'Ignorancia, Error y el Avance de la Comprensión', ***Error y Conocimiento: La Gestatin de la Ignorancia desde la Didáctia, la Ethica y la Filosofía*** ed. Victòria Campos, Anna Estany, Mercè Izquierdo. Grenada: Comares, forthcoming (Translated into Spanish by Remei Capdevila.)

Ignorance, Error, and the Advancement of Understanding

Catherine Z. Elgin

Introduction

Epistemologists standardly assume that ignorance is preferable to error. Ignorance as to whether *p* consists in neither believing that *p* nor believing that not-*p*. Error consists in believing, judging, or concluding that *p* when in fact not-*p*. Descartes (1641/1979) maintains that in cognitive matters we can and should always avoid error by refraining from judging unless our relevant ideas are clear and distinct. By suspending judgment, we remain ignorant, but we make no mistakes. Contemporary theorists (e.g., Rescher, 1973, Lehrer, 1974) contend that our overarching cognitive goal is to believe as many truths as possible and to disbelieve as many falsehoods as possible. This is not so extreme as Descartes' position. They sanction believing a few falsehoods so long as doing so leads to the highest overall balance of true over false beliefs. But suspending judgment about doubtful propositions plainly plays a major role in achieving this objective too. Even the Educational Testing Service agrees. On the SATs, GREs, MCATs, and an entire alphabet of other US university entrance exams, answering incorrectly is penalized more heavily than leaving a question unanswered. The issue I will address is whether this widely held preference is itself an error. I begin by considering three not implausible defenses of the preference – a practical defense, a

Socratic defense, and a strictly epistemological defense. I argue that none of them succeeds. I then offer reasons to think that error marks a cognitive advance over ignorance. Finally I argue that errors can contribute to the advancement of understanding in ways that ignorance cannot.

The Practical Defense

The practical argument rests on the fact that belief forms the basis for action. Acting on a false belief can be dangerous. A skater who believes that the ice is solid, blithely glides onto the pond. The ice breaks, and the skater drowns. The blame for her disaster rests squarely on her false belief. She skated on the pond because she erroneously believed that the ice would hold. This does not show that even in this case error is, from a practical standpoint, worse than ignorance. To establish that, we would have to show that a skater ignorant of the condition of the ice would not have gone onto the pond. This is not so obvious as it might first seem. To be sure, experienced, responsible skaters take pains to ascertain the condition of the ice before they skate. So it is plausible that a skater aware of her ignorance would not skate on thin ice. But children tend to be oblivious of their ignorance of such matters. They are apt to skate on a seemingly frozen pond without giving a moment's thought to the thickness of the ice. At best then, this case seems to indicate that knowing one is ignorant can be preferable to being unaware that one is in error.

In other cases, however, error seems to have the advantage. If the falsity of a belief is due to a currently irrelevant factor, acting on that belief serves an agent's purposes. Falsely believing that the turkey sandwich contains chicken, the hungry man

eats it. Since chicken is as nourishing as turkey, he achieves his goal of slaking his hunger. Assuming that he has no strong aversion to either, he does so at no significant cost to himself. His belief is erroneous in a way that does not matter. From a practical point of view, he is better off than his companion who refrains from eating because she is ignorant of what sort of meat the sandwich contains.

Nor do those who suspend judgment always have the option not to act. Ice-skating is a recreational activity, so not skating if conditions seem risky is feasible. Skipping a meal is an alternative to eating mystery meat. So, unless one is starving, refraining from eating the sandwich is a viable alternative as well. But life is full of forced choices. Driving from New Jersey to Manhattan, Jon comes to a fork in the road. If he goes right, he takes the George Washington Bridge; if he goes left, he takes the Lincoln Tunnel. Assuming he wants to get there as quickly as possible, he wants to take the route that currently has the least traffic. If he chooses wrongly, whether from ignorance or error, he ends up a hideous traffic jam. But the traffic is streaming along at 120 km per hour. Simply stopping because he does not know which route is better is not an option. One way or another he needs to decide. Should he decide badly, it apparently makes no difference whether his decision was due to ignorance or error. Whether he goes right because he mistakenly believed that today the bridge would be faster than the tunnel or because, out of ignorance, he arbitrarily chose to take the bridge, his situation is equally unfortunate.

