Does science give us knowledge?

Science appears to tell us things which go beyond what our senses directly tell us. Here are some examples:

All massive objects attract one another.

Every 24 hours, the earth rotates on its axis.

These claims are not, on a natural interpretation, claims which we can know to be true directly on the basis of sense experience: for example, though we can observe some massive objects attracting each other, we certainly have not observed this of all presently existing massive bodies, let alone all massive bodies past and future. These claims are generalizations.

Much of what science tells us is a matter of generalizations. Other things that science tells us are based on generalizations. An example might be

Halley's comet will next be visible from earth in 2061.

This is not itself a generalization; but our knowledge of it depends on our accepting certain generalizations about the movement of celestial bodies.

It is highly plausible that we should believe some of the generalizations which our best scientific theories endorse. But why is this? What rule of belief might explain this?

Here is a natural answer. We've already encountered the following two proposed rules of belief:

Experience → Belief
If your sense experience tells you that P, and you have no reason to think that your sense experience is misleading, believe P.

Proof → Belief
If you can prove P, believe P.

Even if our experiences don't directly tell us that certain generalizations are true, our experiences do seem to be part of the reason for believing those generalizations. One might think that we can give a kind of proof of the relevant generalizations based on those experiences. If so, the above two rules of belief put together might explain why we should endorse the findings of science.

It will be useful to have a simple example to discuss. Let’s suppose that I have a quantity of water, and I am wondering at what temperature that water will freeze. (Suppose that it is pure water, and that I am at sea level.) Then some elementary science tells me that:

```
This sample of water will freeze at 0°C.
```

What experiences might count in favor of this claim?

The answer seems pretty obvious. We have a whole host of observations of the form:

```
Sample 1 of water froze at 0°C.
Sample 2 of water froze at 0°C.
Sample 3 of water froze at 0°C.
.....
```

Suppose that, in accord with the first rule of belief just listed, I take all of these observations at face value. Why might these observations give me reason to believe the claim about the current sample?

Let’s try to construct an argument in the obvious way.

1. Sample 1 of water froze at 0°C.
2. Sample 2 of water froze at 0°C.
3. Sample 3 of water froze at 0°C.

.....

N. Sample N of water froze at 0°C.

———————————————-

C. This sample of water will freeze at 0°C. (1-N)

This argument is an example of enumerative induction — a kind of reasoning on which we seem to rely all of the time.

Is this argument valid?

This seems to ruin our initial thought that we can justify the claims of science on the basis of experience + proof.

Can you think of any premise which we can add to the argument which would make the argument valid?

Here's a natural choice:

If all past samples of water froze at 0°C, then this sample of water will freeze at 0°C.

1. Sample 1 of water froze at 0°C.
2. Sample 2 of water froze at 0°C.
3. Sample 3 of water froze at 0°C.
.....
N. Sample N of water froze at 0°C.
N+1. If all past samples of water froze at 0°C, then this sample of water will freeze at 0°C.
——————————————-
C. This sample of water will freeze at 0°C. (1-N+1)

Is this argument valid?

This looks like progress. If we should believe all of the premises of this argument, then it looks like we have an explanation of why we should believe the conclusion.

We already have an explanation of why we should believe premises 1-N. What about premise N+1?

We already have an explanation of why we should believe premises 1-N. What about premise N+1?

In our first reading, Hume addresses the question of whether we should believe premises like this by drawing a distinction between two different kinds of claims:

> "All the objects of human reason or inquiry may naturally be divided into two kinds, to wit, *relations of ideas*, and *matters of fact*. Of the first kind are the sciences of geometry, algebra, and arithmetic; and in short, every affirmation which is either intuitively or demonstratively certain. ...Propositions of this kind are discoverable by the mere operation of thought, without dependence on what is anywhere existent in the universe. ...
>
> Matters of fact, which are the second objects of human reason, are not ascertained in the same manner; ...The contrary of every matter of fact is still possible; because it can never imply a contradiction, and is conceived by the mind with the same facility and distinctness, as if ever so conformable to reality. *That the sun will not rise tomorrow* is no less intelligible a proposition, and implies no more contradiction than the affirmation, *that it will rise*. ..."

