

Hume on Causation

I. Recap of Hume on impressions/ideas

Perhaps the best way to understand Hume (1711-1776) is to place him in his historical context. Isaac Newton (1643-1727) had just been laying out the first comprehensive system of mechanics. Like many people, Hume was terrifically impressed with the way a few simple laws of motion sufficed to explain all the workings of physical nature; like many he felt that Newton had essentially completed the project of physics. How did Newton's system work? Stripped to essentials, there are three ingredients. First, the external world is composed of physical objects. Second, each object has associated with it a certain quantity of impetus or motion, what Newton called momentum. Third, the physical world evolves from moment to moment in accordance with a few basic laws, which say in effect that given such and such objects in such and such places with such and such momenta, such and such will happen next.

What Newton had done for the physical realm, Hume proposed to do for the mental one. Thus he would become the Newton of the mind. So, corresponding to Newton's physical objects, Hume has what he calls perceptions: these comprise all the things that occur in the mind. Corresponding to a physical object's momentum is what Hume calls the force and vivacity of a perception. And corresponding to Newton's laws of the physical realm, like the law of gravitation, Hume has laws of the mental realm -- he calls them *laws of association*. For instance, if perception A has frequently occurred together with perception B, and if A resembles B, then they will be attracted to each other, and each will tend to bring the other in its wake.

Recall, Hume's first and fundamental assumption is that perceptions are the mind's sole occupants, just as material objects are the only things in the physical realm. As we have seen, he sets out the following principles:

- (1) All there is in the mind is perceptions.
- (2) All perceptions are impressions or ideas; impressions have greater force and vivacity, ideas less.
- (3) Simple ideas are copies of simple impressions, and complex ideas are combinations of simple ideas.
- (4) Ideas tend to attract ideas of similar things, of contiguous things, and of causally related things.

Note that (4) gives us the "laws" that govern the mind. Remember that it's important for Hume that these are purely mechanical. Just as Newton banned purpose and reason from external nature, explaining the apple's falling by natural forces rather than a desire to return to mother earth, Hume wants to banish purpose and reason from the mind, explaining thinking by blind natural propensities rather than logical rules.

Given this, we can then see a pattern in Hume's work. Whenever Hume considers a philosophical concept -- and today we are going to look at the concept of causation -- he offers first a negative account of that concept and then a positive one. In the negative account, what he calls his 'skeptical problem,' he argues that claims involving this concept:

- A-) can never be justified (because perceptions are not enough to justify them);
- B-) are not even meaningful (because the meaning would be an idea not based on a perception, etc.)

Then in the positive account, his 'skeptical solution', he explains how these same claims

- A+) are natural and beneficial;
- B+) are 'getting at something' nevertheless.

From (A-) and (B-) it follows that if knowledge requires understanding and justification, then the skeptic is right: we cannot know the claims in question. Yet (A+) and (B+) pave the way to a different picture of our position that makes skeptic-proof knowledge seem much less important than before.

II. Recap of Hume on induction

Recall how we gain knowledge. On Hume's account, we need to consider both relations of ideas and matters of fact. Relations of ideas beliefs are justified because there is no way for them to be false, and we can tell that just

by inspection of our ideas. But how do we justify our beliefs about matters of fact, e.g., that the sun will rise tomorrow? His answer is this: “all reasonings concerning matters of fact seem to be founded on the relation of cause and effect” (20).

So, we think that the sun will rise because we think that the earth’s angular momentum will keep it spinning overnight. Or if we walk in and see writing on the board, that someone must have been here earlier because that’s the only way to get marks up there. And so on... Now, if Hume is right that all beliefs about matters of fact – meaning, all beliefs reaching beyond what is immediately present to us – depend on beliefs about cause and effect, then such beliefs are incredibly important and stands in need of immediate justification. But can they be justified?

What’s an example of a causal belief? Suppose, just to have a definite example, that I’ve got a big can of gasoline up here with me; and I’m trying to decide whether to toss a match into it.

(*Causation*) Igniting this gasoline will cause it to explode.

All of us have beliefs like this: what else explains our extreme reluctance to drop a match into the gas tank? If Hume is right that we can’t gain knowledge of (*Causation*) through reason, then someone with no experience of gasoline will have no idea at all what will result when a lit match is thrown into it. A priori reason alone won’t do it, so apparently experience is required. If I think that igniting the gas will produce an explosion, Hume says, that must be because I have observed ignited gas exploding and am projecting these observations into the future.

