Correlation, causation, and ethics

Ben Klemens

11 October 2009

We humans have a strong drive toward explaining events via causal stories. A list of facts is just not satisfying; we want it organized into causal stories where one fact leads to another. That drive to turn plain facts into causal chains has a lot of implications; here, I'll discuss how that drive determines our ethics.

As you well know, statistics has no concept of causality. This is one of those philosophy of science things that one could expound on forever, though I won't go into it too deeply here. But the concept of causation happens only inside the human brain. It's not something we can measure, perhaps with a causality ruler (or a more portable causality tape), and then write down that A causes B with 3.2 causal units, but C causes B with 8.714 causal units. There are intuitive ways to measure a causal claim, like saying that if A always comes before B, then A causes B; in direct correspondence, there are easy ways to break such a simple measure, like how Christmas card sales cause Christmas.

You could take the basic intuition about how causality works and build machinery to draw causal flowcharts, which give a wealth of means to reject the flowchart; look up structural equation modeling or read Perl [2000]. But apply the principle that statistics can never prove a model of the world, but only reject a false model: statistics can never prove a causal model of the world. In fact, this case is only worse because we're not entirely certain about how to measure or even identify causality. As with any model, stats can bolster or cut down our confidence in the causal claim, but that's where it ends.

But people like stories. As kids, we're taught how the world works via causal stories, that were not just a list of incidents but were a chain of events. Because granny was ill, Ms Hood took her basket of food and went walking over the river and through the woods; because the wolf was hungry and evil, he conspired to eat Ms Hood; because Ms Hood was virtuous, she was saved. A story where a bunch of unconnected, seemingly random things happen is just not satisfying, and correlation without causation is dissatisfying in exactly the same way.

Of course, people fake it all the time. You will rarely if ever find a newspaper article declaring a correlation without strongly implying (if not directly stating) that the statistical model showed a causal link. Get your favorite researcher drunk and he or she will stop talking about correlations and start talking about causation, even though everybody in the room knows that it's just a mathematical mirage. The drive for turning random facts into causal stories is just too strong.

Application to ethics   There's a platitude that it is ethics that distinguishes humans from the rest of the natural world. Above, I said that humans are distinguished by their ability and tendency to perceive causal relationships. These two statements are closely related: without causality, there can be no ethics.

Some causal chains are obvious, even to young children: if I drop a plate, then it breaks. If I kick the dog, the dog will bite me. For those that are not so obvious, you can help your child by laying it all out line by line. Here is Joe. Joe committed a misdeed. As a result, Joe's misdeed came back to him and he suffered. Here is Jane. She committed a virtuous act, and as a result, she was rewarded for it. The end.[1]

Person does good, is rewarded; person does bad, is punished may sound simplistic, but it is the canonical format used by most of the stories we hear or see or read. The modern version of Little Red Riding Hood above, all of Aesop's Fables, the one about Snow White and the vainglorious queen, any romantic comedy, they all tie reward to the virtuous and punishment to the misbehaving. We'll get to the stories that don't riff on that theme below.

These stories help us to move up the ladder of causal subtlety from mechanical misdeeds like kicking the dog to societal issues like littering. Thus, causal stories of the form virtue $\Rightarrow$ reward and ill behavior $\Rightarrow$ punishment are really central to building a society.

[It so happens that religious stories directly fit into the same structure: the omnipotent overseer makes certain that good $\Rightarrow$ reward and bad $\Rightarrow$ punishment. That is, where no simple causal mechanism exists, the omnipotent overseer defines one.]

The lit   I think it's so completely obvious that morality is taught through causal chains that I don't feel much compulsion to provide a host of references, but let me give you one or two so you know I'm not entirely making this up.

First, we can point to Jean Piaget, an oft-cited pioneer in the academic study of child development. Among others, he wrote many books on how children develop cause-and-effect relationships, and one entitled *The Moral Development of the Child* (that has almost no discussion of causality). So this could be traced back to Piaget's writings circa 1930 if you were so inclined.

The intro pages to Karniol [1980] give a nice summary of the modern interpretation of Piaget's moral stories, and examples of how kids sometimes take the causal story to what we consider an absurd extreme (e.g., the boy stole the bike $\Rightarrow$ the bridge collapsed). She also ran experiments on about 150 elementary school children. They were read skeletal stories of the form *Joe stole money. Later, Joe fell down the stairs.* or *Jane lied. Later, Jane fell in a puddle.* There were a range of types of causality, including immanent causality (the result is because of something inside the person), asyndetic[2] and/or mediated causality (it was the person's action, but mediated via another force), or chance causality (which is delightfully not jargon). Chance causality explanations were basically the least popular, ranging in use among the five grades from 16 to 34 percent; mediated causality ranged from 58 to 86 percent usage; immanent causality

---

[1]Lest you think I closely associate masculine with vice and feminine with virtue, recall that I flip a coin to determine the gender of all representative agents.

