THE IMPORTANCE OF BEING PARTIAL
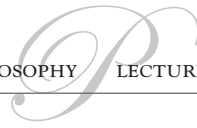
THE CONSTRUCTIVE ROLE OF BIAS IN HUMAN LIFE

LOUISE ANTONY

The Importance of Being Partial: The Constructive Role of Bias in Human Life
Louise Antony

ABSTRACT

The term "bias," as it is generally used, connotes something bad: a prejudging that is unfair or illegitimate, or an improper favoring of some viewpoint or some person's interest over others'. But fundamentally, "bias" means simply a "bent or tendency; an inclination in temperament or outlook" I argue that "bias," in this neutral sense, plays a constructive – indeed, absolutely crucial – role in human life. There are two areas of human life where this is especially true: first, in our cognitive/epistemic life – in our attempts to come to know our world, and second, in our "affective" lives – in our relations and emotional connections to other human beings. When human biases are bad, I contend, it is not because they are biases, but because they are operating in the wrong contexts. In the cases that concern those of us interested in social justice, there are many contexts in which biases – both epistemic and normative – are problematic, but the problem lies with the environment, not the biases.

The Importance of Being Partial:

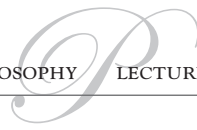# The Constructive Role of Bias in Human Life

Louise Antony

University of Massachusetts

ONE OF THE PRIMARY AIMS of liberatory social movements is the elimination of bias. But what, exactly, is "bias"?

The term "bias," as it is generally used, connotes something bad: a prejudging that is unfair or illegitimate, or an improper favoring of some viewpoint or some person's interest over others'. When we think of bias, we think of things like partisan news reporting and hateful prejudice against members of certain religions or members of minority races.

But fundamentally, "bias" means simply a "bent or tendency; an inclination in temperament or outlook...." Notably, there is nothing normative in this definition, nothing to suggest that there is anything inherently wrong with being biased. A coin is biased if it is more likely to land on one face than on the other. (You can bias a coin by bending it into a slightly concave shape. When the coin is flipped, it will land more often with the concave side up.) Clearly, there is nothing inherently wrong with a coin that is biased in this sense. Stage magicians use biased coins all the time, often to their audiences' delight. What *would be* wrong would be using a biased coin to bet against people who don't know that it is biased. Introducing a biased coin into a betting game in which other parties assume the coin to be fair is a violation of trust, a thwarting of the conditions assumed to be in force by the parties to the game.

This example illustrates the general line I want to take about bias in connection with human epistemic and social activity: neither good nor bad in itself, bias is something that can, in context, work for good or for ill. I will argue that "bias," in this neutral sense, plays
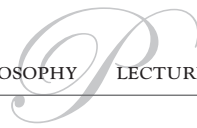
a constructive – indeed, absolutely crucial – role in human life. There are two areas of human life where this is especially true: first, in our cognitive/epistemic life – in our attempts to come to know our world. Second, in our "affective" lives – in our relations and emotional connections to other human beings.

Cognitive biases enable, in a fundamental way, the development of human knowledge. In the cognitive realm, biases explain how we solve the problem of *underdetermination*: Our knowledge of the external world depends on our sensory experience, but because we are finite embodied agents, our sensory access to the external world is sharply limited by time and by space. Whether we are talking about our visual system's constructing a representation of three-dimensional space from a two-dimensional array of information, a human child's success in acquiring language, or a scientist's construction of powerful theories about the origin and evolution of the universe, the tiny body of sensory evidence available to figure stuff out is always consistent with a multitude of different hypotheses. And yet, these epistemic feats – from the mundane to the epic – have been achieved.

Affective biases make possible the sort of focused interpersonal connections that are psychologically necessary for human flourishing. Human beings are deeply dependent on others for much of our lives – as babies, for example, we are helpless and must be provided with every one of our animal needs. Even as we mature, it appears that we need sustained, focused attention from specific others in order to flourish psychologically and emotionally.

So these are the basic ways in which biases are *good*. But what about *bad* biases? Using (or relying on) biases seems to conflict with norms. In the epistemic realm, biases seem in conflict with the norm of objectivity: we want to be able to criticize forms of bias that lead to distortions of the truth. In the affective realm, biases seem to conflict with the norm of ethical impartiality: we want to condemn racism, sexism, and other social practices that privilege some people over others. In short, if we want to endorse objectivity and impartiality in some form, we need to explain the difference between good biases and bad biases.