Premise N+1 appears to be like the claim that the sun will rise tomorrow: it is a matter of fact rather than a matter of the relations of ideas, and so cannot be known "by the mere operation of thought."

Premise N+1 appears to be like the claim that the sun will rise tomorrow: it is a matter of fact rather than a matter of the relations of ideas, and so cannot be known "by the mere operation of thought."

N+1. If all past samples of water froze at 0°C, then this sample of water will freeze at 0°C.

But if this cannot be known just by thought, it seems that we must believe it on the basis of experience. But do we have experiences which tell us that N+1 is true?

N+1 is an instance of a more general claim, which Hume calls the principle of the uniformity of nature:

The Uniformity of Nature
The future will be like the past.

It seems as though, if we should believe in the Uniformity of Nature, we should believe N+1. So the basic question is whether we should believe in the Uniformity of Nature.

The Uniformity of Nature
The future will be like the past.

It seems as though, if we should believe in the Uniformity of Nature, we should believe N+1. So the basic question is whether we should believe in the Uniformity of Nature.

Is the negation of this claim a contradiction?

It seems not. So it seems that, if we should believe it, we must believe it on the basis of experience.

But we don't have any experience which tells us directly that this principle is true. So we must know it on the basis of some series of experiences. And it might seem pretty clear what this series of experiences is.

After all, yesterday the future was like the past. And the same for the day before that. And this suggests an argument for the Uniformity of Nature.

After all, yesterday the future was like the past. And the same for the day before that. And this suggests an argument for the Uniformity of Nature.

1. Yesterday, the future was like the past.
2. The day before yesterday, the future was like the past.
3. The day before the day before yesterday, the future was like the past.

.....

N. N days ago, the future was like the past.

——————————————————-

C. Today, the future will be like the past. (1-N)

Is this argument valid?

What extra premise would make the argument valid?

It is hard to see how we could make the argument valid without adding a premise which was just a restatement of the very claim — the Uniformity of Nature — which we were trying to prove.

This line of argument from Hume is sometimes called “the problem of induction.” Because scientific reasoning seems to rely on induction, it is a problem with understanding why we should believe the claims of science which go beyond our experience.

Notice that we cannot avoid the problem by abandoning our belief in

This sample of water will freeze at 0°C.

In favor of some weaker claim like

It is probable that this sample of water will freeze at 0°C.

To get even this claim, we would need to rely on the claim that it is probable that the future will be like the past. But the negation of that claim also seems clearly intelligible, and it is no easier to argue for it than it is to argue for our original Uniformity of Nature principle.

It is worth being clear about what that problem is. We have not given a direct argument that the use of enumerative induction is irrational; rather, we have shown that it seems very difficult to give a justification of enumerative induction which is not circular, in the sense that it presupposes the legitimacy of inductive reasoning.

It is interesting to compare induction in this respect with deduction: the formation of beliefs on the basis of valid arguments.

Suppose that I believe some claim *Z*, and you ask me to provide my justification for believing it; I might respond with an argument of the following form:

(A) A.
(B) If A then Z.
———————
(Z) Z.

Suppose now that you respond like this: “Yes, I concede that both premises are true. But I still don’t see why I should accept *Z*!”

(A) A.
(B) If A then Z.
———————
(Z) Z.

Suppose now that you respond like this: "Yes, I concede that both premises are true. But I still don't see why I should accept *Z!*"

How should I reply?

It seems that I should say something like this: "Look, if (A) is true, and (B) is true, then *Z* has to be true."

Now suppose you respond like this:

"Yes, I see your point. What you are saying is that if (A) and (B) are true, then (Z) is true. I grant this. So let's add the following premise to our argument:"

(C) If (A) and (B) are true, then (Z) is true.

"Yes, I see your point. What you are saying is that if (A) and (B) are true, then (Z) is true. I grant this. So let's add the following premise to our argument:"

(A) A.
(B) If A then Z.
(C) If (A) and (B) are true, then (Z) is true.
————————————-
(Z) Z.