Even so, it might seem that from the point of view of action, ignorance is on the whole preferable to error, because ignorance favors idling if one can whereas belief favors action. This assumes that it is worse to act on a false belief than, out of ignorance,

to refrain from acting. But even this is not always so. Promptly administering streptokinase to a heart attack victim limits the damage to her heart. So if a cardiologist knows that the patient is having a heart attack, he administers streptokinase. What if he does not know? If he falsely believes she is suffering from a heart attack and prescribes the drug, he unnecessarily risks side effects. But serious side effects occur in less than 0.5% of the cases. The probability of a significant negative consequence is thus negligible. Still, there is a risk. Suppose he withholds the drug because he suspends judgment about whether she is having a heart attack. Then if she is not having a heart attack, she is no worse off than if he had prescribed the drug, and if she is having a heart attack, his inaction will cause her to suffer significantly greater heart damage. In this case, refraining from acting is worse than acting.

The verdict is mixed. Although there are some cases where, for practical reasons, ignorance is preferable to error, there are others in which error is preferable to ignorance, and yet others in which they are on a par. Practical arguments do not in general decisively favor ignorance over error.

The Socratic Defense

Socrates (Plato, 399 BCE/1956) maintained that knowing that one does not know is preferable to not knowing that one does not know. The rationale is this: if someone knows that she does not know whether *p*, she can attempt to find out. But if she believes that she knows that *p*, she has no incentive to investigate the matter further. So falsely believing that one knows is undesirable. For the mistaker does not know, but still has no reason to investigate further. This suggests that an epistemically responsible agent is

more likely to alleviate her ignorance than her error.

Let us concede Socrates' point about the cognitive value of knowing that one does not know. Why does this favor ignorance over error? The reason is this: A fully rational agent cannot believe that, for example, her belief that Andorra is in Spain is mistaken and still believe that Andorra is in Spain. It is Moore's paradoxical to say, or think to oneself, 'I mistakenly believe that Andorra is in Spain.' (Moore, 1942). For if I believe that Andorra is in Spain, I think that the proposition 'Andorra is in Spain' is true;[1] if I consider the belief that Andorra is in Spain to be mistaken, I think that the proposition 'Andorra is in Spain' is false. A rational agent cannot simultaneously think that 'Andorra is in Spain' is true and that 'Andorra is in Spain' is false. Thus insofar as I am a rational agent, as soon as I discover that something I believe is false, I cease to believe it. That being so, it follows from the Socratic argument that someone who harbors a mistaken belief must be unaware of the inadequacy of her cognitive condition, whereas someone who is ignorant of a particular matter need not be.

Again, though, this is not a contrast between ignorance and error, but between knowledge of one's ignorance and the inevitable lack of knowledge that one harbors a specific erroneous belief. But one can be ignorant of one's ignorance as well as of one's error. A person who has never heard of Andorra is not only ignorant of its location, she is also ignorant of her ignorance of its location. She has no incentive to find out. So such ignorance is no more likely than an error to be uncovered and alleviated.

Moreover, not all cognitive errors are false beliefs. Suppose Ben adds a long column of numbers. As soon as he looks at the result, he knows it is wrong. (The result he arrived at is far too great or far too small). Realizing that his result is in error, he

knows that he does not know the sum. The Socratic argument does not clearly favor ignorance over this sort of error, since there is nothing Moore's paradoxical about a situation of this kind. Evidently, knowing that one does not know is possible in a state of error as well as in a state of ignorance

Like the practical argument, the Socratic argument is inconclusive on the issue that concerns us. It affords no clear reason to prefer ignorance to error, or error to ignorance.