Here's the general line I want to take: when human biases are bad, I contend, it is not because they are *biases*, but because they are operating in the wrong contexts. In the cases that concern those of us interested in social justice, there are many contexts in which biases – both epistemic and normative – are problematic, but the problem lies with the environment, not the biases.

First, however, I want to outline explicitly the conception of objectivity and bias to which I am opposed.

Bias, Objectivity and Impartiality

WHEN WOMEN BEGAN, in the 1980s, to enter the academy in (relatively) large numbers, feminist and other progressive scholars began challenging the claimed objectivity of work that had been done in fields hitherto dominated by privileged white men. Feminist historians, sociologists, economists and biologists showed over and over how the domination of their fields by men had led to distortions in either doctrine or method or both. Feminist philosophers began to scrutinize the canonical works of Western philosophy. Evidence of contempt for women in the writings of great philosophers was not hard to find: Aristotle, Rousseau, Locke, Hume, and Kant were all on record as defenders of patriarchy, citing deficiencies in women's minds or characters (or both) as warrant for male dominance.

One response to such discoveries was simply to dismiss them as failures of rationality or observation on the parts of the philosopher-authors – failures which would have been unremarked and unremarkable, given their milieux. On this way of thinking, the offensive bits from Aristotle and Kant could simply be excised, preserving for feminist purposes the central tenets of the philosophers' moral theories. But many feminist theorists believed that the problem went far deeper than a few gratuitous expressions of sexism. They argued that many of the concepts and values at the heart of Western philosophy reflected concerns and ways of thinking peculiar to privileged white men.

One of the concepts that came under this sort of scrutiny was the concept of *objectivity* itself. Feminist critics (e.g., Harding, Jaggar, Alcoff, Code) charged that this notion, at least as it had functioned within mainstream epistemology, set up an unachievable ideal of epistemic practice, even as it disguised the actual conditions of human inquiry. All human knowledge, the critics argued, is *partial* – that is, limited by the material circumstances of its subjects. But the modernist ideal of objectivity suggests that knowledge can and should be *impartial* – that it is possible for human beings to transcend the limitations of their particular epistemic positions and achieve an epistemic standpoint that would somehow constitute a "view from nowhere." That such a standpoint could even be thought possible reflected,

according to feminist critics, a social position which allowed privileged men to ignore the material messiness of human existence, in a way that women, who had to contend with pregnancy, childbirth, care of the sick, care of the home, and so forth could not. Furthermore, the feminist critics charged, such a standpoint was objectionably individualistic – it suggested that human knowers were epistemically self-sufficient, when in fact we all depend crucially on epistemic interaction with others. Finally, the critics argued, mainstream philosophers erroneously and arrogantly presumed that they themselves had managed to achieve this standpoint of perfect impartiality. They assumed and promoted an unearned authority for their own perspectives, while tacitly rationalizing the erasure of other voices, including women's.

Ironically, however, this critique of the notion of objectivity did not comport well with the feminist project of criticizing male bias; the one seemed to undermine the other. *If impartiality was an objectionable ideal, what could be the basis for objecting to male partiality*? Perhaps men philosophers could coherently be criticized for failing to recognize that their views expressed their own partial points of view, but how could such partiality constitute *distortion*? If we reject objectivity, how can we explain what's *wrong* with bias? This problem, which I have dubbed the "Bias Paradox," sets a desideratum for any acceptable account of objectivity: it must be compatible with the facts about the essential locatedness of human knowledge, without disabling critiques of pernicious forms of partiality.[1]

Epistemology and the philosophy of science were not the only branches of philosophy where the notion of objectivity as impartiality came under attack. Some men philosophers, such as Bernard Williams and Michael Slote, had already challenged the role of objectivity within the realm of value theory, arguing that impartialist ethical theories such as utilitarianism and Kantianism neglected and or even denigrated our affective connections with other, particular human beings. Feminist philosophers, however, connected this general critique of impartialist theories to the overall neglect by mainstream theorists of the ethical value of *caring* work – work that is done, by and large, by women. Annette Baier, for example, argued that an exclusive focus on justice neglects the ethical value of caring activity, which necessarily involves our privileging the concerns and needs of particular persons – paradigmatically, our family and friends. Normative theories that neglect care, argued Baier (and others),

---

1   For a fuller exposition of this argument, see my "Quine as Feminist."

fail both descriptively and prescriptively – they fail to characterize an enormously important part of what is humanly valuable, and they give us no guidance about how to conduct ourselves in that part of life. A more adequate ethics would acknowledge the necessity of partiality for the proper performance of caring labor, and would assign proper ethical value to the particularistic relationships that form the fabric of human life.[2]

But here, as in the epistemic case, we risk running afoul of an important norm: justice. When, and to what degree, may we privilege our nearest and dearest, without taking into account the needs of others with whom we happen not to be acquainted?

