attitude inferences are made.

Conversely, if a person donates when the docent is breathing down their neck and telling them the museum is about to go bankrupt because not enough people are donating ( $s=3$ ), the ideal observer will infer something about the patron's disposition ( $E[ds=3,a=1]=.08$ , Figure 2 point C), since a true Scrooge would simply ignore the docent and walk through. As long as situations are not deterministically strong, the ideal observer should make *some* dispositional attribution, but the strength of that attribution should be modulated by situation strength.

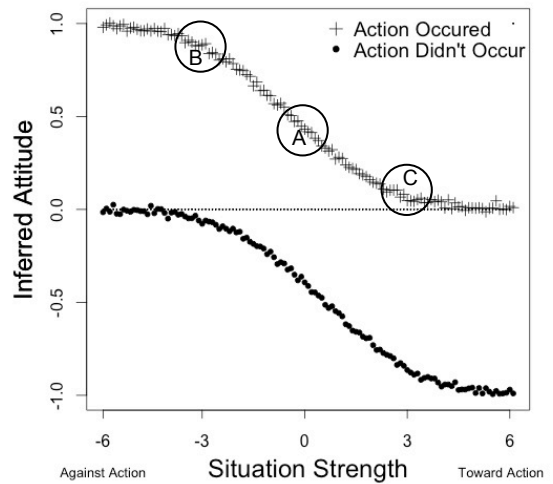


Figure 2: The strength and direction of the attitude that the ideal observer infers depends on the situational pressure in combination with whether the action occurred. When there is stronger situational pressure toward an action and the actor performs the chosen action (+), the ideal observer makes weaker inferences about attitude. However, when there is stronger situational pressure against the action and an actor does it anyway, the ideal observer makes a much stronger inference about attitude. Symmetrically, not doing the action (●) is more informative when the situation was motivating the action, compared to when it was discouraging the action. Until the situation becomes deterministically strong, the ideal observer always infers an attitude that is consistent with an observed action.

### Inferring Attitude From Situationally Motivated Behavior (Jones & Harris, 1967)

In the classic FAE experiment, Jones and Harris (1967) examined how we account for situational pressures when reasoning about other's dispositions. Based on a deterministic (Boolean logic) view of such inferences, they reasoned that a behavior is evidence of a person's disposition, but when there is a situational explanation for the behavior it no longer reveals anything about the actor's disposition. They asked university students to read an essay

that either opposed or endorsed Castro. Participants were told that the essay was written by a classmate who was either instructed to argue for a particular position, or was free to choose whether to write a pro or a con essay. After reading the essay participants answered ten 7-point Likert scale questions (1: strongly anti to 7: strongly pro) about what they thought the author's true attitude toward Castro had been; these ten responses were summed, yielding an overall scale from the strongest anti-Castro beliefs (10) to the strongest pro-Castro beliefs (70). If the essay position were freely chosen, then it obviously reveals the authors' attitude; however, if instructions to write in support of one position or another would make any person — regardless of their disposition — produce a compelling essay for the instructed position, then Jones and Harris suggest that the essay content should not be informative of the authors' attitude.

As predicted, when readers were told the essay position was freely chosen they believed that the author had an attitude about Castro consistent with the views expressed in the essay. However, when the reader was told that the position had been assigned, readers continued to estimate the authors true attitude to be consistent with the essay's position, albeit more weakly (original data re-plotted in Figure 3A). So, people do not behave logically according to Jones and Harris: people infer attitudes when they should explain behavior based on the situation.

But what inferences should we expect from a rational observer who *did not* believe that instructions to write a particular essay was completely deterministic? We can characterize the behavior of such a rational observer via a Bayesian causal inference model: given the observation of either a pro or anti Castro essay (a binary action), and some assumption about the influence of instruction (situation strength) what might the actor's attitude about Castro be (disposition)? From the logic captured in Figure 2, we would expect that such an observer would infer *some* attitude that is consistent with the essay even when the position had been assigned. If the instruction to write a pro-Castro essay does not completely determine behavior, then those with vehemently anti-Castro views might still write an anti-Castro essay; therefore, seeing a pro-Castro essay still tells us *something* about the author's attitude, namely that they are not so strongly against Castro that the instructions were insufficient to compel them to write a pro-Castro essay. So, qualitatively, a rational observer believing in soft (non-deterministic) influence of situations will still infer some disposition, but just how much depends on the observers' assumptions about how compelling the situation is. What assumptions about the "strength of the situation" (of being assigned to write an essay supporting, or opposing Castro) would such an observer need to make to yield not just the qualitative, but also the quantitative pattern of participants' judgments?

