Tragic Flaws

ABSTRACT: *In many tragic plays, the protagonist is brought down by a disaster that is a consequence of the protagonist's own error, his or her* hamartia, *the tragic flaw. Tragic flaws are disconcerting to the audience because they are not known or fully recognized by the protagonist—at least not until it is too late. In this essay, I take tragic flaws to be unreliable belief-forming dispositions that are unrecognized by us in some sense. I describe some different types of flaws and consider what we might do about them. Then I examine three types of policies for managing our tragic flaws: doxastic, dispositional, and methodological.*

KEYWORDS: social epistemology, biases, ignorance, self-judgment, intellectual humility

There are some faults that are difficult to perceive, which have not been classified or determined, and which have no name.
— Joseph Joubert (1754–1824; *Pensées*, published posthumously)

Only small minds cannot tolerate being criticized for their ignorance. This is because, since they are usually quite blind in all things, quite stupid, and quite ignorant, they are convinced that they see clearly what their minds only see confusedly.
— Madeleine de Souvré, Marquise de Sablé (1598–1678; *Maximes de Mme la Marquise de Sablé,* published posthumously)

The fool doth think he is wise, but the wise man knows himself to be a fool.
— William Shakespeare, *As You Like It*, 5.1

Introduction

In some of Shakespeare's tragedies, the protagonist is brought down by a disaster that is a consequence of the protagonist's own error. This error is the protagonist's *hamartia*—the so-called tragic flaw. Macbeth's flaw was an overriding ambition and thirst for power; Hamlet's was his hesitation to act to

CrossMark

avenge his father's death. Tragic flaws are disconcerting to the audience because they are not known or fully recognized by the protagonist—at least not until it is too late. As the protagonist's self-made catastrophe unfolds, we watch with rapt attention or cover our eyes in horror, peeking a bit through our fingers. If the Shakespearean protagonists were destroyed by faults of which they were ignorant, oughtn't we worry about our own faults that we do not recognize?

That worry can be hard to shake when we reflect on our contested opinions about morality, politics, science, history, and philosophy. Our views could be distorted or skewed by invisible errors in judgment—and our opponents might insist that is so. What of this worry? Should it change how we inquire? The idea of tragic flaws can tell us something about intellectually responsible inquiry in light of our limited understanding of our own limitations.

As I speak of them here, *flaws* are epistemic or cognitive in nature, not moral. Epistemic flaws, I will stipulate, are dispositions or tendencies that produce suboptimal epistemic results for believers. For example, flaws are dispositions that prevent knowledge, justified beliefs, or other epistemic goods. The kinds of flaws I focus on are unreliable belief-forming or belief-maintaining dispositions. Flaws are not always manifested, but sometimes they undermine successful inquiry.

Take several examples of epistemic flaws. Being disposed to dismiss or ignore counterevidence without good reason is a flaw (Kelly 2008; Cassam 2019: ch. 2). That disposition systematically biases our total evidence in support of our pre-existing beliefs. Prejudice against testifiers from a particular social or ethnic group is a flaw (Fricker 2007) because that trait inclines us to ignore potentially valuable testimony. A tendency to be overly confident in our opinions is a flaw because it leads us to misjudge the import of our evidence. A tendency to see greater bias in other people than in ourselves is a flaw (Pronin 2007) because it leads us to attribute too much bias to other people and not enough to ourselves. Red-green colorblindness can be an epistemic flaw because it leads people to have difficulty forming correct perceptual beliefs. There are many other examples. Biases of judgment and reasoning are intellectual flaws in my sense of the term.

What makes a flaw tragic? This term of art is inspired by *hamartia*. In tragic drama, errors hidden by ignorance are commonplace. Commentators use the term *hamartia* to cover a variety of unwitting mistakes, hidden missteps, and

miscalculations.[1] But ignorance is a core feature of *hamartia*: protagonists fail to grasp or comprehend whatever brings their downfall. And the same holds for the notion of tragic flaws motivating my discussion. To a rough first approximation, a tragic flaw is an unrecognized flaw. It is not necessarily a vice of character nor a feature we should be blamed for, but it might be. And a tragic flaw in my sense need not literally lead to a tragedy—though it might. I say more about the notion shortly.

First, I will describe a number of ways in which a flaw can be unrecognized by us and thus tragic (section 1). Then I explore some ideas about the management of tragic flaws (section 2). It might appear that we are helpless to do anything about the fact that we have such flaws, but that is not so. I ask whether knowledge of our tragic flaws should change our confidence in our beliefs (section 3) or encourage us to develop special types of dispositions (section 4), or lead us to pursue inquiry differently (section 5).

The topic of tragic flaws calls for attention. Our intellectual efforts are obviously imperfect in ways we do not recognize, and we should hope to do what we can to avoid mistakes. The alternative is to concede that we cannot stop the invisible workings of our failure, allowing our unrecognized flaws to become the mainsprings of our intellectual downfall. Remaining ignorant of our flaws hands over to fate what we ought to try to control. We can do better.
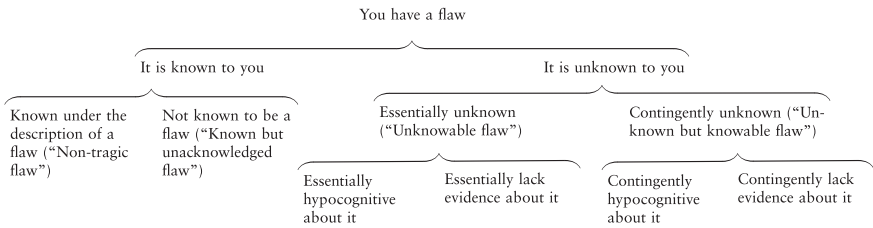
1. What Are Tragic Flaws?

I will unpack the idea that a tragic flaw is an unreliable belief-forming disposition whose unreliability is unrecognized.[2] Before I begin I will set aside one misleading way to think about flaws. In psychology, 'lay dispositionalism' is the idea that people's thoughts and actions flow from internal factors, such as their beliefs, values, and abilities, rather than from the situations they are in (Ross and Nisbett 1991). The lay dispositionalist understanding of flaws is that they are always internal to us and remain stable across situations. But lay dispositionalism is a mistake because often internal factors and situations both matter for explaining what we do. Thus, we should think our flaws can be partly fixed by situations, not only by internal factors. Take color blindness as an example. Color blindness is a

---

[1] There are disagreements over what exactly the term means. One twentieth-century commentator noted that 'there has been perhaps more scholarly blood spilled over the meaning of [*hamartia*] in Aristotle's *Poetics* than over any other concept in the history of literature' (Dyer 1965: 658). The Shakespearian notion of *hamartia* has to do with how character produces unrecognized errors in judgment and decision, whereas the Aristotelian notion is related to forces beyond protagonists' control, producing outcomes they do not anticipate, such as Oedipus's murder of his own father in Sophocles's tragedy. I do not use the term 'tragic flaw' as literature scholars use it though the stipulative notion of tragic flaws I develop has a Shakespearian tincture.

[2] The notion of (un)reliability admits of degrees. Suppose a perfectly or maximally reliable belief-forming process produces 100 percent true beliefs. If a process is not perfectly reliable, is it thereby flawed? That depends on our purposes. If we would regard less than perfect reliability as epistemically suboptimal in some situation, then depending upon such a process to form a belief is a flaw. I assume that our goals and interests can set a threshold for appropriate reliability. Thus, some imperfectly reliable processes—say, someone's visual perception under normal lighting conditions—will not count as flawed across all situations. (Thanks to Samuel Kampa for discussion.)

flaw in many normal perceptual situations, but we can create environments where colorblind people have advantages over those not colorblind. What counts as a flaw may depend on how things are in the world. That point about the nature of flaws is unsurprising given research on heuristics and biases—the 'fast and frugal' processes that give us accurate beliefs in one setting can backfire elsewhere (Gigerenzer 1991).

