Research Methodology
Prof. Soumitro Banerjee
Department of Physical Sciences
Indian Institute of Science Education and Research, Kolkata

Lecture - 56
Importance of Theory-Building in Science

We have seen certain methods of model building or theoretical research, but this is by no means an exhaustive description, because theory building is at the heart of most sciences.

Faced with a question, we produce a hypothesis or a postulate, and then we try to test the hypothesis or postulate. This is the basic method and therefore, the scientists' imagination has to go beyond what is observed and what is measured. What is observed and measured may produce the initial questions that the scientist tries to answer. But in trying to answer, one has to look into what goes beneath, what might not be immediately visible to us.

For example, when we saw that hydrogen and oxygen react in a specific ratio to produce water, we had no clue why that happens. But then the underlying process was pointed out by John Dalton, who said that everything is made of atoms and when any reaction happens, atoms react with each other. Therefore, two atoms of hydrogen are reacting with one atom of oxygen and that is why we get a particular 2 is to 1 ratio. But when he said this, atoms were not visible, molecules were not visible. So, the theorist had to go beyond what is immediately visible.

So far, I was talking mostly in terms of the experimental data giving rise to some kind of a functional form, or deriving the functional form, or say, dimensional analysis. These are empirically biased theory building. But often we have to build theories in order to explain certain phenomena. But the theory may point to things that might not have been observed. So, all theory are not guided or chained by what is observed. The theorist has to unleash the imagination beyond what is immediately observed.

Let me give a few examples. I have already given the example of the atoms and molecules. Theorists had to imagine the existence of atoms and molecules at a time when atoms and molecules were not visible, in order to explain certain observed phenomena.

Some philosophers argued that things that are not observed are not real, and therefore, do not count on that. Ultimately that was shown to be an error.

There may be a time when we might not be able to see that, but still the theorist should not be constrained by not being able to see those things. He/she should try to figure out what are the experimentally testable predictions of the theory. That is exactly what Einstein did.

When Paul Dirac, through his calculations, made the bold statement that if there are particles there should be also be anti-particles; if there is electron there should also be a positron, positively charged, but particle with similar character, then there was no experimental evidence on which he was based. He was based solely on theoretical grounds, but then he said: go forth and check whether you find that kind of particles or not, and it was found.

Similarly, when Wolfgang Pauli talked about neutrinos, there was no inkling of any observation of neutrinos. He theorized that on the basis of purely theoretical consideration. But then the theory also pointed out the conditions of that particle, so that the experimentalists knew what conditions to create in order to observe that particle.

So, all these theories had experimentally testable predictions and that structure told the experimentalist what to do. Similar was the case of the Higgs Boson. Higgs boson was proposed in the early 1960s, and it was actually observed some more than 50 years later. Gravitational waves: it was proposed in 1916 and it was observed 100 years later.

These were possible because the theorists concerned were bold enough to make the statement that my theory predicts the observation of this. So, go forth and try to observe it. If you observe it, then my theory is true. If you do not observe it, then my theory is false. Unless the theorist takes such steps and tries to predict, tries to go beyond what is immediately visible to us, science cannot progress. So, much of the activity of a theorist is essentially going beyond what is immediately observed.

And that is why it is said that theory is a very important pursuit in science. Because unless theory shows the light, the experimentalists cannot do any experiment blindly. All experimentalists do experiments with some theoretical expectation in mind. What is the theoretically predicted result? That is what he or she is trying to check. There has to be a theoretical grounding behind every experiment. Therefore, theory is a very important pursuit in science.

But, as we just said, there is no universal prescription to that. Each theoretical research has its own specificities. When Dirac made his prediction, when Pauli made his prediction, or when Higgs and other people made the prediction regarding the existence of a new particle, they were basing themselves on different considerations. Therefore, there is no universal prescription. It is just that a theorist has to be bold enough to see beyond what is immediately visible.

Now, we are coming to the end of a part of the course that concerned with the philosophical grounding and the actual things that a scientist does. And in the next part of the course, we will come to dissemination of scientific results, scientific communication, scientific ethics, and things like that.

So, let us round up this part with a few words. The part where I was talking regarding the philosophical issues and the part where I was talking about the actual things that a scientist does, these are not disconnected parts of the course. The philosophical issues actually produce the grounding or the logic behind what we actually do.

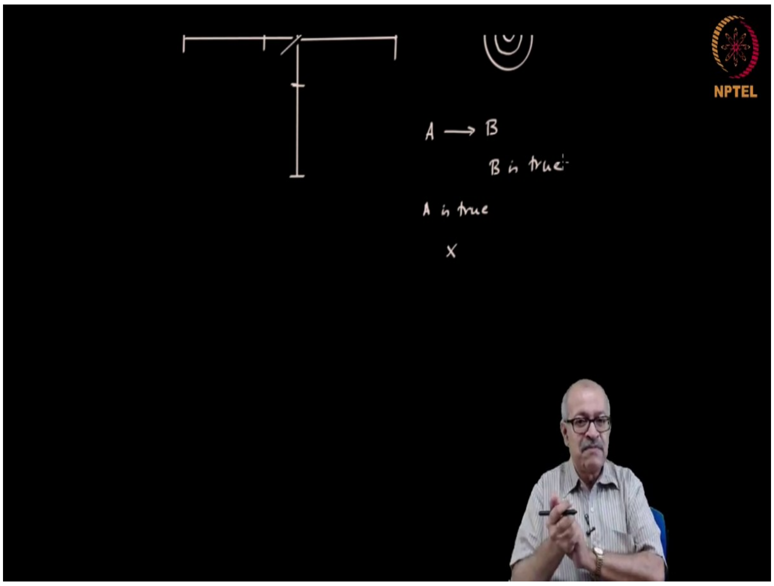
For example we have seen that earlier the way of thinking was subjective and post-Renaissance, the way of thinking became objective. And nowadays science bases itself on objective way of thinking. Then we saw in hypothesis testing, how we try to eliminate all possibilities of subjective judgment from a experimental procedure.