To formalize this we must specify the "situation strength" of being assigned to write an essay supporting or opposing Castro, and those under free choice. Since the majority of

reader-subjects reported being anti-Castro, we assumed that readers had strong prior belief that a majority of people (about 70%) were against Castro. We assume that a neutral person (not average, but split between positions) who chooses what to write would be equally likely to produce a pro or anti-Castro essay: this is equal to a neutral situation strength. Further, we assume that the assignment to write a pro- or anti-Castro essay have situation strengths that would compel a perfectly neutral person to write the assigned essay 88% ( $s=2$ ) of the time. Finally, we scaled posterior beliefs about disposition to the 10-70 point scale using a logistic transformation.

Under these assumptions,<sup>3</sup> an ideal observer infers the same pattern of dispositions as people do: when the situation does not exert any pressure (the "choice" conditions) the ideal observer treats the attitude expressed in the essay as very informative, and infers that the author's true attitude roughly mirrors what was expressed in the essay. When the situation does exert pressure to take a particular position (the "no-choice" condition) both humans and the ideal observer treat the behavior as informative (though less so), and make correspondingly weaker dispositional attributions (Figure 3B).

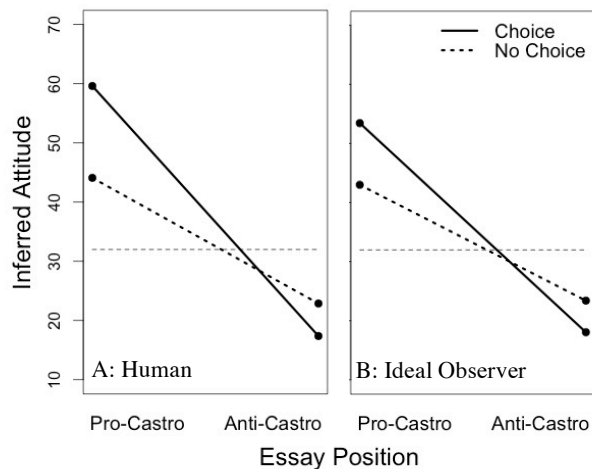


Figure 3: Inferred attitude as a function of essay position, and whether this position was chosen or assigned. A: People inferred that the position expressed in an essay was indicative of the author's true attitude both when they chose their position, and when it was assigned (Jones & Harris, 1967). The attitude attributed to the author was stronger when the author chose his position (solid line), and weaker when assigned (dotted line). B: An ideal observer also infers that the essay is indicative of the author's true attitude, but more informative when the position was chosen, and less informative when assigned (dotted line).

<sup>3</sup> Prior beliefs about situations and attitudes, and interpretations of the strength of the particular situations (for Jones & Harris, 1967) and attitude (for Quattrone, 1982) influences were not collected in the original studies. Therefore we chose plausible parameters to predict peoples' inferences. Importantly, all reasonable parameter choices yield the same qualitative patterns.

## Inferring Situation when Attitude is Known: Inverting the FAE (Quattrone, 1982)

The theory of the FAE suggests that people are prone to overestimate the influence of disposition and to underestimate the influence of situation. If people indeed have this tendency then we would not expect them to infer additional situational pressure when a known disposition has already accounted for an action. Yet there is a curious finding in the literature that suggests the opposite: when people know about an actor's disposition, they are more likely to "over-attribute" the actor's action to situational pressures.

Quattrone (1982) asked subjects to read an essay favoring or opposing the legalization of marijuana, but rather than knowing if the essay position was chosen or assigned, subjects were told that the author was known to have either a neutral opinion about legalization, or an opinion consistent with the attitude expressed in their essay. Subjects were told that the purpose of the study was to determine if extraneous experimental factors (e.g. experimenter bias) might be influencing the opinions people expressed. After reading the essay, subjects were asked to estimate the likely situational pressure on a 30 point Likert scale (-15: pressure to oppose, 15: pressure to favor). Even when subjects were told that the author held a pro-legalization view, they believed there was pressure to write a pro-legalization essay, and vice versa (original data re-plotted in Figure 5A).

According to the logic of classic FAE studies, this could be considered an "over-attribution" of the situation, since the pre-experiment attitude is known to have caused behavior. This finding is inconsistent with typical explanations of the FAE and calls into question the theoretical accounts of the FAE that claim that we have an inclination to over-attribute behavior to dispositions, and not attribute enough to situations (e.g. Taylor & Fiske, 1978; Gilbert, Pelham & Krull, 1988).