Again, tragic flaws are unreliable belief-forming dispositions whose unreliability goes unrecognized. But tragic flaws come in various types. To guide our discussion, I begin with a taxonomy of flaws, each one of which I will then describe and illustrate:

Take some particular flaw. This flaw is either known or unknown by the person who has it. If it is unknown to them, is it tragic? Yes. All unknown flaws are tragic flaws, as I use the term. That includes each of the flaws on the whole right branch of the taxonomy. But if a flaw is known, is it nontragic? Not necessarily. There are different ways to recognize a flaw, and some flaws known to a person are still tragic because they are, in some other sense, unknown. There is an important division on the left branch of the taxonomy.

How can a flaw be known to a person and yet unknown? Drawing some careful distinctions will make this less puzzling than it might sound initially. To begin, we can distinguish between flaws known under the description of a flaw and flaws known but not under such a description. Any flaws known under the description of a flaw I call non-tragic flaws. Flaws that are known, but not known as flaws, I call known-but-unacknowledged flaws. Someone can know he has a feature F, where F is a flaw, without knowing or even believing that F is a flaw. Someone can also know both that he has F and that F is a flaw.

Consider an example of a known-but-unacknowledged flaw. Someone can know she responds dogmatically to counterevidence, rejecting such evidence without good reason, but at the same time she will believe her dogmatism is a perfection and in no way a flaw. This person need not describe herself as 'dogmatic'. She could think she is 'settled' or 'unshakable' in her belief. Although this person is fully aware she is dogmatic (under another description), it seems the dogmatism is a tragic flaw nonetheless. It is the type of flaw that could easily catch her unawares, given she does not acknowledge its danger. We might wonder whether it is coherent for someone to believe both that she rejects evidence dogmatically and that doing so is in no way a flaw. How can someone believe that unreasonably ignoring counterevidence is epistemically good? That upside-down stance seems possible. Human beings can think their flaws are perfections, their weaknesses strengths,

their defeats victories. We are sometimes inclined to reverse good and bad. As the witches in *Macbeth* intone, 'Fair is foul, and foul is fair'.

Turn now to unknown flaws—the right side of the taxonomy. All unknown flaws are tragic, but the class of unknown flaws is diverse. We can distinguish between unknowable flaws and unknown-but-knowable flaws. The difference here turns on what is possible. Barbara Vetter's (2013) work on potentiality is helpful for illuminating the difference between unknowable and unknown-but-knowable flaws. The strength of someone's potential to know a flaw is fixed by all manner of intrinsic and extrinsic factors—her actual evidence, her actual concepts, her capacity to use those concepts, what evidence is available in her environment, how relevant things stand in the world, and so forth. These things are abilities, capacities, powers, or dispositions. Plausibly, the factors determining one's potential to know are like vectors that can add up. If there are no factors that position us to potentially know our flaw, it is unknowable. If the factors give us the potential to know a flaw, it is knowable, to some greater or lesser degree. Imagine there is a spectrum along which all of your unknown flaws are set out, ranging from unknowable to knowable. What determines the place of flaws on the spectrum is your potential to know them. For some flaws, the potential could be extremely remote. Perhaps you would know some flaw if a sorcerer waved his magic wand, bestowing upon you enhanced cognitive powers. If something like that sorcerer represents your best opportunity to know your flaw, it is unknowable for all intents and purposes. But if you can discover your flaw by, say, completing a test or asking a friend, it is knowable for you, even if presently unknown.

I will say more about the spectrum between unknowable and knowable flaws. First, begin with the unknowable flaws. Such flaws are so well-hidden that we cannot discover them in ourselves. Perhaps others can identify such flaws in us, but we will be blocked from learning about them through testimony. There are two critically different ways a flaw can be unknowable to us. I will draw out that difference by introducing a concept: *hypocognition*.

Hypocognition is, literally, a lack of cognition. Nobody has every concept. We cannot grasp every truth. As psychologists Kaidi Wu and David Dunning note, 'people's finite conceptual horizons are a pervasive and powerful constraint on how they make sense of the world. These horizons represent the hard boundaries of where people's possible interpretations of their circumstances can go and define the finite channels into which their understanding is funneled' (2017: 1). Hypocognition abounds (see Wu and Dunning 2017 for a review). Experts grasp many concepts that novices and nonexperts do not and potentially cannot grasp. Outside of academic and technical fields of knowledge, lack of lived experience robs people of a basic conceptual grasp. For example, Wallace Stegner wrote, plausibly, that 'home is a notion that only the nations of the homeless fully appreciate and only the uprooted comprehend' (1971: 159). We all lack expertise and experience in a great many domains, and so we are hypocognitive about what is in them. Further, humanity is hypocognitive about domains that are undreamt of within its finite conceptual systems.

Let me describe one striking instance of hypocognition. In a 1970 article, Cambridge University neuroscientists Colin Blakemore and Grahame Cooper

described how they used newborn kittens to study brain development. The kittens were subjected to different types of visual experiences inside specially designed cylindrical tubes. The interior of each tube was covered with black and white stripes and had a clear glass bottom. One type of tube had only vertical stripes and another type had only horizontal stripes. The researchers divided the kittens into two groups: 'vertical' kittens and 'horizontal' ones. For five hours daily, the 'vertical' kittens were placed inside the vertical tubes, and the 'horizontal' kittens were placed in theirs. The kittens were fitted with collars that prevented them from seeing their own bodies. And so, during the experiment, the vertical kittens looked at only vertical lines and the horizontal kittens looked at only horizontal lines. When the kittens were not inside their striped tubes, they lived in pitch darkness.

Then, at age five months, the kittens were moved into an ordinary perceptual environment with tables and chairs. The experiment had rendered the kittens 'virtually blind' to lines running in an orientation opposite to the lines in their tubes (Blakemore and Cooper 1970: 478). Horizontal kittens recognized the seats of chairs perfectly well and curled up on them for a nap, but when they scampered around the room, they would crash into chair legs. Vertical kittens could navigate around chair legs but could not find a cozy place to nap. In one revealing test, Blakemore and Cooper compared the reactions of two cats simultaneously (1970: 478). When an experimenter held a rod vertically and moved it back and forth, the vertical cat would visually follow the rod and play with it. When the rod was rotated horizontally and shaken, however, the horizontal cat would suddenly pursue the rod while its vertical companion would now ignore it completely. The researchers had reared hypocognitive cats.

Like those visually impaired felines, our cognitive grasp of reality is limited. We are hypocognitive about many things—our own flaws included. Some of our flaws are so carefully hidden from us that we cannot recognize them.

Unknowable flaws due to essential hypocognition can befall us in a couple of ways. First, some flaws can only be accessed using concepts that are too complex or subtle for our thinking. On its face, the idea that we have or could devise the concepts necessary to describe all of our flaws is doubtful. Voltaire suggested we can recount the history of human opinion as 'scarcely anything more than the history of human errors' ([1764] 1901: 61). People get a lot wrong much of the time. But our errors are diverse, and we still have something new to learn about them. Second, maybe it is (practically) impossible for us to self-attribute certain flaws because of ego-defenses. We are primed to think of ourselves as good, reasonable people, and we can't bear the thought that *we* have those flaws. We readily acknowledge other people have them, but never ourselves. This self-protective posture might not be driven by hypocognition; potentially, we possess the relevant flaw concepts but just cannot apply them to ourselves. I also suppose we could lack the concept '*I am flawed in some way*'—even when we grasp both that flaw's nature and that others could have it. Then hypocognition would be the source of our inability to judge ourselves flawed. Insofar as self-judgment is fraught, it will be unsurprising that we cannot attribute some flaws to ourselves because we lack relevant self-concepts.