So while I completely agree with Longino, Baier and other feminist critics of objectivity as impartiality who argue that the essential locatedness of human life conditions both our epistemology and our ethics, I insist that theoretical work still needs to be done to show how this can be accomplished, consistent with the demands of important epistemic and ethical norms. I now turn to my own account of bias.
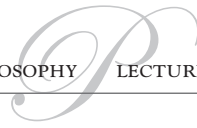
Underdetermination and the Constructive Role of Bias

AN IMPORTANT PART of the story is the conception of objectivity that is the object of the critiques we've been looking at. There is a popular conception of what it is to be objective that derives from an empiricist view of the human mind. The empiricists believed that the mind had only formal structure – that it was a machine for making associations and inferences – and that all substantive knowledge came from the senses. Objectivity, then, involved two things: strictly conditioning one's opinions to the evidence, and preventing interference from emotions, desires, and affections. In short, what was wanted was "Dragnet Objectivity."[3]

But this view of objectivity offers an inappropriate ideal for human knowers. If we ever did – *per impossibile* – attain this ideal, we would actually have less knowledge, not more. Fortunately, there is a more adequate view of human knowledge available now, one that puts at the center the facts about human embodiment that we've been surveying. I'll first look at

---

2   See Baier, "The Need for More than Justice."

3   I'm referring to the avowed ethos of Sgt. Bill Friday on the old television show *Dragnet*. To short-circuit theorizing or speculation by witnesses, Friday would interrupt witnesses with a solemn "Just the facts, ma'am."

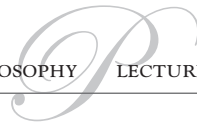the pertinent developments in philosophy and psychology, and then draw out some implications for ethics.

So what exactly does it mean to say that human knowledge is *partial*? One thing it means, of course, is that whatever we come to know is only a small part of what there is to know. But it also means something about the relationship between the evidence we have and the judgments we infer on its basis. Empiricists, as I said above, thought that all substantive knowledge derived from sensory experience. But my particular sensory experience is extremely sparse, especially considered relative to the scope and complexity of contemporary scientific theories, and is limited to local goings-on in my infinitesimally tiny corner of the universe.

One response to these concerns is, of course, skepticism about the external world. But there's another possible response: one could take it as a datum that we *do* manage to acquire knowledge of the world, and treat partiality as posing a puzzle to be solved empirically. This was the course that W. V. O. Quine recommended in "Epistemology Naturalized."[4] He referred to the problem I've been calling "partiality" as the problem of *underdetermination*: the fact that for any finite amount of data, there are always an infinite number of hypotheses logically consistent with that data. The problem for a naturalized epistemology, then, was not *whether* we had knowledge, but rather *how*.

Orthodox empiricists, including the logical positivists of the twentieth century, held that empirical beliefs, to be justified, had to be based exclusively on sensory evidence. Sensation, they held, was the subject's only *objective* connection to the world. Ideally, theoretical judgments could be justified by their being shown to be logically constructible out of sense experience alone. Quine pointed out, however, that even if this idealized "rational reconstruction" of science could be carried out – and he gave reasons to think that it couldn't – we would still be lacking an explanation of how *actual* scientific practice, which flagrantly violated empiricist prescriptions, managed to yield successful theories.

Quine's own solution to the problem of underdetermination, it turns out, marked a decisive break with classical empiricism. He argued that the gap between evidence and theory was filled in for human beings by – wait for it! – *biases*. The early modern empiricist, David Hume, had already come to somewhat the same conclusion in offering his "skeptical solu-
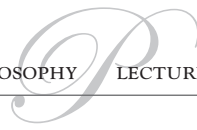
---

4   Quine, "Epistemology Naturalized."

tion" to the problem of induction. What justifies our thinking that the future will resemble the past? he asked. Hume's answer was that we could not expect to find justification, in the ordinary sense. The principle could not be acquired in the familiar way as a generalization from experience. This is because the principle's being true is a condition of our using past experience as evidence of anything. Belief in this principle – or at any rate, the disposition to draw inferences from experience in accordance with the principle – had to be innate. We could therefore only expect a certain kind of pragmatic (rather than epistemic) justification for acquiescing in this ingrained cognitive bias – in short: when we do form beliefs in accordance with this principle, things seem to go well.