(*Induction*) In the past, igniting gas has caused it to explode. Therefore, in the future, igniting gas will cause it to explode.

But we saw last time that inductive inferences are unjustified. There is no rationally justifiable way to get from ‘all observed A’s have caused B’s’ to ‘from now on, A’s will cause Bs.’ Experience of the past simply does not give you a rational route to predictions about the future.¹ (Why? Because the inference relies on a premise that nature is uniform, which is itself a matter of fact that would have to be established by experience; but to establish it by experience, we would have to rely on that very premise, which is question begging.)

But this leaves an enormous mystery. Because even if there’s no rational route from experience to expectation, there has to be some connection! Take a particular case. If you see me dropping a lit cigarette into a barrel of gasoline, you run for cover expecting an explosion. Now, if you had no experience of gasoline, you wouldn’t expect the explosion. The question is, how can experience reasonably influence expectation? According to Hume, there has to be some kind of nonrational influence. Some natural mechanism is leading us to expect A’s to cause B’s when we’ve observed A’s to cause B’s in the past.

Here is how Hume describes the situation:

....wherever the repetition of any particular act or operation produces a propensity to renew the same act or operation, without being impelled by any reasoning or process of the understanding; we always say, that this propensity is the effect of custom. (28)

Applying this to the case at hand, when you get into the habit of having a B perception after an A perception, then the next time you have an A perception, custom leads you to have a B perception once more. According to Hume, nothing else can explain why we expect B’s to follow A’s after witnessing the A/B succession in the past:

This hypothesis seems...the only one which explains the difficulty why we draw, from a thousand instances, an inference which we are not able to draw from one instance, that is, in no respect, different from them. Reason is incapable of any such variation. The conclusions which it draws from considering one circle are the same which it would form after surveying all the circles in the universe. But no man, after seeing only one body move after being impelled by another, could infer that every other body will move after a like

Hume, David. *An Enquiry Concerning Human Understanding*. 2nd Edition. Hackett Publishing, 2012. © Hackett Publishing. All rights reserved. This content is excluded from our Creative Commons license. For more information, see <https://ocw.mit.edu/help/faq-fair-use/>.

¹ Note that the problem of induction need not be stated in temporal terms. The deeper question is how we can conclude that what we have experienced is a basis for knowledge of anything we haven’t experienced, past or future, i.e., to what extent are we entitled to assume that our experience is representative?

impulse. (28)

So, it's habit that tells us what to expect next:

Custom is the great guide of human life. It is that principle alone which renders our experience useful to us, and makes us expect, for the future, a similar train of events with those which have appeared in the past. (29)

Is Hume embarrassed about our reliance on blind natural impulse for our beliefs about the future? Not at all. Rather he's grateful that habit kicks in exactly where reason fails us.

Without the influence of custom, we should be entirely ignorant of every matter of fact beyond what is immediately present to the memory and senses. We should never know how to adjust means to ends, or to employ our natural powers in the production of any effect. There would be an end at once of all action, as well as of the chief part of speculation

But the point isn't just that habit is the only game in town. Hume actually thinks that even induction was rationally defensible, still it'd be better to have our future-directed beliefs handled by blind natural propensities. Prediction is so essential to our survival that our method of prediction should be as secure and efficient as possible. And reason would be in fact less secure, and less efficient, than custom is. Why?

- (1) Predicting would take too long (reasoning is slow).
- (2) Mistakes would occur (reasoning is error-prone)
- (3) Not all could predict (reasoning ability is variable).

This last takes some explanation. If we had to rely on reason in forming expectations about the future, then at times of diminished we would have no idea what to expect. So for example infants and old people, whose reasoning powers are weak or nonexistent, would be in effect future-blind. Babies could never learn to keep their hands away from the flame; and after their reason had deserted them, senior citizens would lose the ability to form expectations altogether.

So although Hume admits that induction is not rationally justifiable, he tries to explain why this is not really such a problem. As luck would have it we're so put together we're going to form inductive expectations whether it's reasonable or not. Moreover he tries to provide induction with a kind of practical justification: if we didn't form these expectations, we wouldn't be able to live.