The Epistemological Standoff

Each separate bit of knowledge does not stand alone. If I know that iron is magnetic, I can build on that knowledge. I can safely and responsibly draw inferences from the proposition 'Iron is magnetic', and use 'Iron is magnetic' as a reason to support other contentions. When I do these things, 'Iron is magnetic' becomes interwoven into my fabric of cognitive commitments. Other beliefs depend for their cognitive standing on the support that 'Iron is magnetic' supplies. If I mistakenly believe that copper is magnetic, I integrate 'Copper is magnetic' into my cognitive system. I draw inferences from 'Copper is magnetic', and adduce 'Copper is magnetic' as a reason. Taking my inferences to be sound, and the other contentions to be supported, I consider them knowledge as well. Errors thus ramify; they spread throughout a cognitive system, weakening or depriving of support a host of directly and indirectly related commitments.

Ignorance seems safer. If I am ignorant as to whether copper is magnetic, I refrain from incorporating 'copper is magnetic' into my cognitive system. I do not adduce it in support of other contentions. And if I use it in inferences at all, I do so only

in inferences I know to be hypothetical. I do not draw conclusions based on 'Copper is magnetic'.

It is not unreasonable to want to prevent errors from entering one's cognitive system. But how far we should go to avoid incorporating errors is less clear. Traditional foundationalist epistemology is highly risk averse (Chisholm, 1982). Because it considers accepting a proposition – that is, admitting it into one's cognitive system – virtually irrevocable, it strongly prefers ignorance to error. If excising a false proposition from a belief system is exceedingly difficult, we should avoid introducing one. To be sure, by setting the standards for acceptance high, we consign much to the realm of ignorance. But what gains admittance into a cognitive system is epistemically secure.

Such a policy seems attractive. If we adopt it, not only can we be confident about the propositions we have accepted, we do not need to constantly review our epistemic situation to see whether previously accepted propositions merit continued acceptance or need to be rescinded. By making the standards of acceptability sufficiently high, we put ourselves in a position to treat acceptance as virtually irrevocable.

There are at least two worries with such a policy. First, if this is our policy, then the system of accepted propositions has enormous cognitive inertia. Since the standards are so high, there is a very strong presumption that whatever has been accepted is true. This means that a new result is overwhelmingly likely to be rejected if it fails to accord with what has already been accepted. In case of a mismatch, the novel consideration is deemed unacceptable.

It is reasonable to hold that there is some presumption in favor of what has already been accepted. That body of beliefs has stood the test of time, and has served us,

we suppose, fairly well. The worry is that if the presumption in favor of what has already been accepted is too strong, revolutionary insights will be incapable of discrediting received views. Einstein could then never displace Newton; Darwin could never unseat Aristotle; democracy could never discredit despotism. But no matter how high our standards, we cannot guarantee that error will *never* pass for knowledge. So if we make discrediting received views too difficult, we may make errors effectively irrevocable.

Second, although consigning suspect considerations to the realm of ignorance may assuage our epistemological consciences, it cannot always justify excluding them from our deliberations. If, for example, a physician has to decide between two courses of therapy, or a driver has to decide which fork in the road to take, she cannot simply say, 'Well, I don't have enough information. I am ignorant of the matter.' Her claim may be true, but she has to decide anyway. Admitting, 'I just don't know' does not solve her problem. Nor does it give her any way to proceed.

Holistic theories like mine (Elgin, 1996) are less risk averse. Constraints on acceptance are weaker than the constraints that epistemology traditionally imposes, because acceptance is always provisional and revocable. Thus according to my account, it is sometimes reasonable to take a cognitive chance and accept a doubtful proposition. Because there are resources for identifying and rejecting previously accepted contentions, mistakes can be corrected. On my theory, we are always attempting to arrive at the constellation of beliefs that is best on balance. Because the constellation available at any time incorporates many of our cognitive commitments and satisfies our standards of acceptance, there is some presumption in its favor. If a new finding clashes with the accepted system, it is apt to be rejected. But some new findings force revisions in the

previously accepted fabric of commitments. If a better overall account would emerge if we incorporated the new finding and revised or repudiated some of what we had previously believed, that is what we should do. Thus, despite the weight of evidence in favor of Newtonian physics, the deflection of starlight observed during the 1919 eclipse constituted a good reason to accept the theory of relativity and reject the previously accepted theory.