Quine, similarly, argued that we bring a variety of other, more specific biases – what he called "extra-empirical" assumptions – to any epistemic problem. This suite of native biases cuts down to manageable size the set of hypotheses that we must consider seriously. For example, we display a clear preference for simpler hypotheses over more complicated ones. What reason do we have for thinking this preference is reasonable? On the one hand, it seems presumptuous to imagine that nature conforms itself to the limits of our minds; on the other hand, the assumption seems to have stood us in pretty good stead. Moreover – and this echoes Hume – we have no choice; without such assumptions we cannot learn anything at all. Even the austerely empiricist behaviorist psychology to which Quine was committed needed to posit an "innate similarity space" that made animals selectively more sensitive to certain parameters of resemblance than to others in the stimuli to which they were exposed.

Chomsky famously canonized this form of argument for native structure as the "poverty-of-the-stimulus" argument. In his work, he focused on an amazing but perfectly mundane human achievement – the acquisition of natural language. Human children, he observed, acquired mastery of this complex symbol system by the age of four, without formal instruction, simply on the basis of exposure to other speakers. In language acquisition, the information available to children vastly underdetermined the grammatical structure of the language spoken around them. The children, therefore, must be bringing some kind of innate knowledge to the acquisition situation. Chomsky posited, then, a native structure he called "universal grammar" (UG) that served to constrain sharply the set of hypotheses

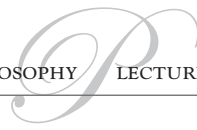a child could consider in response to the linguistic evidence around them, thereby vastly simplifying her learning task.[5]

Since then, cognitivists in linguistics and several areas of psychology have posited a wide variety of native mental biases, or "instincts," in humans and in non-human animals alike, to fill in the gaps between sensory evidence and cognitive attainment in many domains. In vision science, Irvin Rock and his followers posited "hidden assumptions" within the visual system to account for the computation of three-dimensional spatial facts on the basis of two-dimensional data. Developmental psychologists have posited several innate cognitive "modules" that operate, analogously to Chomsky's language acquisition device, to enable rapid learning within typical human environments. These include Elizabeth Spelke's "core cognition" and Alan Leslie's "theory of mind" modules. Some clinical researchers believe that an absence of this theory of mind module might be at the root of the difficulties autistic individuals have interpreting the behavior of neurotypical individuals.

In short, biases of a certain sort play a constructive – indeed an enabling – role in the development of human knowledge. They solve a great deal of the underdetermination problem. And they also speak to the issue of idiosyncrasy: because they are species-wide, they do not reflect peculiarities of any individual's epistemic position. But this can't be the full story about human bias. The biases I've discussed are all benign. But we know that there are also biases that lead us astray epistemically, ethically, or both. I'll discuss two kinds of "bad bias": social prejudice and unethical partiality. My claim is that in both cases, the badness results from the exercise of unobjectionable cognitive or ethical biases in unpropitious circumstances.

Social Prejudice

I SPOKE EARLIER of Quine's "innate similarity space," posited to explain how we group things for the purposes of generalization. Contemporary developmental and cognitive psychologists and psycholinguists (e.g., Susan Gelman, Marjorie Rhodes, Sarah-Jane Leslie) have added detail to Quine's picture (though not in a way he would have liked). According to these re-

---

5   See Chomsky, *Reflections On Language*.

searchers, human beings operate with a native folk metaphysical theory Susan Gelman has dubbed "essentialism":

> Roughly, essentialism is the view that categories have an underlying reality or true nature that one cannot observe directly but that gives an object its identity. In other words, according to essentialism, categories (such as "boy," "girl," or "intelligence") are real, in several senses: they are discovered (rather than invented), they are natural (rather than artificial), they predict other properties, and they point to natural discontinuities in the world.[6]
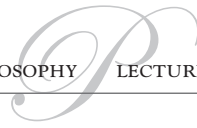
If Gelman and others are right, then even if initial groupings of things are made on the basis of observable similarities, human children (and human adults, it turns out) unconsciously posit unobserved (if not unobserv*able*) essences as the properties that determine membership in the group. Physical appearance is a factor that often triggers essentialism, but it is not the only cue human beings rely on to posit essentialist groupings. For human beings, language is a very powerful cue. Not only does the language a person speaks indicate whether they are a member of one's own "tribe", grammatical construction can signal which attributes of other people are regarded as stable and consistent across contexts, and thus which attributes indicate groupings that are candidates for essentialization. Gelman and Heyman, for example, found that when children were introduced to a group by means of a noun ("Rosie is a carrot-eater") as opposed to a verb phrase ("Rosie eats carrots whenever she can") they were much more likely to project the property of carrot-eating to others introduced in that context. Generic nominal expressions, such as "bare plurals" (plural nouns with no determiners or quantifiers, such as "ducks" in "ducks lay eggs"), seem both to trigger and to express the essentialization of groups.[7]