You might suspect I have exaggerated the difficulty. You might challenge me to point out just one flaw about which you are essentially hypocognitive. Could I meet

your challenge? Certainly not. If you could see one such flaw, it would not be essentially unknowable. But we can all sneak up on such flaws, as it were—knowing that they exist without knowing exactly what they are. The distinction here is simple. We can know that we have a feature *F* and that *F* is a flaw, but we cannot determine which one of our features is *F*. That is existence knowledge of a flaw. By contrast, we have identification knowledge when we know, of some particular feature, that it is a flaw.

One general point about knowledge is that we can know we have a false belief without knowing what it is. That is just existence knowledge without identification knowledge. And we can know an essentially unknowable flaw, in the sense of having existence knowledge of it, even though we can't by definition have identification knowledge of it. Consider how this could be true. First, we could start with observations of other people's unknowable flaws. We see others have blind spots they cannot ever see. We can infer we are like them: we have blind spots we cannot ever see ourselves. In my experience, people appear to be essentially in the dark about some flaws. Second, facts about cognitive development suggest how we could easily lack identification knowledge of unknowable flaws. Toddlers are essentially hypocognitive about the biases of judgment and reasoning. They cannot understand the notion of confirmation bias, the fundamental attribution error, the bias blind spot, and so on. That is true even though toddlers' thinking is influenced by those biases. Children have intellectual flaws they cannot grasp because their grasp is immature. An analogy suggests our own predicament as adults: we are like little children when it comes to some of our flaws and can't so much as understand what they are.

Some unknowable flaws are due to what I call essential hypocognition. That is not the only type of unknowable flaw. Another type can be grasped but remains (practically) unknowable because someone cannot acquire evidence sufficient to gain identification knowledge of it. Unknowable flaws like these have been discussed throughout the history of skepticism. In his *Meditations*, for example, Descartes made the Evil Genius a famous personage, at least among students of philosophy. Being systematically deceived by the Evil Genius is a flaw anyone could have because that property leads us to form problematic perceptual beliefs. The Evil Genius's cunning scheme might be impossible for us to detect, and as a result, we might never know our flaw. In that case, the flaw would be unknowable, though in a different sense than we have considered so far. It is not unknowable due to hypcognition. We can grasp the idea of the Evil Genius perfectly well, but we just can't acquire evidence to know that we are deceived.

The Evil Genius shows how an unrecognized flaw can lead to a mistaken belief. But similar flaws can also lead to the truth. Imagine you have an Epistemic Guardian Angel who always arranges circumstances so that whenever you make a competent effort to know some contingent proposition, you are rewarded with a true belief (Greco 2012: 10). But imagine it is impossible for you to detect your Guardian Angel's activity. Now suppose you competently reason to a false belief, and your Guardian Angel covertly changes the world, ensuring your belief is actually true. Your justified or competently formed true belief falls short of knowledge due to the coincidence or luckiness of the Guardian Angel's intervention. This amounts to a Gettier case (and provides a counterexample for

some versions of reliabilism). Here, being systematically guided to truths by the Guardian Angel is a tragic flaw for you—not because it leads you to a falsehood, but because you are 'Gettiered' and thus do not know a truth you are justified to believe.

Our lack of cognition or evidence or both might make some of our flaws unknowable, but recall that I earlier distinguished between unknowable flaws and unknown-but-knowable flaws. Among unknown-but-knowable flaws, we can mark another distinction. Our flaws might be unknown to us because we are contingently hypocognitive about them or contingently lack evidence to know them.

It is easy to admit that we have—or have had in the past—unknown-but-knowable flaws. Sometimes they move from being unknown to known, and when they are known under the description of a flaw, they become nontragic. But none of us can seriously entertain the idea that we recognize all of our flaws.

Color blindness is an intriguing example of an unknown-but-knowable flaw, showing how elusive evidence of flaws can be. The condition was first investigated in the late eighteenth century by the English chemist John Dalton. For years, Dalton had found the nomenclature of colors somewhat confusing, and had 'often seriously asked a person whether a flower was blue or pink, but was generally considered [by them] to be in jest' (1798: 29). Dalton was in his late twenties when he observed a pink geranium flower and noted it appeared to him to be 'an exact sky-blue' (1798: 29–30). He queried some friends about the flower and found his vision was unlike theirs. He had a chromatic visual abnormality. (Dalton's brother also reported seeing a blue flower—an early clue that color blindness runs in families. For more on Dalton's eyes, see Hunt et al. [1995].) Before Dalton's investigation, reports of color blindness are surprisingly uncommon. Yet the condition is relatively widespread in some populations—for example, red-green color blindness affects perhaps 8 percent of males and 0.5 percent of females of Northern European descent (National Eye Institute 2015).

During the nineteenth century, scientists and physicians investigated color blindness, and as knowledge of the condition circulated, activists wanted to protect society from its dangers. Imagine a colorblind railroad engineer, operating his steam engine, riding atop a few hundred tons of cast iron at 60 miles per hour. The engineer watches for signal lights and colored semaphores. Green means go and red means stop—unless he's colorblind, and then all bets are off. By the last quarter of the nineteenth century, after reports of railroad incidents caused by color blindness, railways instituted vision screening for engineers, brakemen, stationmasters, and signalmen.[3] Yet even to the present day, color blindness can make railroading risky. In Oklahoma in 2012, two Union Pacific freight trains crashed head-on, killing three people onboard, causing nearly $15 million in

---

[3] After an infamous head-on collision between two trains in Sweden in 1875, many railways adopted vision screening for employees. Contrary to standard reports about the crash, it is unclear whether color blindness caused the accident. Following the accident, however, a Swedish ophthalmologist convinced railway officials to administer a color-vision test to all employees. This test was one of the first psychological tests performed on a large group. For further details about the Lagerlunda crash and its aftermath, see Mollon and Cavonius (2012).

damage when diesel fuel tanks ruptured and exploded. The National Transportation Safety Board (2013) found that one engineer, who had a history of color vision difficulties, had been unable to visually detect crucial wayside signals.

Color blindness is a type of a flaw—not in any way a character flaw, to be sure, but a belief-forming disposition that is unreliable in particular circumstances. It can be known yet often goes unrecognized. One reason why that is so traces back to hypocognition: some colorblind people may be unable to grasp the nature of their visual capacities. Before the nineteenth century, everyone lacked the language and testing instruments to understand color blindness. A second reason why color blindness may be unrecognized traces back to a lack of evidence concerning the flaw even when people can conceptualize it. There are many other types of flaws, of course, and our contingent ignorance about many of them is a fact of life. Some of those flaws are ones we can at least hope to know about if we equip ourselves with more concepts and more evidence.

Thus far, I have suggested we can know tragic flaws exist even when we cannot identify them. I also described a taxonomy of tragic flaws. Some tragic flaws are unknown to us, either because they are unknowable by us or because we have not come to know them. Other tragic flaws are known to us but not acknowledged: we are aware of them but not under the description of a flaw. I should add that what I call tragic flaws are not the only type of flaw—and various other flaws do not fit into the taxonomy because they are *not* unrecognized in some or other relevant sense.[4] Being unrecognized is an essential part of the sort of Shakespearean *hamartia* I am exploring.