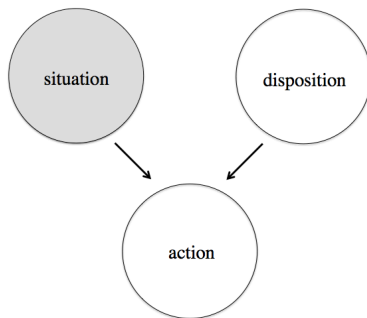


Figure 4: Graphical model show that situation and disposition influence the probability that an action will occur. Instead of conditioning on (observing) situation and the action, and inquiring about situation, here we condition on disposition and action, and inquire about the strength of the situation.

A Bayesian Inference account, however, predicts this pattern of results. When someone behaves in a way that is

motivated by their known disposition, it is still rational to infer that the situation was also motivating the action, given that probabilistic dispositions do not completely determine behavior. Assuming the same generative process as explained previously (Figure 1), inferring the unknown situation strength given a known disposition is symmetric to the inference process applied previously to the attitude-attribution paradigm. Here, we observe an action and know the disposition but now must reason about the likely situation strength (Figure 4). Knowing the disposition and what action the agent chose, but having a prior distribution over types of situations people encounter, we can use Bayes formula to derive a posterior probability of the impact of the situation:

$$P(s | a, d) = \frac{P(a | s, d)P(s)}{\sum_{s'} P(a | s', d)P(s')}$$

This framework provides mirrored inferences to the framework used to reason about disposition: an ideal observer should always infer that the situation has some impact on an observed action. Just as the rational observer would infer that a museum patron who gives a donation is generous even when a docent is watching, it is also rational to infer that if a generous friend donates, there was some pressure for her to donate. And just as observing an action is more informative about an actor's disposition in some situations compared to others, how much an action informs you about the situation also depends on the how strong that actor's disposition is. But so long the actor's disposition does not compel them to act identically in all situations, it is rational to infer that the situation had some impact.

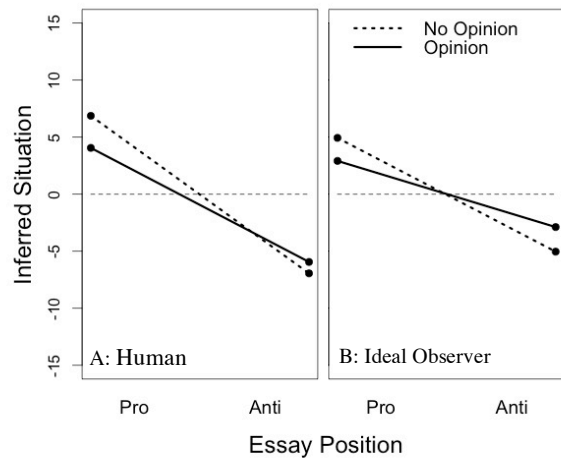


Figure 5: Inferred situation as a function of the essay and the known attitude. A: Subjects inferred that the position expressed in an essay was evidence for how much the situation was motivating behavior, both when they thought the author had a pre-existing attitude, and when the didn't. The situation inferred was stronger when they thought the author had no pre-existing opinion. B: The ideal observer also infers that the situation was pressuring the essay position, but more so when the author had no existing opinion.

In the Quattrone (1982) task, the ideal observer once again observes a binary action (either a Pro- or Anti-Castro Marijuana essay) but now they know the strength of the author's true attitude and must infer the strength of the situation. To formalize the "no opinion" condition we assume that people would be equally likely to write a pro- or anti-legalization essay, and in the "opinion" condition we specify that the proclivity to write a pro- or anti-legalization essay has a strength of 1 and -1, respectively (meaning that 73% of people would write an essay consistent with their attitude in a neutral situation). Just as before, we used a logistic transformation and rescaled the expected posterior situation strength to place it on the same scale as Quattrone (1982).

Consistent with a Bayesian Inference account, and in contrast to what a FAE framework would predict, when the reader thought the author had a no opinion before the experiment they inferred that the experimental situation had exerted some force on the writer. When the reader thought the author had a pre-experiment opinion about marijuana legalization, they still inferred that the experimental situation exerted some pressure, just less than what readers in the "no opinion" condition inferred (Figure 5A). An ideal observer demonstrates the same pattern of results (Figure 5B), suggesting we can explain why people will in some situations behave opposite to the predictions of the FAE.

## Discussion

Our results show that human attribution of behavior to situational and dispositional causes – which has long been considered systematically biased – can be reconciled with the inferences made by a rational observer reasoning in a probabilistic world. We demonstrated how judgments in the prototypical Fundamental Attribution Error paradigm are not errors, but rather are the result of inferring internal qualities in a world where situations do not fully constrain behavior. Rational attitude attribution might be rendered consistent with human judgments merely by adding a biased prior about the strength of situations (i.e. by supposing that situational constraints are systematically underestimated); however this would amount to merely reframing the FAE in Bayesian jargon. For instance, Jennings (2010) assumed that reasoning about dispositions could be explained by Bayesian inference using a biased prior, and showed that people's attributions could still be internally consistent. Critically, however, we show that a situation in which people behave inconsistently with the typical explanation of the FAE – over-attributing behavior to the situation when disposition is known – is also a natural consequence of rational probabilistic reasoning, and could not be explained simply by miscalibrated expectations about situations.