It is the fact that very young children display strong evidence of essentializing that the tendency to essentialize is thought to be innate. But this tendency does not fade away; adults show the same patterns. Philosopher Sarah-Jane Leslie has brought many of these findings together in her theory of "striking-property generics."[8] Contemporary adults reliably judge as true statements like "ticks carry Lyme Disease" despite the fact that only a tiny fraction

---

6   Gelman, *The Essential Child*, 3.

7   Pertinent research is summarized in Leslie, "The Original Sin of Cognition."

8   Leslie, "The Original Sin of Cognition."

of ticks carry the Lyme bacterium. Leslie's explanation is that, first of all, carrying Lyme disease is a *striking* property – it is distinctive of ticks, not common to all arachnids – and it is, secondly, a *dangerous* property – Lyme Disease is a serious illness. If the kind "ticks" has been essentialized, then we will be prone to inferring that members of the kind, even if they do not *actually* carry Lyme Disease, all have the *potential* or *propensity* to carry it.

While it's reasonable to suppose that this human habit of essentializing and projecting is often epistemically efficient, and usually benign, the habit can lead to pernicious biases in certain social conditions. Leslie argues that when people's religions, ethnicities or cultures are experienced as alien or "other," shallow linguistic cues can trigger or support false essentialist views. If some members of the essentialized group are observed to have a "striking" property – even if it is only a miniscule minority of the whole – it is likely that some people will come to accept striking-property generics about the group. Significantly, adults do *not* essentialize groups of which they themselves are members. For their own groups, people recognize the existence of individual differences and do not project a striking property displayed by one member to the other members of the group. All of this is, sadly, borne out by data about Caucasian Americans' negative attitudes toward Muslims since the destruction of the Twin Towers on September 11, 2001.[9] These data must be considered in the context of the actual facts about domestic terrorism: perpetrators of mass shootings, bombings and other forms of politically motivated violence in the United States have disproportionately been white, male, and non-Muslim.[10] The true demographics of domestic terrorism in the United States may finally be coming into focus as a result of the mob attack on the US Congress on January 6, 2021.

Unconscious and automatic processes of generalization sometimes result in explicit beliefs – propositions that people will avow as their own. But often these processes result in what psychologists call *implicit* attitudes – thoughts or preferences that are evinced in subtle behavioral differences toward members of different social groups. Such differences often

---

9   See Hartig and Doherty, "Two Decades Later, the Enduring Legacy of 9/11," especially the section entitled "Views of Muslims, Islam grew more partisan in years after 9/11."
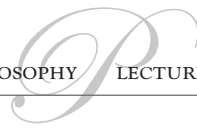
10   For data on mass shootings (which does not include the recent tragedy at a Pittsburgh synagogue), see Berkowitz and Alcantara, "The Terrible Numbers That Grow with Each Mass Shooting." For data on acts of domestic terrorism (which does not include the recent spate of mailed bombs), see CNN ___.

emerge in conditions where one must act quickly or without conscious reflection, and can affect everyday interactions and evaluations. Implicit attitudes can and do co-exist alongside contrary explicit beliefs and preferences. This can make change difficult: it is hard to persuade people that their practices are compromised by biases they sincerely deny they have.

But of course not all biases are implicit. Unfortunately, there are still plenty of people who will forthrightly espouse racist, sexist, homophobic or other socially prejudiced beliefs. Frequently, such beliefs represent an externalization of the essentializing processes described above. Philosophers of race agree that the perjorative term "racism," when applied to individuals, rather than to social structures or historical circumstances, is most apt when made in application to a false belief in the existence of racial "essences." Such beliefs are typically colored by negative emotional dispositions, such as antipathy and disgust, evinced toward members of the essentialized group.

But beliefs of this sort are not the only kind of belief people have in mind when they think of social biases. Consider claims like the following: "women want to be mothers" and "blacks have lower tested IQs than whites." Many people who are committed to social justice are flummoxed when confronted with examples like this. There seems to be *something* sexist or racist about them, but what? They are, according to polls and psychometric research, both true and justified. My claim, however, is that they are *only* true because the reality they report has been *deformed* by *injustice*. Furthermore, there are issues about the *expression* of such claims. In certain contexts, the assertion of such claims can function, pragmatically, to perpetuate the unjust structures that made them true in the first place.