Standing back for a minute from the details, Hume is saying something very radical about human nature. All of our most important beliefs derive not from any rational faculty but from the brute mechanical operations of our minds. To the extent then that we identify ourselves as fundamentally reasoning beings -- and this is how Descartes sees us -- we are misunderstanding what it is to be a human being. In effect then he's recommending to us a new, radically non-Cartesian, self-conception (closer to the animals than the angels.)

III. Hume on causation

Remember that Hume's skeptical problem about causation breaks up into two parts: on the one hand, we have no good evidence for our causal claims (this is the (A-) part; on the other hand, our claims about what causes what aren't even meaningful (this is the (B-) part). So take a typical causal statement like:

Heating ice causes it to melt.

This we can think of as made up of two components. First we have the conjunction component: Whenever heat is applied to ice, it melts. This Hume has already shown to be unjustifiable due to the problem of induction.

However there's more to the causal statement than just the *conjunction component*. What's missing? That there is some kind of causal oomph, some kind of power, some kind of "necessary connexion" between the heat and the melting, such that the heat is not just followed by the melting but *makes the ice melt*. Call this the *connection component* of the causal statement: the ice melts because it's heated. How do we know, though, that this is something genuinely additional? Because we can imagine, can't we, that the conjunction component holds and the causal statement fails. Examples?

So the connection component is genuinely additional to the conjunction component. And for Hume it is just as problematic; indeed in a way it's more problematic. Because where we know what the conjunction claim means, we don't really know what the connection component means. In other words when we say that one thing makes another happen, we really have no idea what we're talking about! And the reason is, that we have no real idea of "necessary connexion."

Isn't this crazy, though? We certainly seem to know what we mean by one thing making another happen. So there are two things Hume has got to explain to us. First, why it is that we don't have an idea of necessary connexion; second, how we fall under the illusion that we do have such an idea, that we do understand what we're talking about.

How can Hume possibly convince us that we have no idea of power, or force, or of one thing making another happen? Surely we can have an idea of anything we like! No, says Hume, because our powers to form ideas are not unlimited. Any idea we can frame to ourselves must break down to simple ideas, all of which are faded copies of impressions we have had. All right, then, let's look for the impression or impressions our idea of connexion comes from. First question is, is this a simple idea or a complex one? Introspection, Hume thinks, shows it to be simple. That is, the idea of the heat's making the ice melt does not seem to break down into anything simpler. This means that our idea of connexion, if we have one, has got to come from an impression of connexion. And now Hume asks, have we ever had such an impression? Granted that we often see heat applied to ice, and shortly thereafter ice melting, do we ever see the heat's forcing the ice to melt?

Hume looks three main places for this impression: our experience of causal interactions between external objects; our experience of our control over external objects, such as our limbs; our experience of our control over internal objects, that is, our ideas. But let's stick to his treatment of the first case, which is by far the most interesting. Do we ever experience the power by which one external thing produces another? Hume gives two interesting arguments that we don't (41ff).

i) If we observed the heat's power to melt the ice, then from this observation alone we could deduce –we could tell by reason alone, without benefit of any other experiences – that the ice was going to melt. Because given this power, the ice cannot help but melt. Yet it seems absolutely clear that the ice's melting can't be inferred from the observation of the cause: we can readily imagine the ice staying frozen despite what we perceive. So we can't be perceiving the power.

ii) Second kind of evidence is introspective: look inside yourself and ask yourself what you really experience when you see one billiard ball knocking another out of place. Two choices: either you see a simple conjunction, the first ball hitting the second followed by the second moving; or else you see a real connection, the one making the other move. Be honest now! If you carefully attend to your experience, Hume says, you find you perceive one thing after another, not one thing because of another.

Now, maybe you want to argue with Hume about this, claiming you do perceive the causal nexus. But ask yourself, how would your experience have been different if there were no necessary connection, say because the balls were controlled by magnets hidden underneath the table? Not at all: conjunction and connexion are in fact experientially indiscernible in this case. But if you'd have had the same experience even in the absence of connexion, how can you claim to be experiencing the connexion? (Suppose I say: I can see you through the door. Then if you walked away, my experience should change, and in fact it doesn't.)