Because the theory is holistic, the ramifications of error may be less serious than on a traditional account. If a contention is secured by multiple, relatively independent chains of reasoning, the discovery that any one of the chains is defective can leave the contention sufficiently well supported. This is not to deny that the loss of support would weaken the fabric of commitments; it is only to say that such a loss of support need not be fatal. If we discover that one of the five witnesses to the crime is colorblind, we may still have plenty of reason to believe that the bank robber was wearing a red shirt. The testimony of the other four witnesses suffices. An epistemological foundationalist would advocate distrusting all five witnesses unless their capacity to discriminate colors has been firmly established. On my view, given that we have grounds for believing that most people are not color blind, it may be reasonable to take the witnesses' sincere claims at face value, and only attempt to discredit them if a problem emerges. Since acceptance is revocable, incorporating a contention that might turn out to be false is a viable option.

Ignorance has costs as well. If we fail to incorporate a well supported, but not maximally well supported, contention on the grounds that it might be false, we weaken our fabric of cognitive commitments. Suppose, for example, we refuse to believe a seemingly credible eyewitness's claim because we are in no position to establish that he

is not colorblind. (Maybe he is a historical figure, hence unavailable to be tested.) While we do not incorporate a falsehood into our understanding of the robbery, we deprive ourselves of the information that if, like most people, he was not colorblind, he could supply.

We seem to have arrived at an impasse. There are epistemological considerations that favor ignorance over error and epistemological considerations that favor error over ignorance. Apparently the jury is still out.

Error as a sign of success

Although none of the foregoing arguments decisively favors ignorance over error or error over ignorance, I suggest that there are reasons to think that error can be cognitively more valuable. One reason is Davidsonian. According to Donald Davidson (2001), the identity of a belief derives from its place in a rich, textured array of relevant, true beliefs. A person who lacks the requisite background cannot have a particular belief. So, for example, unless Paul knows a good deal about physics, he cannot rightly or wrongly believe that electrons lack mass. He would not know what an electron is or why the question of its mass even arises. Unless Pat knows a good deal about Catholicism, she cannot rightly or wrongly believe that Mother Teresa was canonized before she died. She would not know what it is to be canonized, what steps are required, or whether it must be done posthumously.

Few, if any, specific beliefs are required to back a given contention. One can approach a question from different directions. So apart from very general beliefs, such as ‘electrons are subatomic particles’ or ‘canonization is a Roman Catholic honor’, no

specific body of beliefs is required to equip the agent to form beliefs about a topic. Moreover, there are bound to be undecidable cases. It is not clear how sparse a complement of relevant beliefs a person can have and still, rightly or wrongly, believe that *p*. Indeed, the answer might depend on what sort of belief is in question. If little is known or reasonably believed about the subject, a relatively sparse constellation of relevant beliefs might suffice. If plenty is known, then the agent's constellation of beliefs might have to be considerably denser. Nevertheless, a fairly wide and dense cluster of relevant beliefs is, Davidson maintains, required to anchor each particular belief.

Davidson insists that most of the beliefs that constitute the background must be true. I am not convinced. Rather, I believe, sufficiently many of the agent's relevant beliefs be roughly true and reasonably believed. If we insist that for Fred to have beliefs about quarks, most of his beliefs about quantum mechanics must be true, we ask too much. Perhaps most of what is currently accepted in quantum mechanics is roughly true. But it would be unsurprising if many of the contentions of current theory turn out to be not exactly true. So I modify Davidson's claim and say that to be in a position to harbor (true or false) beliefs about a topic, an agent must have a constellation of beliefs, most of which are at least roughly true and reasonably believed.[2]

Whether or not one accepts the entire Davidsonian position, this point seems right. If Paul knows no physics, he has no idea what an electron is or what it takes for a particle to have or lack mass. If Pat knows nothing of Catholicism, she has no idea what is required for canonization, hence no basis for telling whether the Church's activities vis à vis Mother Teresa amount to canonizing her. Too great an ignorance of a topic leaves a person out of touch with its subject matter. So he can have no views about it. If Paul

cannot tell an electron from a quark, then nothing in his cognitive system equips him to have views about the one but not the other.

