General Identification of the Argument in "On Methods of Inference"

Post by “Elayne” of October 29, 2020 at 8:58 AM

Is everyone clear on the difference between analogies, making observations before conclusions, and the scientific method, which is a more accurate way of testing observations? It's hard for me to tell from the conversation so far.

Observations would be:

I've observed that on the visible level, everything I see is made of parts, and those particles have various features like being smooth or rough, heavy or light, etc. I haven't seen any exceptions.

Analogy based on observations would be: I've also seen phenomena that could only be explained if these visible particles were composed of particles too small to see. I am concluding that the smallest particles also have features like smooth, rough, different shapes, etc, like the visible ones, and many visible behaviors of matter would make sense to me if that is correct. This is the kind of process we see outlined in Lucretius.

A modern analogy would be: I have tested this drug in mice, which are mammals, and it works, so it will work in humans. Or, I've tested this drug in a bunch of humans and it worked for them so it _will_ work for you. Which many people, including some doctors, think is what we do, but it's really not, lol.

The modern process that I have not seen in Epicurus is: I've made a lot of observations about matter. I have some ideas about how that might be happening-- about the mechanisms and the composition of matter too small to see. I am going to form a specific, falsifiable hypothesis which makes a prediction about how matter will behave under certain circumstances-- how it would behave if my hypothesis is right-- and I'm going to run experiments. I'm going to repeat those experiments many times and see if other scientists can repeat them and get the same results. I'm going to get direct sensory observations by building instruments sensitive enough to register what my eyes can't, as an extension of my senses. I will be using my senses to make direct observations under strictly controlled conditions.

I'll remain aware of possible confounding factors I haven't controlled my experiments for, and I will also do other experiments to test my hypothesis. After multiple different types of evidence have been obtained by multiple people, I will accept the conclusions as factual enough to use. The science definition of fact is pragmatic, neither skeptical nor dogmatic. But once a fact has

been repeatedly observed as reliable, very strong evidence would be required to falsify it. For instance, now that we have observed electrical activity in the brain to cause seizures, and even in some case particular genetic mutations resulting in ion channelopathies and seizures, if someone wants to say naw, it's demon possession, they need to produce a demon.

This difference between the simpler observation to conclusion process vs observation to hypothesis to testing process is in the reliability of the conclusions, something which has also held up (that's a sort of meta experiment-- what kind of data collection turns out to be reliable over time).

This is so well established that doing "post hoc" analysis of data collected to test another hypothesis is called derisively the "spaghetti method." This is why it is standard in meta-analysis to exclude post-hoc papers. In the publish or perish world, researchers will mine their data for other patterns. Intuitively it seems that should be fine, but it is notoriously unreliable. Instead, the reliable approach is to take any post hoc observations as a new hypothesis and design tests for it.

Cassius, what you will immediately notice is that there is the issue of exactly how many times and in how many ways results need to be replicated before we are going to accept them as reliable. That is an important issue, and it's where statistics come in. We use things like p values and control groups to tell us how likely it is that our results are to be different from chance. We can choose how certain we want to be about a particular conclusion.

That issue is present always, but the point is that we can compare reliability, not that we can make anything 100% reliable.

For some hypotheses only one counter example is required. For instance, the plane crash. Those are the always/never type hypotheses. But usually in biology, it's more about "will this drug work for more people than not using it, or for more people than a current treatment?"

As far as levels of evidence-- because we are not typically using always/never hypotheses, in medicine we have levels which are not arbitrary but based on how reliable conclusions are from each type of evidence. Some people include expert consensus but I think that's silly-- I would only include the level of evidence they used.

The lowest level of evidence in medicine is the kind that is historically least likely to be correct. For example, if one person given orange juice recovers from flu. We don't have enough data to decide if that is different from chance. But a single case is sometimes enough to be attention grabbing. What if we give someone OJ for a disease no one is known to have survived, and they survive? It's a hypothesis worth testing, but since OJ itself isn't known to be risky unless you are

allergic, there would be a low threshold for using it before any controlled study-- and depending on how that went, you might decide it is unethical to design a higher level of evidence study.

Case series are the next level, more reliable than one case. And so on up to double blinded, randomized, controlled studies with replication _and_ a documented mechanism of action.

So when we are looking at the conclusions, it is not consensus whether a level of evidence is more reliable than another. That's something we have directly observed by comparing methods. That's the case even though sometimes a single case will turn out to be reliable but a large study turn out to have unexpected confounding variables.

This is what lets us say with confidence things like "there is a plausible mechanism for seizures that does not involve demons, so we are going to disregard your demon idea unless you produce evidence" or "there are plausible mechanisms for the repeated observations we've made about the neuroscience of dreams which do not require a missing element of some particle or energy transmission from outer space penetrating the skull, and furthermore, we have evidence that people often erroneously perceive agency where none exists upon closer examination, so we are going to disregard that idea unless you can produce evidence stronger (more replicable) than what we have so far."