The context in which such claims are made is generally one in which an explanation is being sought. Why are there so few women scientists at Harvard? Because women want to be mothers. Why are blacks so badly underrepresented in the professions? Because blacks have lower IQs than whites. When proffered as *explanations*, claims like this carry a strong *implicature* – a message that is conveyed without being explicitly asserted – that the attributes mentioned have a *necessary* and *immutable* connection to the groups involved. In this way, a certain amount of essentializing is encouraged *pragmatically* by the mere *expression* of claims that, taken as mere empirical generalizations, are true. In the case of many groups, especially women, this pragmatic process can create feedback loops of self-confirmation. The *expression* of general claims – even if true – can promote false ideas about women's or Black persons'
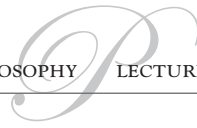
natures, all the while deflecting attention from the extrinsic factors that structure choices for members of these groups: for example, the facts about the racist or sexist expectations of others that make certain jobs difficult for Black men and women of any color to perform, social structures that make it necessary for women to choose between motherhood and careers, and the social norms that penalize women who appear not to value motherhood above all else. Thus, the claim "most women value family over career," while probably *true*, can, if offered in answer to the question "why are there so few women scientists in elite universities?" serve pragmatically to obscure the social conditions that force women but not men to choose between the two, and to send the covert message that women's unconditioned preferences are simply different from men's.[11]

What, then, can be said, in general, about the difference between "good bias" and "bad bias" in the cognitive domain? This distinction cannot be made on the basis of etiology – good and bad may both spring from the same sensory and cognitive dispositions. *I contend that the difference between the good ones and the bad ones depends, rather, on the environments in which they are exercised.* In the case of social biases, we can expect that if our social environment has been shaped by injustice, the regularities and the saliencies that our native faculties are prone to register will not tell us what we want to know.

But what, then, becomes of the virtue of epistemic objectivity? Were we simply wrong to think that there is any such virtue?

I think that objectivity is still an ideal to be pursued. But I think we should understand it not in terms of *epistemic* biases, but rather in terms of *biases of interestedness*. Objectivity, all will agree, has something to do with a connection to things as they are in themselves. Viewed one way, this is an unachievable ideal, because we cannot transcend our own cognitive perspectives. Viewed in another way, however, it has to do with our policies about how our inquiry is to be governed. I submit that inquiry is more or less objective according to how neutral our *desires* are with respect to the inquiry's outcome. Bearing in mind the essential role played by native and (some) acquired *epistemic* biases we possess, we may still constrain

---

11   It is difficult to get polling data directly pertinent to the question, what do women want, but available data makes it clear that women choose to alter their worklives in response to family obligations at a higher rate than do men. See Parker, "Women More Often Choose Family Over Work."

ourselves to respect the methods and assumptions to which we have committed ourselves, by remaining *conatively* neutral – neutral with respect to our desires and interests – about the results of our efforts. To the extent that we have something non-epistemic *at stake* in research, objectivity is compromised. Research into the health effects of smoking should not be financed by tobacco companies; public organs of communication should not be controlled by companies who sell personal data. A high-ranking official of a contending political party should not be in charge of running elections. (In Ohio, during the 2004 presidential election, Ken Blackwell was both the secretary of state, with jurisdiction over election policy, and the honorary co-chair of George W. Bush's re-election campaign in Ohio.)

A bias toward conclusions that one *wants* to be true obviously lacks the kinds of justification enjoyed by the enabling biases I describe above. And indeed, we can say more: unlike epistemic biases that pay their way, eventually, by facilitating empirical success, biases of interest are in no way disciplined by the objective state of the world. If what we want to be true turns out to *be* true, our practices must ensure that that is nothing but an accident.

Against this background I can endorse philosopher Helen Longino's model of objectivity as diversity – the difference is that, whereas she holds that we need diversity of beliefs, I think it is *interests* that must be diverse. The best way to guarantee *disinterestedness* in inquiry is to ensure that there is *asystematicity* in the interests that inquirers bring to the table. This means democracy. It means making sure that no one racial or ethnic group, no one age group, no one gender – no group defined by *any* common interest – controls inquiry. We must have diversity among the people who do science, but also among the people who report the news, and among the people who make the news.