This means, however, that to be in a position to make a mistake marks a significant cognitive accomplishment. Only someone who knows a good deal about a topic has the resources to have mistaken beliefs about it. Only because Fred knows a good deal about quantum mechanics, can he harbor the erroneous belief that quarks and antiquarks exchange charms. (They exchange muons.) So being in a position to have erroneous beliefs about a topic is itself a cognitive accomplishment. It reveals a significant measure of understanding of that topic.

Ignorance, on the other hand, can be blind. Paul may have no opinions whatsoever about quarks, and indeed be incapable of forming any opinions about them, since to do so would require an understanding of quantum mechanics which he does not begin to have. He seems considerably further from an understanding of the subject than Fred does. He has a long way to go before he can be in a position to make Fred's mistake.

Error is often a way station on the road to truth. An eight-year-old child who believes that humans descended from apes is in error. Humans and the other great apes descended from a common hominid ancestor who was not strictly an ape. But the child displays some understanding of evolution. She is clearly better off cognitively than her playmate who is wholly ignorant of evolution. Unlike her playmate, she is poised to learn the truth about human evolution, since her opinion, although false, is not all that far from the truth. Because the error is nested in a network of true beliefs, resources for correction are available. The child who is ignorant of evolution is considerably further

from understanding our relation to the other great apes.

The pattern exhibited here is endemic to scientific education. We typically begin with rough characterizations that properly orient us toward the phenomena, then refine the characterizations as our understanding of the science advances. We follow a natural trajectory from folk biology (with fixed species) to evolutionary biology, and from folk physics (which is roughly Aristotelian), through Newtonian mechanics, to relativity and quantum mechanics.

One might argue that such false beliefs are, at most, useful heuristics. Possibly the best way to arrive at the truth is to follow a well-trodden path of errors. Still, it might seem, the way stations on the path are *merely* heuristic. There is nothing cognitively valuable about them, aside from the curious fact that they lead fairly reliably to the acquisition of true beliefs. Although they are pedagogically better than ignorance, then, they are not epistemologically better.

I disagree. First, the eight-year-old has already made a significant cognitive advance over her playmate who is ignorant of evolution. She clearly understands something that her playmate does not about the relations among the species and the grounding of those relations in the history of life on this planet. That level of understanding, although it has not yet eventuated in a true belief about our relationship to the other great apes, is itself an epistemologically estimable cognitive achievement. Second, and perhaps more importantly, the pattern that students display in learning science is the same pattern as science displays in the sequence of theories it develops.

A central tenet of Copernicus's theory is the contention that the Earth travels around the sun in a circular orbit. Kepler improved on Copernicus by maintaining that

the Earth's orbit is not circular, but elliptical. Later astronomers reasoned that the orbit is not precisely elliptical, for the gravitational pull of the other planets on the Earth perturbs the ellipse. Currently, having abandoned the commitment to absolute space, astronomers can no longer say that the Earth travels around the sun, simpliciter; they must speak of how the Earth and the sun move relative to each other. Nevertheless, despite the fact that Copernicus's central claim was strictly false, it and the theory that embeds it constitute a major advance in astronomical understanding over the Ptolemaic theory it supplanted. Kepler's theory is a further advance. Later theories are further advances. This seems obvious. The advances are clearly cognitive advances. With each step in the sequence, we understand the motion of the planets better than we did before. But no one claims that contemporary astronomy has arrived at the exact truth about the motion of the planets. Like Copernicus and Kepler and the eight-year-old child, we are at a way station on the path to truth. Each mistaken theory paves the way for the next step. Only because Kepler looked at the heavens in a Copernican frame of mind was he able to modify and improve on Copernicus's theory.

Some errors, we have seen, are false in ways that do not matter. The man whose hunger was satisfied by what he falsely took to be a chicken sandwich made an error of this kind. In other cases, however, an error may be felicitous. Not only does it do no harm, it is an asset to understanding. The child who thinks that humans descended from apes embeds that contention in a general account that reflects both a commitment to evolution and an idea that humans and other apes are closely related. So although there is a falsehood involved, it is one that enables her to connect, synthesize, and grasp a body of information that is grounded in the biological facts and is supported by the evidence

available to her. Evidently, her false belief fosters her understanding and puts her in a position to learn more. Still, it seems possible to insist that her error is only of heuristic or pedagogical value.