All of this raises a question: Given that we are not completely ignorant about our tragic flaws, what shall we do?

2. Dealing with Our Flaws

It may seem to be impossible for us to do anything consciously about a problem that by definition we do not recognize as a problem. After all, tragic flaws are so troubling because they are unrecognized, and so we appear helpless to stop them from thwarting our inquiries. But if that is correct, isn't the topic of tragic flaws practically irrelevant and not worth another thought? Maybe we are tragically flawed—but so what? Admitting that we are flawed in that way, someone might insist, can make no possible difference for how we ought to lead our intellectual lives.

That idea is tempting but wrong. People regularly find out what they did not recognize before. Learning happens. One upshot is that we can, at least in principle, find ways to do something about both known-but-unacknowledged

---

[4] Here is just one example. Someone can know about a flaw—and even know it under that description—while lacking an appropriate intellectual responsiveness to that recognition. Suppose, for example, you know about confirmation bias and realize it is a problem whenever you think about politics. But let us imagine your thinking about politics is unresponsive to your knowledge about confirmation bias even when you are directly reflecting on it. Your knowledge of your flaw is in an important sense 'inert' or 'idle'—your knowledge fails to influence your thinking as it ought to. We can call this a known-but-disregarded flaw. Known-but-disregarded flaws are a problem because they can lead our judgment astray even when we fully recognize what is going on. But because such flaws are recognized, I will not treat them as tragic flaws. (Thanks to Alex Arnold for discussion.)

flaws and unknown-but-knowable flaws. We can possibly acquire new knowledge that shines a spotlight on those flaws, rendering them non-tragic.

Let us not forget unknowable flaws, though. Is it impossible for us to do something about unknowable flaws because they are by nature unrecognizable by us? I don't think so. There are ways to address tragic flaws that do not involve making them nontragic. Sometimes unknowable flaws might be counteracted or even eliminated by shrewd planning. Suppose that other people can recognize my unknowable flaws. By definition, I cannot have identification knowledge of those flaws, but plausibly I could acquire existence knowledge of them by testimony (Cassam [2019: 149 and 155–57] discusses the possibility of gaining self-knowledge of epistemic vices by testimony.) And I could authorize other people to implement correctives. To be sure, their fixes might strike me as utterly wrongheaded—we are assuming these flaws are unrecognizable by me. But sometimes I can bind myself to doing as they say. Indeed, it has been noted by social scientists that techniques for debiasing and mitigating the bad consequences of biases are often launched at the social or institutional level (Heath, Larrick, and Klayman 1998; Kenyon 2014). Our communities could help, and I will soon say more about that possibility. An epistemic benefit of long-term relationships, such as marriage, might be that they give people a mechanism and motivation for mitigating flaws not recognized by one of the parties. In our endeavors, one good reason for joining forces with people who hold different viewpoints than ours is simple: we might help each other overcome our unknowable flaws.

But let us set aside unknowable flaws for now. Instead, I will focus on contingent tragic flaws that might become nontragic. Even if we cannot eliminate such flaws, it is presumably better for us to recognize them than not. That is to say, non-tragic flaws are better than tragic ones, at least from the perspective of seeking truths and avoiding errors. Tragic flaws can influence our inquiry in untoward ways, leaving us clueless about our shortcomings, but non-tragic ones can sometimes be counteracted or eliminated.

So, again, what can we do? One idea comes from Quassim Cassam, who introduced the idea of 'stealthy vices' (2019: ch. 7; 2015). These are epistemic vices that obstruct their own detection. Cassam gives the example of a completely closed-minded person who does not know she is closed-minded; self-knowledge of her vice needs a modicum of open-mindedness she lacks (2019: 146). Cassam notes other stealthy vices can include arrogance and dogmatic belief in one's own rightness. We can add to Cassam's list the trait of being a jerk, following Eric Schwitzgebel's observation that being a jerk is an epistemic vice that 'works to prevent its own detection' (2020)—jerks don't know they are jerks because they are jerks. So, what is the relationship between stealthy vices and tragic flaws? All stealthy vices are tragic flaws: they are unreliable belief-forming dispositions whose unreliability is unrecognized. But not all tragic flaws are stealthy vices. That is because many tragic flaws are not in any way implicated in their being unrecognized by us. In other words, the hypocognition or lack of evidence that makes a flaw tragic does not necessarily flow directly from that flaw itself.

Cassam discusses the possibility of overcoming or outsmarting epistemic vices. Generally, he is moderately optimistic about managing non-stealthy vices (2019:

170–71). But he does not say much about dealing with stealthy ones. His main example is a military officer who failed to prevent an impending attack; this officer had been closed-mindedness toward mounting evidence that forewarned of an invasion (2019: 183). The officer did not recognize his own closed-mindedness —it was a tragic flaw. But then Cassam imagines that the officer, after the disastrous attack, could undergo a kind of 'breakthrough' experience of failure, '[recognizing] the need for remedial action' (2019: 184). This breakthrough experience leads the officer to reflect on the skills he needs to improve, with a corresponding boost in open-mindedness. Cassam comments on the process:

> There are indications here of a kind of virtuous circle, with greater open-mindedness at each stage facilitating further remedial action against closed-mindedness. The initial breakthrough is provided by the experience of failure, where the nature and magnitude of the failure leave no room for doubt as to its origins in the subject's intellectual character. (2019: 184)

Cassam describes the officer's experience as traumatic (2019: 158–65)—a kind of 'sudden, unexpected, potentially painful event' (2019: 158). Traumatic experience of failure may surely reveal to us some tragic flaws, by giving us new evidence or concepts or by priming us to make better use of the evidence or concepts we already have.

I do not see how this possibility helps us much. First, we are not normally in control of 'breakthrough' experiences of failure. When disaster strikes, it is not as though we can arrange things in just such a way that the tragic flaw responsible for the trouble will be revealed. Second, history gives us many examples of people who failed catastrophically and yet learned no lesson about their flaws, or learned the wrong lesson. Cassam's military officer may not be so representative—or, at any rate, some antecedent conditions must be met for traumatic experiences to have an enlightening effect.

If we want actively to do something about tragic flaws, waiting for traumatic experiences of failure is not it. What could help? In what remains of this essay, I will examine three answers: self-doubt, intellectual humility, and what I call 'double-loop' inquiry, a special type of thinking strategy that scrutinizes our methods and assumptions. Each answer is a policy for doing something, suggesting ways to be more intellectually responsible in light of our admission that we are tragically flawed.

3. More Self-Doubt?

If we admit we have tragic flaws, perhaps we should lower our level of confidence in our beliefs. The idea is that recognizing we have these flaws calls for a healthy dose of self-doubt. Plausibly, knowing we have some such flaw provides evidence of our fallibility. That evidence could prompt us to hold almost any of our opinions more tentatively. We must be careful about the scope here, however. For some opinions, our tragic flaws might not be relevant. Consider the Augustinian 'cogito': *si fallor,*

sum (if I am mistaken, I am) (Matthews 2005: ch. 5). Admitting that I am tragically flawed should not lead me to doubt that I exist. If I have flaws, I exist. But for most or at least many other opinions, should my acknowledgement of my tragic flaws not induce in me a bit of self-doubt?