However, we mustn't go so far as to condemn interestedness in general. Certain kinds of bias play a foundational role in our affective and ethical lives analogous to that played by cognitive biases in the epistemic realm.
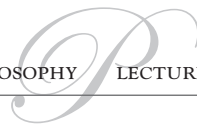
Ethical and Unethical Partiality

THE BIAS PARADOX, as I outlined it earlier, also emerges in the context of affection and conation as a conflict between the partiality of our concern for those we hold dear (including ourselves) and the apparently impersonal demands of morality. As has been argued by the

philosophers I cited earlier, neither deontological (rule-based) moral theories, nor conse-
quentialist (outcomes-based) theories seem to make room for the kind of focused and partial
concern that characterizes the best forms of friendship and love. Deontological theories, like
Kant's ("Act only on those maxims that one could will be to universal") seem to prohibit
any favoring of individuals whom one happens to love – all human beings should be treated
similarly. Consequentialist theories like utilitarianism ("Always choose the act which will
result in the greatest good for the greatest number") may permit you to act on your partiality
toward a loved one, but only accidentally, and for what are intuitively the wrong reasons.

A well-known thought-experiment illustrates the criticisms: suppose you are in a life-
boat with room for only one more, and suppose that there are two more survivors struggling
in the water: one is someone you love deeply, and the other is a perfect stranger. Many of us
have the intuition that it is at least morally permissible for you to choose to save your loved
one. But neither Kantianism nor consequentialism seems to give the right answer. Kantian-
ism, it seems, requires you to set aside your partiality toward your loved one, and instead, to
choose between the two survivors by some random process. (Maybe you still have a coin in
your pocket that you can flip.) Utilitarianism *might* allow you to save your loved one, but only
if that person's survival would contribute more to the general good than would the stranger's.
You are allowed, within this framework, to take into account the distress it would cause you
to abandon your loved one, but that "disutility" would have to compete with the overall
good effects that rescuing the stranger would have. (The stranger would stop global warming.
What would your lover do for humanity?)

The problem, as Annette Baier points out, is that human beings cannot conduct their
lives in ways that conform to these strictures. We have – most of us – a deep psychological
need for loving relationships, and such relationships are partly constituted by the reliability
of our partner's *partiality* toward us. This is easiest to see in the case of parents and children.
As we know from many tragic cases, children who are abandoned or orphaned or neglected
tend to be physically stunted, and psychologically damaged. Sometimes they die, from a con-
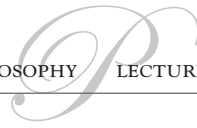dition with the offensively impersonal name "failure to thrive."

In theory, a Kantian could accommodate this datum about the developmental needs
of human children. Anyone charged with fostering the well-being of a child would, on the
Kantian view, be required to act in the ways necessary for the child's well-being. This would

involve more than simply feeding and clothing the child. The Kantian caregiver would also have to behave toward the child in ways that fostered its emotional and psychological health: smiling, cuddling, talking, and playing with the child. As the child grows up, the caregiver would need to show interest in the child's ideas and pastimes, (in terms of modern American parenting – attending their sports events and musical performances), and see to the child's moral development (partly by disciplining the child in appropriate ways).

This means two things, however. First, it means that a minimally adequate parent must devote a great deal of time and attention to *that particular child*. There are limits to the number of children that the average human being can successfully parent. (Some of us reach our limits with one or two.) So it seems to be inherent in the nature of parenting that some partiality is necessary – some adults must privilege the well-being of the children in their care over the interests of other rational agents (including, frequently, themselves.) That much, a Kantian can allow: one can have special duties in virtue of a role that one occupies. But the second challenging thing about the partiality that a successful parent must display is this: it is extremely *challenging* to be a good parent. It can be exhausting, painful, infuriating, frustrating, boring – not to mention financially costly. (And by the way, in case you're wondering, I love my kids. Really.) And *duty* is an inadequate motivator for most human beings to take on such a large bundle of other-directed responsibilities of such magnitude and for such an extended period of time. (Not to mention the liability to pain – watching one's child suffer, or God forbid losing one's child – these are some torments one risks in taking on the role of a parent.) For most of us – certainly for me – if all we had was a sense of duty to our children, we would fall terribly short in our responsibilities. What we need, then, in order for human beings to do what is necessary is some other powerfully motivating factor. And that, of course, is *love*.

Now I am not saying that we love our children *because* it enables us to do our duty towards them. We love our children because they are our children, just as we love our friends because they are our friends. Love is, to a large extent, an instinct, and one on which the survival of the species probably depends. But that does not mean that love is an illusion, or a disguised version of self-interest. Contrary to what some philosophers would have you think, such deflationary views of love do not follow from and are not supported by the theory of natural selection. Evolution cares that our offspring survive; it doesn't care how or why it
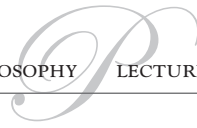
happens that their survival is ensured. My contention here is that the existence – and per-sistence – of *genuine* filial love could easily have co-evolved with our species' long period of juvenility. And if our ancestors ever did have to compete with hominids who operated only on the basis of self-interest – and *actual* competition with such beings is a prerequisite for love to be an *adaptation* – it's not surprising that we loving hominids won.