Impossible though it seemed a few minutes ago, Hume has poked large holes in the assumption that we have an impression, and hence an idea, of necessary connexion. On the other hand, it certainly seems to us we have such an idea. So Hume has his work cut out for him. He's got to do justice to our subjective certainty that we know what 'this causes that' means when in fact we have no idea of necessary connexion.

Where does this subjective conviction come from? Here is how Hume reasons it out. Take first the case where you're having your first experience of the cause followed by the effect. Say you've never watched Star Trek before and you see Spock apply the Vulcan death grip to someone who then collapses onto the floor. This by itself isn't enough to induce the feeling of necessary connexion; if you're like me, it looks almost accidental that the Klingon collapsed at exactly that time. Watch Spock do this some more, though, and you do begin to

experience the death grip as making the Klingon fall. So whatever exactly the mechanism that induces this feeling of connexion, it must be something that requires several experiences of succession to start operating.

Now, what can the mechanism be that it requires several experiences of succession to go into operation? Think back to what we said about the nature of inductive inference. After one instance of E following C we have no tendency to expect the next E to be followed by a C. After several repetitions of Es following Cs though this tendency does in fact set in, e.g., Pavlov's dog. Since this internal tendency the only thing repeated exposure to C/E sequences provides us that a single C/E sequence does not, it must be the source of what we call our impression of necessary connexion. That is, this so-called impression, rather than deriving from anything in the outside world, is really just a reaction to our subjective tendency to expect the effect when we observe the cause. All our so-called impression of necessary connexion is a kind of by-product or spin-off of the association experience has set up between C-perceptions and E-perceptions.

Now call this feeling a 'so-called' impression of necessary connexion? Take the example of an impression of heat. By an impression of heat we mean an impression caused by heat and which informs us that heat is really out there. But our feeling of connexion does not really inform us of connexion. By far a better analogy is with the feelings of pleasure you get warming your hands before a fire. This feeling is not an impression of pleasure in the fire but just something that happens in your mind when you're perceiving a fire. Even better: the feeling of excitement you used to get when the ice cream truck rolls down the street. Surely there isn't excitement rolling down the street. So it is with the feeling of necessary connexion: it's just something that happens in your mind on perceiving frequently correlated events. This then is the analogy to bear in mind: the feeling of necessary connexion is not like the feeling of heat -- caused by real external heat -- but like that of pleasure or excitement.

Next question is, if the feeling of necessary connexion is like that of pleasure, or excitement, why are we inclined to attribute necessary connexion to external events? Hume explains this, in a wonderful phrase, by saying that the mind has a tendency to 'spread itself' on objects; that is we read the subjective feeling, actually thrown off by the operations of the mind, onto the external events, and thus see the cause as forcing the effect to come about. Observing external reality is thus for Hume like watching a movie. All you are really perceiving is a succession of causally unrelated images, but it doesn't feel that way because you read our feelings of connexion into the observed succession and thereby feel as though you are observing real connexions.

Because we lack any real idea of connexion the word "cause" is not strictly speaking meaningful. Is this to say that causal statements have no significance at all? Not quite. What is true is that they can't have the meanings we intended for them. But maybe some other and related meanings can be found. Hume's project here is to explain what causal statements will have to mean if they are going to mean anything.

Start by asking what we wanted causal statements to mean. Two things seem crucial. First, causation was supposed to be an objective relation. Second, it was supposed to involve some kind of connexion. But these conditions are irreconcilable it now turns out. What Hume does is propose two revised notions of causation, each of which meets one of the two conditions while violating the other. Thus

Objective conception: a cause is 'an object, followed by another, and where all the objects, similar to the first, are followed by objects similar to the second' (51).

This is all the external world can offer in the way of a content for our causal beliefs. But there is something crucial missing, namely the element of connexion. This can be had but only at the cost of objectivity. Thus Hume's second definition:

Subjective conception: a cause is 'an object followed by another, and whose appearance always conveys the thought to that other' (51).

What Hume is saying with these two definitions is basically that you can't have it both ways: no objective account can handle necessary connexion; no subjective account can do justice to our conviction that causation is "out there" and not merely in the mind.

MIT OpenCourseWare
<https://ocw.mit.edu>

24.01 Classics of Western Philosophy
Spring 2016

For information about citing these materials or our Terms of Use, visit: <https://ocw.mit.edu/terms>.