The mechanisms of our ideal observer model are not particularly sensitive to parameters: all parameter values capture the qualitative effects in the classic FAE studies, and *a priori* plausible values yield a good quantitative fit. Nonetheless, in future work we will empirically obtain observers' expectations about situations and dispositions, as

well as how they interpret the influence of specific situations, to compare parameter-free predictions of our model to human behavior.

In short, our work suggests that results from decades of attribution experiments, which have been classically interpreted as evidence that our social inferences are fundamentally flawed, might instead be the natural outcome of reasoning about a complex and uncertain world.

## References

- Ajzen, I., & Fishbein, M. (1975). A Bayesian analysis of attribution processes. *Psychological Bulletin*, *82*(2), 261.
- Baker, C. L., Saxe, R., & Tenenbaum, J. B. (2009). Action understanding as inverse planning. *Cognition*, *113*(3), 329-349.
- Gilbert, D. T., & Malone, P. S. (1995). The correspondence bias. *Psychological Bulletin*, *117*(1), 21.
- Gilbert, D.T., Pelham, B.W., & Krull, D.S.(1988). On cognitive busyness: When person perceivers meet persons perceived. *Journal of Personality and Social Psychology*, *54*(5), 733.
- Gladwell, M. (2000). *The Tipping Point*. Boston: Little, Brown
- Griffiths, T. L., & Tenenbaum, J. B. (2005). Structure and strength in causal induction. *Cognitive psychology*, *51*(4), 334-384.
- Griffiths, T. L., & Tenenbaum, J. B. (2009). Theory-based causal induction. *Psychological review*, *116*(4), 661.
- Holyoak, K. J., & Cheng, P. W. (2011). Causal learning and inference as a rational process: The new synthesis. *Annual Review of Psychology*, *62*, 135-163.
- Jennings, K. E. (2010). Determining the Internal Consistency of Attitude Attributions. In *Proceedings of the 32nd Annual Conference of the Cognitive Science Society*. Austin, Texas: Cognitive Science Society (pp. 978-984).
- Jones, E. E., & Davis, K. E. (1965). From acts to dispositions: the attribution process in person perception. *Advances in Experimental Social Psychology*, *2*, 219-266.
- Jones, E. E., & Harris, V. A. (1967). The attribution of attitudes. *Journal of Experimental Social Psychology*, *3*(1), 1-24
- Jones, E. E., Worchel, S., Goethals, G. R., & Grumet, J. F. (1971). Prior expectancy and behavioral extremity as determinants of attitude attribution. *Journal of Experimental Social Psychology*, *7*(1), 59-80.
- Kelley, H. H. (1973). The processes of causal attribution. *American Psychologist*, *28*(2), 107.
- Kemp, C., & Tenenbaum, J. B. (2009). Structured statistical models of inductive reasoning. *Psychological Review*, *116*(1), 20.
- Morris, M. W., & Larrick, R. P. (1995). When one cause casts doubt on another: A normative analysis of discounting in causal attribution. *Psychological Review*, *102*(2), 3
- Nisbett, R. E., Caputo, C., Legant, P., & Marecek, J. (1973). Behavior as seen by the actor and as seen by the observer. *Journal of Personality and Social Psychology*, *27*(2), 154
- Pearl, J. (1988). *Probabilistic reasoning in intelligent systems: networks of plausible inference*. Morgan Kaufman.
- Quattrone, G. A. (1982). Overattribution and unit formation: When behavior engulfs the person. *Journal of Personality and Social Psychology*, *42*(4), 593-607.
- Ross, L. D., Amabile, T. M., & Steinmetz, J. L. (1977). Social roles, social control, and biases in social-perception processes. *Journal of Personality and Social Psychology*, *35*(7), 485.
- Ross, L.D., & Nisbett, R.E. (1991). *The Person and the Situation: Perspectives on Social Psychology*. New York: McGraw-Hill
- Sherman, S. J. (1980). On the self-erasing nature of errors of prediction. *Journal of Personality and Social Psychology*, *39*(2), 211.
- Taylor, S. E., & Fiske, S. T. (1978). Salience, attention, and attribution: Top of the head phenomena. *Advances in Experimental Social Psychology*, *11*, 249-288.