So it's a fact of human embodiment that we do not flourish and may not even survive but for the existence of highly partial emotional connections, just as it's a fact of human em-bodiment that we could not know much without the existence of native cognitive biases. But just as we asked what becomes of objectivity once we recognize our dependency on cognitive biases, we must ask what becomes of *justice* once we recognize our dependency on *affective* partiality.

I want to make an analogous answer, in keeping with the strategy I have outlined. *For affective partiality to operate in an ethically justifiable way, it must operate within the right kind of context.* And the "right" kind of context here is justice. When cognitive biases operate in a world shaped by injustice, they produce harmful beliefs. When affective biases operate in a world shaped by injustice they produce pernicious inequalities. *If there is already an unjust con-centration of wealth in the hands of a few, then the partiality of those few toward the members of their own tribe means that other people are not able to realize the benefits of affective partiality.* If there are no good public schools, if there is not adequate health care, if not every child is safe from war and violence, then the partiality I show my own kith and kin is morally unjustifiable. The solution to this problem is not to enjoin the affluent to stop favoring their own children; rather we must work to structure society so that *every* parent may offer *every* child the proper benefits of affective bias. If there really were "school choice" – if every parent had the ability to send their children to a school that was precisely tailored to their children's needs – there would be no problem about each parent looking out for his or her own.
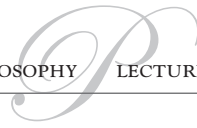
Kant actually made an argument something like this in the context of his theory of property. (John Rawls made the idea explicit, and spelled things out in detail, in *A Theory of Justice*.)[12] Here's the idea: Kant recognized that, given the fact of human embodiment, it is necessary in order for any of us to pursue our ends as rational agents, that we each are able to

---

12   Rawls, *A Theory of Justice*.

exercise exclusive control of a certain amount of space and of physical material. This meant that there was a "transcendental" justification for some kind of institution of *property*. But this institution would have to be set up in a way that was consistent with the demands of justice – with the categorical imperative, which dictated the inherent and equal moral worth of every person. That means that every person would have to have ensured the benefits of there being an institution of property.

In sum, there's nothing wrong with bias. We need bias; we cannot do with it. *What we must have though, are the structures and safeguards that create the environments in which biases can perform their constructive functions, without their potentially distorting effect.* Democracy and distributive justice!!! Make the human world safe for bias.

References

Antony, Louise. "Quine as Feminist: The Radical Import of Naturalized Epistemology."
    In *A Mind of One's Own: Feminist Essays on Reason and Objectivity*, edited by Louise
    Antony and Charlotte Witt. Boulder, CO: Westview Press, 1993.

Baier, Annette C. "The Need for More Than Justice." *Canadian Journal of Philosophy* 17, no.
    S1 (1987): 41–56.

Berkowitz, Bonnie, and Chris Alcantara. "The Terrible Numbers That Grow with Each
    Mass Shooting." *Washington Post*, May 9, 2021. https://www.washingtonpost.com/
    graphics/2018/national/mass-shootings-in-america/.

Chomsky, Noam. *Reflections on Language*. London: Temple Smith, 1975.

Gelman, Susan. *The Essential Child: Origins of Essentialism in Everyday Thought*. Oxford:
    Oxford University Press, 2003.

Hartig, Hannah, and Carroll Doherty. "Two Decades Later, the Enduring Legacy of
    9/11." Pew Charitable Trust. September 2, 2021. https://www.pewresearch.org/
    politics/2021/09/02/two-decades-later-the-enduring-legacy-of-9-11/.

Leslie, Sarah-Jane. "The Original Sin of Cognition: Fear Prejudice, and Generalization."
    Journal of Philosophy 114, no. 8 (August 2017): 393–421.

Parker, Kim. "Women More Often Choose Family Over Work." *The Atlantic*, October
    6, 2015. https://www.theatlantic.com/politics/archive/2015/10/women-more-often-
    choose-family-over-work/433029/.

Quine, W. V. O. "Epistemology Naturalized." In *Ontological Relativity and Other Essays*. New
    York: Columbia University Press, 1969.

Rawls, John. *A Theory of Justice*. Rev. ed. Cambridge, MA: Harvard University Press, 1999.