The claim that existence knowledge of tragic flaws undermines confidence in most of our beliefs is underdeveloped, if not false. Consider why. I believe I have two hands. I concede it is possible I am tragically flawed in virtue of being systematically deceived by an Evil Genius. My belief that I have hands could be wrong because my sensory experience could be the product of the Evil Genius's trickery. But why should the mere admission of a possible flaw change my level of confidence about my basic perceptual beliefs? It appears that I should be no less confident in my beliefs after I reflect on the mere possibility of an Evil Genius than before I do. Of course, if I have positive evidence to believe that I am systematically fooled by an Evil Genius, I should doubt some of my beliefs. But knowing about the bare possibility of such a flaw—again, a flaw that is unrecognized by me—does not seem to require me to reduce my confidence in any perceptual belief.

Try a more promising idea. When I admit that I have tragic flaws in a domain, I acknowledge reasons to believe that I currently neglect or overlook evidence about my limitations in that domain. There are flaws I am subject to but, for whatever reason, I am ignorant of them. That admission is a kind of 'higher-order' evidence. It is evidence that tells me about my competence or ability to respond appropriately to relevant evidence in the domain. This higher-order evidence should lead me to temper my assessments of my opinions (see Feldman 2005; Kelly 2010; Christensen 2010; Sliwa and Horowitz 2015; Ballantyne 2019).

An illustration is in order. Imagine Earhart is piloting a small aircraft above 10,000 feet. Earhart knows that people in small aircraft flying at high altitudes often suffer from hypoxia—a condition where the brain is oxygen-deprived—and, as a result, their judgments become unreliable (a similar example is proposed by Elga [unpublished] and discussed in [Ballantyne 2019: 139–42]). Once hypoxia has taken effect, it will typically seem to the hypoxic subject that her reasoning is good even when it is bad. Hypoxia does not 'leave a trace' in consciousness. Because Earhart recognizes she might be hypoxic at this high altitude, she has reason to invest much less confidence in her calculations about her aircraft's remaining fuel. Her reason here is a kind of higher-order evidence. She has learned she is not competent in making judgments about her fuel.

Learning you might be hypoxic is like learning you might be tragically flawed. If you have hypoxia, everything seems fine with your judgment and reasoning, but it is not. Therefore, you should adjust your confidence downward insofar as you suspect you are hypoxic. Likewise, if you have tragic flaws, everything seems fine with your judgment and reasoning on some topic, but it is not. Therefore, you should adjust your confidence downward insofar as you suspect you are tragically flawed.

The point is a general one. When we see there are things we do not see, we should think seeing is not everything. We should try to balance the observation that we are missing something about our flaws against the weight of what we do see. In other words, if we know that what we do not see is evidence indicating we are

overconfident in light of what we do see, that insight should moderate our self-judgment.[5]

Arguably, our acknowledgement of tragic flaws in some domain can induce self-doubts about our competence there. The rationale for those doubts is reminiscent of so-called precautionary principles, which encourage restraint in situations when there is risk to human health or the environment (Jordan and O'Riordan 1999). We sometimes know there is a risk but do not know how significant it is or what factors exacerbate it. But before all of the information comes in, precautionary principles tell us to act to mitigate risk. Along similar lines, my suggestion is that when we recognize that we have tragic flaws in a domain, we should take caution in advance of finding out what our flaws are and thus lower our level of confidence in our beliefs.

All of this suggests there might be plausible norms or principles that prompt us to lower our doxastic confidence in view of our tragic flaws. But such a policy faces difficulties of application. First, it is unclear how much doubt is required by our concession of tragic flaws. I do not have a sense for how to understand, in qualitative or quantitative terms, how much reduction of confidence is called for. We should not forget Prince Hamlet's chronic indecision and self-doubt. Experiencing self-doubt in view of tragic flaws could itself be a tragic flaw—and too much self-doubt is a common problem. How much is appropriate? Even if we can answer, a second issue is that our knowledge of tragic flaws is tenuous. We know of them but cannot identify them. It is easy to admit as a general, abstract matter that we are tragically flawed, but hard to admit we are flawed in some specific way. Observe how this works. On the one hand, we can readily acknowledge that we have some tragic flaws relevant to our beliefs about the vast range of topics in the intellectual world, from politics to economics to religion. Our thinking in that vast domain is doubtless hampered by some unreliable belief-forming dispositions. On the other hand, are we equally willing to concede that we have tragic flaws relevant to our thinking on some topic about which we have views? Consider the morality of capital punishment or the justice of some taxation policy or the existence of God. Are we tragically flawed in ways that could prevent us from seeking truth and avoiding error concerning those matters? Here we may feel less inclined to confess that we have tragic flaws. Acknowledging our hidden flaws is easier when nothing in particular is at stake.

Notice why that matters. The idea behind tragic flaws is that we do not know what they are. Our lack of specific knowledge prevents us from articulating a convincing argument for greater self-doubt about some particular belief. Of course, whenever we recognize that we are novices and nonexperts about domains where there are legitimate experts and authorities, that tips us off about the appropriateness of self-doubt. But then perhaps it is not clear whether our recognition that we have tragic flaws itself induces our self-doubt or whether what matters is another factor, such as recognized disagreement from experts or relevant evidence we know we do not possess.

---

[5] Ballantyne [2015] and King [2016] defend the following idea: learning there is evidence we lack about a topic can be evidence that influences what we should believe about that topic.

My hunch is that reflecting on our tragic flaws should sometimes induce in us greater self-doubt about some beliefs, beyond the truism that we are not infallible. But I am unsure how to work out that idea in detail. Let me try something different.

4. More Humility?

Once we recognize that we are tragically flawed, we may seek to develop a type of intellectual character that can compensate for our unrecognized flaws. If tragic flaws are sparks near a dry forest, then a trait like intellectual arrogance is gasoline. Mix together tragic flaws and arrogance and then brace yourself for an epistemic catastrophe. Plausibly, if we cannot help but be tragically flawed about some matters, we should aim to develop dispositions opposed to arrogance. A suitable type of character may make more of our flaws nontragic. This is something we can do to manage our unrecognized flaws.

What trait is the opposite of intellectual arrogance? One plausible candidate is intellectual humility, a notion philosophers and psychologists have examined in some detail recently. Researchers have not found a consensus about the nature of intellectual humility, and it appears that there are distinct notions marching under the same banner (Ballantyne, forthcoming).

According to the most influential account of intellectual humility, articulated by Dennis Whitcomb, Heather Battaly, Jason Baehr, and Daniel Howard-Snyder, the trait 'consists in proper attentiveness to, and owning of, one's intellectual limitations,' where 'owning . . . consists in a dispositional profile that includes cognitive, behavioral, motivational, and affective responses to an awareness of one's limitations' (Whitcomb et al. 2017: 520 and 518; cf. Haggard et al. 2018; cf. Leary et al. 2017). The humble person 'owns' her limitations in something like the way a team 'owns' its loss on the field.

We have limitations we do not recognize. But what do tragic flaws have to do with the limits-owning account of intellectual humility? Is the humble person on this account supposed to be attentive to and 'own' tragic flaws? It is not immediately clear. On the one hand, the phrase 'one's intellectual limitations' could be taken to mean 'one's actual limitations,' including tragic flaws. I will call that the objective interpretation of the limits-owning account. Alternatively, the phrase 'one's intellectual limitations' could mean 'the limitations one's evidence indicates one has' or, equally, 'one's recognized limitations'. I will call that the subjective interpretation.