But it is implausible to say the same about Kepler. The Earth's orbit is not elliptical. Believing it to be elliptical is an error. Nevertheless, the Earth can be accurately represented as traveling around the Sun in an orbit that is not all that far from an ellipse. So his false belief figured in, and enabled Kepler to unify, a body of information that answered to the evidence better than his predecessors could. Moreover, at a certain level of precision, the difference between the shape of the Earth's orbit and an ellipse makes no difference. At this level, Kepler's laws embody a sufficient understanding of the motion of the planets that they foster further understanding. They give astronomers something to build on. For certain purposes, then, the contention that the Earth's orbit is elliptical may be as good as the truth. Indeed, it may be better than the truth. Accommodating the perturbations in orbit that it overlooks might lead to intractable cognitive or computational complications, masking the regularities that scientists seek.

To be sure, both Kepler and the eight-year-old would be even better off if they knew that their beliefs were only approximately correct. I am not here arguing that error is just as good as knowledge. But science abounds with errors of this kind, and they seem preferable to ignorance. For they are manifestations of a significant level of understanding of the phenomena they concern, and they put those who harbor them in a good position to discover their mistakes and correct them.

Learning from our Mistakes

We are regularly reminded that we should learn from our mistakes. Ordinarily, perhaps, we think that this means that we should learn not to make a particular mistake again. Karl Popper (2002) considers this the key to science. Science, he urges, can never prove anything. It can only disprove. Thus, he concludes, scientific progress consists in formulating hypotheses and attempting to disprove them. Each time we generate a disproof, we make progress. We know that that particular hypothesis is false. If we do not make that mistake again, we get closer to the truth. What we learn from our mistake then is that things are not the way the hypotheses says they are.

Where the mistake is a matter of isolated fact, such as 'King Philip I of Spain was King Philip II of Portugal', that the proposition is false may be all that is available to learn. Similarly, when mistakes are random guesses. When a student discovers that he is mistaken in thinking that the value of Avogadro's number is 14, he is not much closer to knowing what the actual value of Avogadro's number is. Learning from mistakes through a process of elimination is extremely labor-intensive. There are indefinitely many ways the world might be. So the idea that we will get at the truth by sequentially eliminating hypotheses that have shown themselves to be mistaken is not promising. Neither individually nor as a species will we live long enough to eliminate all the false hypotheses.

A Popperian characterization of the issue treats every error as a stab in the dark. When we do not know much about a situation, we are not in a position to learn much from a mistake. But our situation is rarely so bleak as this suggests. Usually, we know enough that the range of plausible alternatives is restricted. If Maria know that there are

only three configurations in which a particular protein might fold, then eliminating one of them yields considerable information about the protein – it folds in one of the remaining two ways. If Bill is aware that only three candidates are running for office, the news that one has been eliminated again yields considerably more information than if he thought (as Popper sometimes seems to recommend) that any one of the nearly 7 billion people on earth might be elected.

But even if we grant that our cognitive situation is frequently better than that of the student who is wildly guessing values for Avogadro's number, it might seem that our prospects of learning much from a mistake are bleak. The problem is a consequence of holism. Quine (1961) argues, correctly I believe, that 'statements about the external world face the tribunal of experience not individually, but only as a corporate body'(p. 41). This means, as Walden (2007) explains, that we never get evidence for an isolated proposition; we get evidence for a proposition relative to a cluster of background assumptions and methodological prescriptions about what counts as evidence for what. The entire collection is confirmed or disconfirmed together. It follows that what we learn when we discover that we have made a mistake is that something in a wide constellation of beliefs is false. But the error alone does not target any particular element of the constellation.

What this pessimistic assessment overlooks is that typically not all of the elements of the constellation are equally vulnerable. Since constellations of beliefs overlap, some elements of the constellation that contains a mistake may be independently supported by their roles in other constellations that have been confirmed. Then we have prima facie reason to believe that they are not the locus of error. The more solidly grounded our

background assumptions, the more precise the focus a mistake can provide.