"But, you continue, I still don't accept Z."

This little story comes from a short piece by Charles Dodgson (also known as Lewis Carroll) entitled "What the Tortoise said to Achilles." (In the story, Tortoise is playing your role, and Achilles is playing mine.)

Here's the continuation of the story from this point.

If you accept *A* and *B* and *C*, you *must* accept *Z*."

"And why *must* I?"

"Because it follows *logically* from them. If *A* and *B* and *C* are true, *Z* *must* be true. You don't dispute *that*, I imagine?"

"If *A* and *B* and *C* are true, *Z* *must* be true," the Tortoise thoughtfully repeated. "That's *another* Hypothetical, isn't it? And, if I failed to see its truth, I might accept *A* and *B* and *C*, and *still* not accept *Z*, mightn't I?"

"You might," the candid hero admitted; "though such obtuseness would certainly be phenomenal. Still, the event is *possible*. So I must ask you to grant *one* more Hypothetical."

"Very good. I'm quite willing to grant it, as soon as you've written it down. We will call it

(*D*) If *A* and *B* and *C* are true, *Z* must be true.

Have you entered that in your note-book?"

"I *have*!" Achilles joyfully exclaimed, as he ran the pencil into its sheath. "And at last we've got to the end of this ideal race-course! Now that you accept *A* and *B* and *C* and *D*, *of course* you accept *Z*."

"Do I?" said the Tortoise innocently. "Let's make that quite clear. I accept *A* and *B* and *C* and *D*. Suppose I *still* refused to accept *Z*?"

"Then Logic would take you by the throat, and *force* you to do it!" Achilles triumphantly replied. "Logic would tell you 'You ca'n't help yourself. Now that you've accepted A and B and C and D, you *must* accept Z!' So you've no choice, you see."

"Whatever *Logic* is good enough to tell me is worth *writing down*," said the Tortoise. "So enter it in your book, please. We will call it

(E) If A and B and C and D are true, Z must be true. Until I've granted *that*, of course I needn't grant Z. So it's quite a *necessary* step, you see?"

"I see," said Achilles; and there was a touch of sadness in his tone.

Here the narrator, having pressing business at the Bank, was obliged to leave the happy pair, and did not again pass the spot until some months afterwards. When he did so, Achilles was still seated on the back of the much-enduring Tortoise, and was writing in his note-book, which appeared to be nearly full. The Tortoise was saying "Have you got that last step written down? Unless I've lost count, that makes a thousand and one. There are several millions more to come.

What does this story tell us about the justification of deductive reasoning? Can we think of the Tortoise as someone who refuses to accept the legitimacy of deductive reasoning unless given a noncircular justification for it?

What does this story tell us about the justification of deductive reasoning? Can we think of the Tortoise as someone who refuses to accept the legitimacy of deductive reasoning unless given a noncircular justification for it?

One thing this example suggests that is that deductive reasoning can only be justified by deductive reasoning; if so, perhaps induction is not in such bad shape, even if we can give no answer to Hume's problem.

Perhaps, then, we should just adopt the following as a rule of belief:

Induction → Belief
If you have inductive support for P, believe P.

Maybe Hume shows that we can give no non-circular argument for this rule. But perhaps the example of Achilles and the Tortoise shows that we can also give no non-circular justification for this rule:

Proof → Belief
If you can prove P, believe P.

Perhaps, then, we should just adopt the following as a rule of belief:

Induction → Belief
If you have inductive support for P, believe P.

Maybe Hume shows that we can give no non-circular argument for this rule. But perhaps the example of Achilles and the Tortoise shows that we can also give no non-circular justification for this rule:

Proof → Belief
If you can prove P, believe P.

But surely even if that is right then we should not reject Proof → Belief. So perhaps even if Hume is right we should not reject Induction → Belief.

Induction → Belief
If you have inductive support for P, believe P.