Do Whitcomb and colleagues settle the matter about their intended meaning? They seem to embrace the subjective interpretation. This can be seen in the way they liken people's errors to blips on a radar screen—the arrogant person ignores the blips when they ought not to be ignored whereas the humble person takes the blips seriously when they should be taken seriously (2017: 516–17). Whitcomb and co-authors emphasize that humility requires attentiveness to and owning of *recognized* limitations. As I have already noted, it is unclear how we can be attentive to tragic flaws. We have existence knowledge, not identification knowledge, of them. I am asking about whether we can 'own' tragic flaws by doing something to address them. What sort of cognitive, behavioral,

motivational, and affective response is appropriate given our awareness that we are tragically flawed? That is the question, but I cannot see how humility as limits-owning offers us guidance. While this account of humility is certainly consistent with the idea that we ought to do something about such flaws, it is silent about what that something is.[6]

For all I have said, intellectual humility, open-mindedness, or some trait like these may help us manage tragic flaws. Humility or open-mindedness influences information processing. These dispositions help us give appropriate credit to contrary evidence or to seek out such evidence. When some of the incoming evidence suggests that we are flawed, it is plausible that we will be more inclined to take such evidence to heart if we are more humble or open-minded than not.

But even if throwing a spotlight on our hidden flaws is more likely when we are humble or open-minded, what are the potential mechanisms for making that happen? I do not know how stable dispositions or virtues such as humility reveal the dynamics of managing our tragic flaws. How can we humbly *try* to overcome those flaws? Let me consider a third type of policy that reveals some possibilities.

5. Double-Loop Inquiry?

One of the aims of inquiry is to recognize and correct our errors and limitations. How can we do that? To conclude this essay, I will consider one type of model for learning about errors and then suggest what it tells us to do about tragic flaws.

A twentieth-century British psychiatrist and cyberneticist named Ross Ashby thought about how organisms and technical systems adapt to new situations. One of Ashby's illustrations was an autopilot in an airplane (Ashby 1960: 108; Umpleby 2009: 234). An autopilot system is designed to maintain the airplane's stability, but a technician could install the autopilot incorrectly by mixing up some wiring. If an autopilot with faulty wiring kicks in, the airplane will be imperiled. Ashby envisioned an 'ultrastable' autopilot system that could detect when crucial variables exceed their limits and then rewire itself until the aircraft stabilizes. This autopilot features two feedback loops: one loop operates by making small corrections to the aircraft's flight path, whereas the second loop operates by changing the functioning of the system, whenever the system's 'essential variables' get out of whack.

Ashby's idea of an 'ultrastable' feedback system influenced an American organizational researcher named Chris Argyris. Argyris distinguished between what he called 'single-loop' and 'double-loop' learning (1976; 2002). Double-loop learning occurs when we correct errors by modifying our 'governing values',

---

[6] Perhaps the limits-owning account suggests humble thinkers need to 'own' the following fact: that they have unrecognized flaws. Consequently, humble thinkers will tend to avoid holding beliefs about which flaws they are *not* subject to and thus remain more open to self-attributing a wider range of flaws than non-humble thinkers would tend to be. While it is possible to 'own' the fact at issue, how does doing so help us do anything about tragic flaws? Even if we are open to the possibility that we have some flaws we can conceptualize, we surely remain hypocognitive about many others, as I suggested. Furthermore, it is unclear how being open to having some unrecognized flaws obviously makes us less likely to hold epistemically suboptimal beliefs because of them. This sort of 'owning' strategy does not appear up to the task.

whereas single-loop learning occurs when we correct errors without modifying our governing values. Those values are assumptions, sometimes unconscious, about what we want to do and what is possible for us to do, and Argyris and others have applied this model of learning to study organizations and group behavior.

I want to use the idea of double-loop learning, but I will employ it a bit differently than Argyris and others. I focus on inquiry, not learning. Inquiry is any attempt to answer a question using evidence. Suppose I wonder to myself: Is it raining outside? I can inquire by looking out my window, by reading a weather report, or the like. Single-loop inquiry involves trying to answer a question exclusively using evidence that bears directly on that question. For the question 'Is it raining?' my single-loop inquiry involves collecting and interpreting evidence about rainfall. Double-loop inquiry, on the other hand, is a more complex process. It involves trying to answer a question using evidence that bears on that question as well as evidence that bears on a further question: Are my methods to answer that first question appropriate or reliable? (I intend that question to encompass both evaluation of methods and the ways those methods are deployed. Even if some method is generally reliable, it is not reliable when deployed incorrectly.)

To illustrate, suppose I wonder whether it is raining. If I try to answer by looking out the window, double-loop inquiry could involve asking whether my eyesight is good enough to detect rain, asking whether the outdoor light level is sufficient to reveal rainfall, and so forth. In double-loop inquiry, we turn to scrutinize the methods and evidence we use to answer a question. (I should add that there can be double-loop inquiry about double-loop inquiry. In such an inquiry, there is a first-order question whether proposition p is true and a second-order question whether our methods for inquiry whether p is true are appropriate. In investigating the second-order question, we can ask whether our methods for answering precisely that question are appropriate.)

If we want to do something to manage tragic flaws, I suggest we engage in double-loop inquiry. It is the sort of activity that can, at least under favorable circumstances, make unrecognized flaws recognized. Consider an example of an unknown-but-knowable flaw. Magoo is looking out the window to try to determine whether it is raining, but he is nearsighted and does not know it. Even when it is raining, Magoo will not see rain. But he can scrutinize his visual methods and perhaps come to recognize his flaw. Magoo should visit an optometrist.

Pursuing double-loop inquiry seems humble. What is the relationship between double-loop inquiry and intellectual humility? They are not equivalent: the former is a process or strategy for inquiry whereas the latter is widely thought to be an intellectual character trait or epistemic disposition. Engaging in double-loop inquiry, just once or even habitually, falls short of being disposed to 'own one's limitations', in the sense of one prominent account of intellectual humility (Whitcomb et al. 2017). Interestingly, some other accounts of humility say the virtue helps people to form epistemically appropriate beliefs about the epistemic status of their beliefs (Hazlett 2012: 220; Church and Barrett 2016: 69). If Magoo believes it is raining and he is intellectually humble, he will be disposed to form an appropriate belief about the epistemic status of his first-order belief that it is raining. Magoo could use double-loop inquiry to reach a proper higher-order

belief about the status of his rain belief, but he could also come by that higher-order belief another way, such as by directly introspecting his first-order belief's grounds (a process that need not involve his scrutinizing his methods or evidence). Thus, even on the account of humility at issue, it is possible to be humble without ever executing a double-loop inquiry. Double-loop inquiry is one cognitive strategy among others that could promote humble dispositions, but the relationship between this strategy and the trait is not tight. That is good news for less-than-humble thinkers, because they might still choose to practice the strategy in spite of their dispositional arrogance.

Double-loop inquiry, at least when pursued properly, may show us that we are not as knowledgeable or wise as we had thought. For various reasons, we are not always motivated to engage in double-loop inquiry. Think of dogmatism and fundamentalism in all of their varieties. People are content to feel confident in their opinions without scrutinizing how they reached those opinions. Whoever digs at the foundation of a building risks collapsing it, and so people are often motivated not to jeopardize their deepest convictions. We do not want to topple our self-image, our social lives, or maybe even our livelihood. As Upton Sinclair noted, 'It is difficult to get a man to understand something, when his salary depends on his not understanding it' ([1935] 1994: 109).