At the opposite extreme from random guesses are cases where we are quite convinced that we know what will happen, but turn out to be wrong. In such cases, our mistaken conviction is a telling error. An experiment that to our surprise fails to confirm a seemingly well established theory is such a case. Let me briefly describe such an experiment (Suzuki et al, 2007). A plasmid is a circular DNA molecule found in bacteria. Inserting mutations into plasmids is a common technique in genetic engineering. The bacteria containing those plasmids then transfer the mutation to other bacteria through conjugation, a process by which bacteria exchange nuclear material. That is the established, well confirmed background against which the following experiment took place. A bacteriologist in Marcin Filutowicz's laboratory attempted to insert two different mutations into a bacterial plasmid. Given the previous successes and the understanding of bacteriology, he had good reason to believe he would succeed. Introducing mutations into plasmids is a boringly routine thing for bacteriologists to do. But, his belief was mistaken. No matter how often he tried, and no matter how carefully he proceeded, he could not produce a live bacterium containing the two mutations. This constituted a surprising experimental failure. In light of the then current understanding of bacteriology, the procedure ought to have worked.

A Popperian might conclude that what we learn from this mistake is that one cannot introduce the two mutations into the same plasmid and obtain a live bacterium. A Quinean might conclude that there is a mistake somewhere in the cluster of assumptions that led bacteriologists to believe that the introduction of the two mutations was possible. Both would be right. We do learn these things. But, I suggest, we learn something more.

Introducing either of the mutations alone did not produce the unwanted result. The background assumptions and methodology were the same whether one or two mutations were introduced. So rather than simply concluding that something must be wrong somewhere in the vast constellation of substantive and methodological assumptions, Filutowicz was in a position to zero in on the mistake -- to recognize that the mistake arose from thinking that if you can successfully introduce each of the mutations, you can successfully introduce both. That is, given the depth and breadth of the background knowledge, he was in a position to use the mistake to probe current understanding of plasmids.

The mistake put him in a position to look more carefully at what occurs within a plasmid when mutations are introduced. Having evidence that the two mutations together did something that either alone did not do, he hypothesized that when the two mutations were introduced together, they caused the plasmid to overreplicate and destroy the bacteria. He confirmed the hypothesis and went on to devise a bacterium that suppressed the overreplication, thereby creating a 'Trojan Horse'. The new bacterium contains the normally overreplicating plasmid. In conjugation, it passes on the propensity to overreplicate to bacteria that cannot survive the process. This yields an antibacterial agent that evidently does not trigger antibiotic resistance.

For our purposes, the important point is that because so much was understood about bacteriology and about the process of introducing mutations into plasmids, once the error was discovered, the mistaken assumption was extremely informative. Filutowicz and his associates did not just learn that you cannot introduce the two mutations at once; or that there is something wrong somewhere in the cluster of assumptions that led them to

think that they could do so. Because the mistake focused attention on the replicating behavior of plasmids, it opened the way to new and fruitful insights. If it turns out, as it now looks as though it might, that pathogenic bacteria are incapable of developing a resistance to the 'Trojan Horse', then the fruits of the scientist's mistaken belief will be far more valuable than the insights that would have been gained had his expectation about the experimental outcome been true.

One other case worth mentioning is the Michelson-Morley experiment. At the end of the 19th century, physicists recognized that light consists of waves. They assumed that waves require a medium of propagation. So since light waves travel through space, they posited that space was filled with a medium called 'the luminiferous ether'. The objective of the experiment was to measure the flow of ether across the Earth as the Earth travels around the sun. Although the experiment was designed and conducted with exquisite care, the result was null. No ether drift was detected. Over a period of years, the experiment was redesigned and ever more sensitive measuring devices were used. But to no avail. The belief that light consists of waves in an ethereal medium is mistaken.