Let's return to our original argument:

1. Sample 1 of water froze at 0°C.
2. Sample 2 of water froze at 0°C.
3. Sample 3 of water froze at 0°C.
.....
N. Sample N of water froze at 0°C.
——————————————————-
C. This sample of water will freeze at 0°C. (1-N)

This argument is invalid. But it shows that the conclusion has strong inductive support; so if the above rule of belief is a good one, this shows that we should believe the conclusion. (And, by extension, we should believe other scientific claims made on similar inductive grounds.)

Induction → Belief
If you have inductive support for P, believe P.

On the other hand, this point still leaves us with a bit of a puzzle. It seems clear that inductive and deductive reasoning are better ways of forming beliefs than, for example, astrology. But what would this difference consist in, if we could give an argument, using premises from astrology, for the reliability of the astrological method of belief formation?

Induction → Belief
If you have inductive support for P, believe P.

Let's turn now to a different kind of worry about this proposed rule of belief. The worry is that, surprisingly, there are cases which appear to be straightforward counterexamples to it. This is a different, and in some ways more serious, challenge to scientific reasoning than the one that Hume raised.

This challenge is due to Nelson Goodman, one of the most important American philosophers of the 20th century.

Goodman's aim in his book *Fact, Fiction, and Forecast* was to show that rules of belief like ours are false; he did by defining a made up word, "grue," as follows:

x is grue if and only if either: (i) x is green, and has been observed before 2019, or (ii) x is blue, and has not been observed before 2019.

x is grue if and only if either: (i) x is green, and has been observed before 2019, or (ii) x is blue, and has not been observed before 2019.

It is important to see, first, that this is a perfectly legitimate definition; it succeeds in classifying all objects as either grue or non-grue.

But suppose that we enumerate all of the emeralds which have been observed so far, and consider the following pieces of data:

1. Emeralds first observed in 2019 were grue.
2. Emeralds first observed in 2018 were grue.
3. Emeralds first observed in 2017 were grue.

....

Now suppose that it is January 1, 2020, and you are going emerald hunting. If you accept our Induction → Belief rule, the following argument might occur to you.

Now suppose that it is January 1, 2020, and you are going emerald hunting. If you accept our Induction → Belief rule, the following argument might occur to you.

1. Emeralds first observed in 2019 were grue.
2. Emeralds first observed in 2018 were grue.
3. Emeralds first observed in 2017 were grue.

……..

———————————-

C. The next emerald I find will be grue.

Would it be reasonable for you to believe the conclusion of this argument?

Of course not; the next emerald you discover will be green and, since it was not observed before 2020, will not be grue. So it looks like Induction → Belief is false.

A very natural reaction is: this is a silly example! It would be crazy just to throw out all inductive reasoning on the basis of "grue."

Perhaps what we need to do is to restrict the cases of induction that we use to avoid annoying examples like "grue;" a natural thought is that we should restrict them to cases in which only suitable scientific vocabulary is used. (Words like "grue" that we want to rule out are sometimes called "gruesome predicates.")

Perhaps what we need to do is to restrict the cases of induction that we use to avoid annoying examples like “grue;” a natural thought is that we should restrict them to cases in which only suitable scientific vocabulary is used. (Words like “grue” that we want to rule out are sometimes called “gruesome predicates.”)

To pursue this thought, we need to be able to say what a gruesome predicate is - that is, we need to be able to say what, exactly, is so bad about “grue.” This turns out to be harder than you might think.

A first thought is that the problem is due to “grue” being a made-up word. But this won’t get us very far — after all, scientific theories introduce new scientific terms all the time, and these are “made up” in just the way that “grue” is — they are new terms defined in terms of existing vocabulary. At one time, “electron” was made up.

A more promising idea is that the problem with “grue” is that it is defined in terms of a particular time.

A more promising idea is that the problem with "grue" is that it is defined in terms of a particular time.