We are constrained by what Chris Argyris called 'defensive routines' (1985). These are thoughts and behaviors that protect people's assumptions about themselves and the world—ideas and mechanisms that prevent us from 'experiencing negative surprises, embarrassment, or threat' (Tagg 2007: 39–40). We have ways to avoid the 'breakthrough' experiences of failure that Cassam mentions. We often avoid forthright feedback about our inquiry and instead seek feedback from sources guaranteed to tell us that we have done well. One cautionary example is the Nobel Prize winning chemist Linus Pauling, who could be supremely overconfident in some of his opinions, including the erroneous idea that megadoses of vitamin C effectively treat cancer. The molecular biologist James Watson noted that Pauling's fame made others 'afraid to disagree with him,' adding that Pauling could only speak freely with his wife, 'who reinforced his ego, which isn't what you need in this life' (1993: 1813). Pauling's insight into his intellectual limits was apparently impoverished. Long-term relationships can be a mechanism to manage tragic flaws, as I noted earlier, but they can also be part of the problem.

In any case, well-executed double-loop inquiry can expand the information we get about our flaws. In the 'postmortem' debriefing, people who have failed at a task are interviewed with the aim of revealing what went wrong and what could be improved. In the 'premortem' discussion, people imagine that a hypothetical plan goes sideways and then discuss measures that could prevent the outcome. These exercises stimulate double-loop inquiry. As I think of it, double-loop inquiry is a kind of *thinking against oneself*.[7] The normal direction of inquiry moves outward from the mind to the world, but double-loop inquiry invites investigation back home, asking about the suitability of our methods and our

---

[7] This is a phrase I first happened upon in an essay by the American essayist Phillip Lopate, but I later learned it goes back to the Romanian pessimist, Emil Cioran.

'governing values' as inquirers. To be sure, thinking against oneself will not help much if we are oblivious to what we are—what would we be trying to think against exactly? Plausibly, many tragic flaws involve a lack of recognition about our nature as inquirers.

Double-loop inquiry can be helpful for at least two reasons. Most obviously, it can render tragic flaws non-tragic when it gives us the evidence and the cognition required either to know our unknown flaws or to acknowledge our known-but-unacknowledged flaws. For example, the chemist John Dalton discovered his color vision was abnormal by engaging in double-loop inquiry; under controlled conditions, he scrutinized his own perceptual method by comparing it to others' methods. Second, even when double-loop inquiry fails to illuminate tragic flaws directly, it can offer us indirect evidence about where those flaws lie. Suppose we cannot determine what our methods are. As a matter of fact, sometimes we use untutored intuitions to answer questions; our prejudices shape our views; we are moved by groupthink. Tragic flaws can easily hide inside our unanalyzed and unknown methods and in the subterranean depths of our nature. If we give an answer and do not really know how we got there, we might think to ourselves: my tragic flaws could be taking me for a ride. Our inability to scrutinize our methods can lead us to doubt their reliability and, in turn, to doubt the beliefs we reached by using those methods.

As should be obvious, double-loop inquiry is not by any means a surefire solution to our problem. Our tragic flaws can influence second-order inquiry just as they influenced first-order inquiry. The flaws we do not recognize at the first level may remain unrecognized when we move to the second level. This would be unsurprising, given that our ignorance at the first level is implicated in our ignorance about our ignorance (Kruger and Dunning 1999; Dunning et al. 2003). The light just will not get in. But double-loop inquiry is one type of activity we can sometimes meaningfully engage in.

Realistically, if double-loop inquiry is going to help in important cases, it will probably be supported by groups and institutions. We cannot expect to do this alone. To be sure, as I noted, social realities might undermine double-loop inquiry. Upton Sinclair's remark suggests that if coming to understand our flaws means losing our salary, finding them out will be hard. Groups may inculcate and promote 'defensive routines' in all sorts of ways. But group dynamics may also encourage effective double-loop inquiry. Consider the practice of risk management in organizations.[8] Risk managers aim to understand risk and uncover potential limitations in the knowledge of risk (Taleb and Pilpel 2007). Managers take into account both the likelihood and the costs of risky events occurring. Correctly recognizing and sizing up the risks can be a life-or-death matter. Think about an automobile with a gas tank that explodes when it is rear-ended—the Ford Pinto.

---

[8] Thanks to Alex Arnold and Shane Wilkins for helpful conversation. Another example comes from the experience (however uncommon) of an intensely critical but constructive Q&A session at an academic philosophy conference. The speaker does not receive adoration and high-fives nor half-baked objections nor yawns and incredulous stares. Instead, the audience shares ideas the speaker had not earlier recognized and so may improve the speaker's thinking.

Or think about a space shuttle's frozen O-rings—*Challenger*. Or a bank that fraudulently creates accounts for customers and gets nailed with fines of almost $200 million—Wells Fargo. Organizations do not always correctly recognize or assess the risks they run, but good risk management averts failure occasionally and minimizes risk frequently.

Fire, freezing, and fraud are risks for some endeavors, but we ought not to forget one ubiquitous 'human' factor in risk management: tragic flaws. Failing to recognize our unreliable belief-forming dispositions can cost us and others dearly. There are interesting questions about the ways in which different organizations try to root out people's flaws. One type of strategy used in risk management is to share responsibility for double-loop inquiry across two people or distinct roles. Someone pursues the inquiry and represents it in a form that can be communicated while someone else checks for 'compliance'. We cannot always be expected to execute double-loop inquiry unless someone is watching over our shoulder. What actually works and what does not, and why, is worth finding out.

The questions here are part of what we could call 'corporate epistemology'—the branch of social epistemology that studies the dynamics of groups and organizations in creating both knowledge and ignorance. (This sort of investigation has also been called 'epistemic systems design'; see Goldman and Blanchard [2018: §5] for an introduction.) This is a traditional, albeit underexplored theme in the history of epistemology, going back at least to Francis Bacon's seventeenth-century work *New Atlantis*. Bacon imagined a scientific research institute, Salomon's House, in which division of labor and specialization would let researchers tackle investigations too demanding for any one person or unorganized group. To grapple with the idea that double-loop inquiry can be launched using social processes, we must attend carefully to the social world. Thinking about how to deal with our hidden flaws demands knowledge about the circumstances where our inquiry happens.

NATHAN BALLANTYNE
FORDHAM UNIVERSITY
n.ballantyne@gmail.com

References

Argyris, Chris. (1976) 'Single-Loop and Double-Loop Models in Research on Decision Making'. *Administrative Science Quarterly*, 21, 363–75.

Argyris, Chris. (1985) *Strategy, Change, and Defensive Routines*. Boston, MA: Pitman.

Argyris, Chris. (2002) 'Double-Loop Learning, Teaching, and Research'. *Academy of Management Learning and Education*, 1, 206–18.

Ashby, W. Ross. (1960) *Design for a Brain: The Origin of Adaptive Behaviour*. 2d ed. London: Chapman and Hall.

Ballantyne, Nathan. (2015) 'The Significance of Unpossessed Evidence'. *Philosophical Quarterly*, 65, 315–35.

Ballantyne, Nathan. (2019) *Knowing Our Limits*. New York: Oxford University Press.

Ballantyne, Nathan. (Forthcoming) 'A Review of Recent Work on Intellectual Humility: A Philosopher's Perspective'. *Journal of Positive Psychology*. DOI: 10.1080/17439760.2021.1940252

Blakemore, Colin, and Grahame F. Cooper. (1970) 'Development of the Brain depends on the Visual Environment'. *Nature,* 228, 477–78.

Cassam, Quassim. (2015) 'Stealthy Vices'. *Social Epistemology Review and Reply Collective,* 4, 19–25.

Cassam, Quassim. (2019) *Vices of the Mind: From the Intellectual to the Political.* Oxford: Oxford University Press.