[2]*Syndetic*: Serving to unite or connect; connective, copulative.

ranged from 23 to 47%.

That's the first experiment; the final experiment, using only kids who'd given a mediated causality response in the first experiments, and a story in which the kid in the story gets struck by lightning, was able to induce a greater recourse to chance causality among the listeners (70%). But the first two experiments (and another story in the third experiment where the boy breaks his leg) still show that if there is no causal story spelled out, the brain of the listener will probably invent one. If you want more, Karniol gives a dozen or so other papers that come to similar conclusions: even the youngest kids will see a link between a person's actions and the eventual outcome when there is a relationship to be had, and will invent one when there isn't.

Variants of the story   Now that the canonical story is ingrained in us, hard, there are all sorts of variants that turn our causal expectations around. Some just make for a better story, but others begin to show flaws in the system.

[The ending to *Moby Dick* was so gut-wrenching because it was so outside of the entire framework. I'm a bit amazed that it got published and sold well enough that we've heard of it, given how much it bucks convention.]

Adult fiction is filled with what we call moral ambiguity, by which we mean that the virtuous aren't rewarded and the evil aren't punished. This is not to be confused with stories that create tension by allowing the bad guy to win halfway through, getting the princess or the thousand pounds of gold bullion [both props play the same rôle in the typical story]. In those half-win stories, tension comes from our knowledge that the inevitable downfall will only be worse after the temporary victory.

Many bookshelves have been filled with Dark Knight-type stories about characters of ambiguous virtue. But we humans have an easy solution for these stories: if we are firmly wired to see virtue $\Rightarrow$ reward, then we eventually start to see reward $\Rightarrow$ virtue. In logic class, it'd be a blatant error to conclude the second relation from the first, but we're not talking about logic, we're talking about how people think.

If you're an Objectivist, you learn that whatever it takes to gain reward is by definition virtuous. If you follow other sorts of commerce-oriented ethical systems, then you follow a similar but looser line. As the cliché goes, *might makes right*. On the other side, I've heard more than enough people give me a line like 'it's not illegal, so it's not unethical', which in this context means *no punishment $\Rightarrow$ not evil*.[3]

Or, it's easy for both kids and adults to misread what the cause was that led to the final outcome. It's downright cliché that the protagonist is attractive and the antagonist ugly, from which we are taught that *attractive $\Rightarrow$ reward*; *ugly $\Rightarrow$ punishment*.

If the virtuous are always rewarded and the evil always punished, then anybody who is being punished must be doing something wrong. If we see a person, or a group of people (grouped by language, size of nose, or genitalia), and find that they are doing worse than others, our brains work overtime to fill in the blank in the relation _____ $\Rightarrow$ *punishment*. E.g., if they hadn't eaten from the Tree of Knowledge of Good and Evil, they wouldn't be worse off. [Add this to the last paragraph, and we find that *unattractive = evil*, which

---

[3] It's not as if I know what the True and Correct ethical system is, but an ethical system that directly equates individual benefit with ethics is really just the state of nature calling itself ethics, and a rejection of the idea that we humans can develop beyond biology.

I find is really how a lot of people think.]

Now, all those stories are really just practice for what happens here in reality, where we write our own stories. The non-fiction evening news is making a huge effort to fulfill our expectations: the evil have to be punished, and (from time to time) the virtuous have to be rewarded. As viewers, our expectations about how the world should be are very high. If the assailant doesn't go to jail, then we're left with the frustration of a story cut short just before the resolution. If their country is evil, and our country is virtuous, then there is tension until we find a way to bring about some sort of punishment for them, preferably in a manner that brings rewards to our contractors.

And so we see a great deal of our legislative and interpersonal effort put into making sure that rewards and punishments are eventually paid out, even though the only real benefit may be the sense of resolution that comes from making the world fit the stories we were told as kids.

We all have these *virtue* $\Rightarrow$ *reward* and *evil* $\Rightarrow$ *punishment* relations tattooed to the inside of our foreheads. Our parents made sure of it, by teaching us ethical causal stories at the same time that we were learning more mechanistic causal stories. If they didn't present us such stories, we'd just make up our own. But the mechanical relationships like *I drop the plate* $\Rightarrow$ *the plate breaks* are much more robust than the relationship between nice behavior and reward, to the point that we can easily invent unverifiable relationships, like how a pretty face and big muscles implies virtue, or being homosexual is evil, or that whatever person we've never met before is getting exactly what he or she deserves. The ability to develop and understand causal stories, which makes us human, gives us ethical beliefs, and allows us to construct a society, is exactly the same force that lets us dress up self-interested behavior as virtue, makes us pine for retribution against perceived slights, and nudges us to wish ill upon those who look or behave differently from our ideal.

References

Rachel Karniol. A conceptual analysis of immanent justice responses in children. *Child Development*, 51(1):118–130, 1980.

Judea Perl. *Causality*. Cambridge University Press, March 2000.