However, there are a few problems with this suggestion. One is that any predicate can be given a similarly time-indexed definition. For suppose that we define a new term, "bleen", as follows: x is bleen if and only if either x is blue, and has been observed before 2020, or x is green, and has not been observed before 2020. Using "grue" and "bleen" we can then give the following definition of "blue":

x is blue if and only if either: (i) x is bleen, and has been observed before 2019, or (ii) x is grue, and has not been observed before 2019.

One might reply: "Yes, one can define "blue" this way - but we don't have to. The difference between "grue" and "blue" is that no one could understand "grue" without this sort of time-indexed definition." This suggests that we should exclude terms which are impossible to understand except via a time-indexed definition.

One might wonder why we should be so sure that, for example, aliens quite different from ourselves could not find "grue" quite easy to understand without such a definition, and find "blue" rather confusing. But set that aside; there are two further worries about the proposed restriction on admissible vocabulary.

One might wonder why we should be so sure that, for example, aliens quite different from ourselves could not find "grue" quite easy to understand without such a definition, and find "blue" rather confusing. But set that aside; there are two further worries about the proposed restriction on admissible vocabulary.

The first is that this restriction is not restrictive enough: one can concoct gruesome predicates which are not defined in terms of times - for example, if all the emeralds which have been observed are from 17 emerald mines, we could define "grue" in terms of place. Or, if all the emeralds in the world have been seen by one person, we could define "grue" in terms of what has been observed by that person.

The second worry is that it is too restrictive: after all, we might be interested in investigating theories which are only about particular times, and places, and people - we don't want our theory of confirmation to simply fail to apply to such theories.

The idea that we can save Induction → Belief by restricting it to a certain privileged class of vocabulary is thus — while initially promising — hard to carry out.

Let's pursue a different idea, which involves a more sweeping rejection of the idea that one should in general accept the consequences of enumerative inductive arguments.

Let's pursue a different idea, which involves a more sweeping rejection of the idea that one should in general accept the consequences of enumerative inductive arguments.

This idea involves the claim that whether a piece of evidence counts in favor of a theory depends partly on our background beliefs about the subject matter in question.

Consider, for example, the following piece of evidence:

Every lobster I have seen has been pink.

Now suppose that every lobster I have seen has been in a restaurant; and I know that lobsters in restaurants are pink because they are boiled. Given this knowledge it would, it seems, be absurd for me to take my observations of lobsters to confirm the generalization:

Every lobster is pink.

Why? A natural thought goes something like this: I know that all the instances of this generalization I have observed have a certain property — being boiled in a restaurant — which explains why they are instances of the generalization. Moreover, I know that not all lobsters have this property — some are still in the wild. Whenever this is the case, one might think, the instances of a generalization fail to count in favor of the generalization.

Why? A natural thought goes something like this: I know that all the instances of this generalization I have observed have a certain property — being boiled in a restaurant — which explains why they are instances of the generalization. Moreover, I know that not all lobsters have this property — some are still in the wild. Whenever this is the case, one might think, the instances of a generalization fail to count in favor of the generalization.

This might be laid out in the following (cumbersome) rule of belief:

Induction → Belief
If you have observed that many A's are B, and there is no property F such that (i) you believe that the observed A's are B because they are F, and (ii) there are some A's that are not F, then believe that all A's are B.

What would this rule say about our original "grue" argument?

Induction → Belief
If you have observed that many A's are B, and there is no property F such that (i) you believe that the observed A's are B because they are F, and (ii) there are some A's that are not F, then believe that all A's are B.

One interesting consequence of this sort of approach - something which Goodman also took the example of "grue" to illustrate - is that there can be no such thing as the "logic" of scientific theory confirmation. If the above is right, we can never tell when some evidence confirms a theory just by looking at the evidence and the theory - in the way that we can look at a deductive argument and tell, just by looking at the premises and conclusion, whether it is valid.

If this is right, it does not really make sense to ask, without specifying a person or set of background beliefs, whether some evidence supports a theory — or even whether the theory is, in general, well-supported by the evidence. In general, it will be true that evidence can confirm a theory relative to person A but not relative to person B. Does this undercut the idea that the scientific method provides a method of belief formation which is rational for everyone?