Christensen, David. (2010) 'Higher-Order Evidence'. *Philosophy and Phenomenological Research,* 81, 185–215.

Church, Ian, and Justin Barrett. (2016) 'Intellectual Humility'. In E. L. Worthington, Jr., D. E. Davis, and J. N. Hook (eds.), *Handbook of Humility: Theory, Research, and Applications* (New York: Routledge), 62–75.

Dalton, John. (1798) 'Extraordinary Facts relating to the Vision of Colours: With Observations'. *Memoirs of the Literary and Philosophical Society of Manchester,* 5, 28–45.

Dunning, David, Kerri Johnson, Joyce Ehrlinger, and Justin Kruger. (2003) 'Why People Fail to Recognize Their Own Incompetence'. *Current Directions in Psychological Science,* 12, 83–87.

Dyer, Robert. (1965) '*Hamartia* in the *Poetics* and Aristotle's Model of Failure'. *Arion: A Journal of Humanities and the Classics,* 4, 658–64.

Elga, Adam. (n.d.) 'Lucky to be Rational'. Unpublished manuscript.

Feldman, Richard. (2005) 'Respecting the Evidence'. *Philosophical Perspectives,* 19, 95–119.

Fricker, Miranda. (2007) *Epistemic Injustice: Power and the Ethics of Knowing.* Oxford: Oxford University Press.

Gigerenzer, Gerd. (1991) 'How to make cognitive illusions disappear: Beyond "heuristics and biases"'. *European Review of Social Psychology,* 2, 83–115.

Goldman, Alvin, and Thomas Blanchard. (2018) 'Social Epistemology'. In Edward N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy* (summer 2018 edition). Accessed September 20, 2020. https://plato.stanford.edu/archives/sum2018/entries/epistemology-social/

Greco, John. (2012) 'A (Different) Virtue Epistemology'. *Philosophy and Phenomenological Research,* 85, 1–26.

Haggard, Megan, Wade C. Rowatt, Joseph C. Leman, Benjamin Meagher, Courtney Moore, Thomas Fergus, Dennis Whitcomb, Heather Battaly, Jason Baehr, and Dan Howard-Snyder. (2018) 'Finding Middle Ground between Intellectual Arrogance and Intellectual Servility: Development and Assessment of the Limitations-owning Intellectual Humility Scale'. *Personality and Individual Differences,* 124, 184–193.

Hazlett, Allan. (2012) 'Higher-order Epistemic Attitudes and Intellectual Humility'. *Episteme,* 9, 205–23.

Heath, Chip, Richard Larrick, and Joshua Klayman. (1998) 'Cognitive Repairs: How Organizational Practices can Compensate for Individual Shortcomings'. *Research in Organization Behavior,* 20, 1–37.

Hunt, David M., Kanwaljit S. Dulai, James K. Bowmaker, and John D. Mollon. (1995) 'The Chemistry of John Dalton's Color Blindness'. *Science,* 267, 984–88.

Jordan, Andrew, and Timothy O'Riordan. (1999) 'The Precautionary Principle in Contemporary Environmental Policy and Politics'. In Carolyn Raffensperger and Joel A. Tickner (eds.), *Protecting Public Health and the Environment: Implementing the Precautionary Principle* (Washington, DC: Island Press), 15–35.

Kelly, Thomas. (2008) 'Disagreement, Dogmatism, and Belief Polarization'. *Journal of Philosophy,* 105, 611–33.

Kelly, Thomas. (2010) 'Peer Disagreement and Higher-Order Evidence'. In Ricchard Feldman (ed.), *Disagreement* (New York: Oxford University Press), 111–74.

Kenyon, Tim. (2014) 'False Polarization: Debiasing as Applied Social Epistemology'. *Synthese,* 191, 2529–47.

King, Nathan. (2016) 'Religious Skepticism and Higher-Order Evidence'. In Jonathan Kvanvig (ed.) *Oxford Studies in Philosophy of Religion*, vol. 7 (Oxford: Oxford University Press), 126–56.

Kruger, Justin, and David Dunning. (1999) 'Unskilled and Unaware of It: How Difficulties in Recognizing One's Own Incompetence Lead to Inflated Self-assessments'. *Journal of Personality and Social Psychology,* 77, 1121–34.

Leary, Mark R., Kate J. Diebles, Erin K. Davisson, Katrina P. Jongman-Sereno, Jennifer C. Isherwood, Kaitlin T. Raimi, Samantha A. Deffler, and Rick H. Hoyle. (2017) 'Cognitive and Interpersonal Features of Intellectual Humility'. *Personality and Social Psychology Bulletin,* 43, 793–813.

Matthews, Gareth B. (2005) *Augustine.* Malden, MA: Blackwell.

Mollon, J. D., and L. R. Cavonius. (2012) 'The Lagerlunda Collision and the Introduction of Color Vision Testing'. *Survey of Ophthalmology,* 57, 178–94.

National Eye Institute. (2015) 'Facts About Color Blindness'. Available at: https://nei.nih.gov/health/color_blindness/facts_about.

National Transportation Safety Board. (2013) 'Head-On Collision of Two Union Pacific Railroad Freight Trains Near Goodwell, Oklahoma, June 24, 2012'. Railroad Accident Report NTSB/RAR-13/02, PB2013-107679. Washington, D.C.

Pronin, Emily. (2007) 'Perception and Misperception of Bias in Human Judgment'. *Trends in Cognitive Sciences,* 11, 37–43.

Ross, Lee, and Richard E. Nisbett. (1991) *The Person and the Situation: Perspectives of Social Psychology.* London: Pinter and Martin Publishers.

Schwitzgebel, Eric. (2020) 'The Jerks of Academe'. *Chronicle of Higher Education,* 31 January 2020. Available at: https://www.chronicle.com/interactives/20200131-the-jerks-of-academe.

Sinclair, Upton. ([1935] 1994) *I, Candidate for Governor: And How I Got Licked.* Edited by James Gregory. Berkeley: University of California Press.

Sliwa, Paulina, and Sophie Horowitz. (2015) 'Respecting All the Evidence'. *Philosophical Studies,* 172, 2835–58.

Stegner, Wallace. (1971) *Angle of Repose.* New York: Penguin.

Tagg, John. (2007) 'Double-Loop Learning in Higher Education'. *Change,* 39, 36–41.

Taleb, Nassim N., and Avital Pilpel. (2007) 'Epistemology and Risk Management'. *Risk and Regulation,* 13, 6–7.

Umpleby, Stuart A. (2009) 'Ross Ashby's General Theory of Adaptive Systems'. *International Journal of General Systems,* 38, 231–38.

Vetter, Barbara. (2013) '"Can" without Possible Worlds: Semantics for Anti-Humeans'. *Philosophers' Imprint,* 13, 1–27.

Voltaire. ([1764] 1901) 'A Philosophical Dictionary, Vol. V—Part II'. In John Morley and Tobias Smollett (eds.), William F. Fleming (trans.), *The Works of Voltaire: A Contemporary Version in 21 Vols.* (New York: E. R. DuMont).

Watson, J. D. (1993) 'Succeeding in Science: Some Rules of Thumb'. *Science,* 261, 1812–13.

Whitcomb, Dennis, Heather Battaly, Jason Baehr, and Daniel Howard-Snyder. (2017) 'Intellectual Humility: Owning our Limitations'. *Philosophy and Phenomenological Research,* 94, 509–39.

Wu, Kaidi, and David Dunning. (2017) 'Hypocognition: Making Sense of the Landscape Beyond One's Conceptual Reach'. *Review of General Psychology,* 22, 25–35.