This familiar story may seem to make the same point as the previous one. But it is important to notice that the payoff is different. The cognitive consequences of the mistaken belief that Filutowicz and his colleagues shared were local and limited. Because the tightly woven web of commitments they relied on consisted in large part of independently confirmed strands, they could reasonably quickly identify the erroneous belief and see what routes of inquiry it opened up. The rest of the commitments largely held firm when the mistake was corrected. The Michelson-Morley experiment was

different. There too, scientists had a tightly woven fabric of established cognitive commitments, which led them to believe that the experiment would succeed. Their failure eventually convinced them that luminiferous ether does not exist. But this did not, and could not, lead to a local and limited revision of their understanding. It tore the fabric of cognitive commitments apart. For if light waves do not require a medium of propagation, light is radically different from what science supposed. And if that is so, then many of their other assumptions about matter and energy had to be false as well. I do not claim that the Michelson-Morley experiment was a 'crucial experiment' that by itself falsified Newtonian physics. Rather, I think, its effect was Socratic. It made manifest to the scientific community the extent to which they did not know what they thought they knew.

If a belief is supported by a tightly woven tapestry of reasons, and the belief turns out to be erroneous, more than that particular belief must be revised or rejected. The question arises: How could it be wrong? What have we been missing, or overlooking, or underestimating, or misconstruing? The realization that this is *not* the way things are in a particular area can often afford avenues of insight into the way things are. At the very least, it enables us to focus attention on particular aspects of our system of beliefs. It not only motivates us to ask questions we would not otherwise ask, it often also provides the resources for answering them.

The human propensity for error is typically regarded as a regrettable weakness. Certainly the propensity to make careless mistakes is a weakness. So, presumably, is the propensity to jump rashly to erroneous conclusions. But, I have suggested, the propensity to make thoughtful, educated mistakes may be a strength. This is not to say that every

error is felicitous. Some, like random guesses, provide virtually no information beyond the fact that things are not the way we supposed. Some are cognitively (and often morally) culpable. They fail to take account of available information and therefore lead us in the wrong direction. South African President Thabo Mbeki's denial that HIV causes AIDS is such a culpable error.[3] But sometimes, despite our best efforts, we get things wrong in fruitful ways. Once discovered, such errors provide both incentives and resources for serious, focused inquiry. By revealing not only that, but also where we have got things wrong, they point us in the direction of advancing our understanding. We are lucky we are disposed to make such mistakes.

Catherine Z. Elgin

Harvard University

References

Chisholm, Roderick. (1982). *The Foundations of Knowing*. Minneapolis, Minnesota: University of Minnesota Press.

Davidson, Donald. (2001). 'Rational Animals,' in *Subjective, Intersubjective, Objective.* Oxford: Oxford University Press, 95-106.

Descartes, Rene. (1641/1979). *Meditations on First Philosophy*. Indianapolis, Indiana: Hackett.

Elgin, Catherine. (1996). *Considered Judgment.* Princeton, N. J.: Princeton University Press.

Lehrer, Keith. (1974). *Knowledge*. Oxford: Clarendon.

Moore, G. E. (1942). 'A Reply to My Critics,' in P. Schillp (ed.) *The Philosophy of G.E.*

Moore, New York, N.Y: Tudor.

Plato. (399 BCE/1956). 'Apology', *Euthyphro, Apology, Crito.* Indianapolis, Indiana: Bobbs-Merrill.

Popper, Karl. (2002). *The Logic of Scientific Discovery.* London: Routledge.

Quine, W. V. (1961). 'Two Dogmas of Empiricism.' *From a Logical Point of View*. New York: Harper & Row.

Rescher, Nicholas. (1973). *The Coherence Theory of Truth.* Oxford: Clarendon.

Suzuki, H. et. al. (2007). 'Novel Strategy to Treat Multi-drug-resistant Gram-negative Pathogens Using Bacterial Conjugation,' *Plasmid* 57, 237-38.

Walden, Kenneth. (2007). 'A Practico Priori'. Unpublished paper. MIT.

[1] A proposition, on my view, is an equivalence class of intertranslatable sentences.
[2] I also disagree with Davidson's contention that only language users are capable of having beliefs. But that issue does not arise in this paper so I shall not discuss the matter here.
[3] I am grateful to Stéphane Leyens for this example.