For example, we exclude our possibility of our being subjectively biased by incorporating the single blind test; we eliminate the possibility of the experimental subjects bias by planning a double blind test with placebo. So, all these are actually ways of eliminating subjective bias of the experimenter or the experimental subject.

Now you might notice: in all these examples that I have given, I mostly chose examples from biology. That was deliberate, because I come from non-biological background. Normally students or scientists have, sort of, a compartmentalized view, in which they are concerned about their discipline and their discipline only. Since many of the people I will be addressing to in this class may be physicists or chemists, so I wanted to give biology examples, so that their vision is widened.

But you might ask: why did not you give any example from physics? Yes there can be. Take, for example, the issue of the discovery of gravitational waves. The theory predicted that there would be expansion and contraction of lengths if a gravitational wave passes through the Earth and the expansion and contraction would be less than the radius of a proton. As small as that. If I take a stick about 100 meters in length, that will vibrate and that vibration's intensity will be of the order of the radius of a proton. How to measure that?

(Refer Slide Time: 11:47)



What they did was essentially that there was a laser source coming and that was broken into parts by a mirror. The result of the mirror was that a part would go through, and a part would be reflected. These two parts will go in perpendicular directions, There will be a mirror here, there will be another mirror here. I am grossly describing this without going to the specificities and details. Because of these mirrors, the laser beams will undergo repeated reflections and will oscillate like this. At the end, they would again be combined and observed somewhere here.

So, they will be allowed to go in this direction from this and then it will fall on a screen. Now the whole arrangement is such that at the centre of the screen, in this part, it should be dark. Why? Because the light coming from this branch and that coming from that branch will destructively interfere with each other.

So, at the middle there should be darkness, but then there should be a lit part and a dark part, a lit part and a dark part, so on and so forth. But their focus was on what is observed at the centre and normally no light would come at the centre. But then, the phase of the two lights will depend on their lengths.

If, now, a gravitational wave comes, and this one elongates by one proton size and this one does nothing because they are perpendicular to each other. If a wave comes in this direction then there would be a change in this direction not in this direction, in that case the length will change. As a result, the phase will change, as a result some light will be detected at the center.

So, if there is a photosensitive detector at that point, it will register some oscillation. That was the essential idea. But then, notice that such a very small oscillation can happen due to various reasons: there may be a highway passing nearby and cars passing on that highway might induce that kind of a oscillation. Maybe a caretaker going into that room and his footsteps might induce that kind of oscillation.

How would you distinguish between them? Because of that, they put two similar apparatus in two different parts of the United States: one in the North West part and another in the South East part. So, are they are spatially very far apart, so that one oscillation will not have a counterpart in the other place if there is an oscillation due to such reasons. They took enormous care to ensure that such external vibrations do not influence the result.

But nevertheless, there is still a non-zero probability of such vibrations coming. But the vibration that is coming in one place should not be the same as the vibration coming in another place. So, by tallying these two, they would be able to say whether it is a gravitational wave or not.

But still, scientists often do believe things and most scientists did believe that Einstein was correct. Because of that sort of belief, they might be biased in concluding in favour of gravitational wave. So, there was an additional thing. There was an additional arrangement of artificially inducing oscillations which the scientists working there would not know. Somebody other than these people would come and at any arbitrary time induce that oscillation.

And if the scientists can detect that this is actually not a gravitational wave, an oscillation that looks like a gravitational wave, that is not a gravitational wave, then you know that we are safe from personal biases. So, all these arrangements are put into the system just to eliminate any possibility of subjective bias, to ensure that the understanding that ultimately emerges out of it is objective.

So, you see, objectivity is built into the system. We have also learnt in the logic part that there is a logical error called affirming the consequent that happens when you have a logical structure ‘A leads to B’ and ‘B is true’, ‘therefore A is true’. This is logically incorrect. We have learnt that and because of that, we always try to propose as many hypotheses as possible and we try to check out each one.

But in testing each hypothesis, we say that we count on the null hypothesis and only when we have enough evidence to say that the null hypothesis is not true, then only we say that the alternative is true. Why do you do that? Because of this. There is another reason. I said that a null hypothesis is the one that leads to an equality. It is always an equality kind of hypothesis. It is an equality that can be tested using the statistical techniques.

If there are two populations and the means are assumed to be equal, then we can find out the probability of obtaining a difference that we have actually obtained by experiment. If that is very unlikely to be observed, then we can say that it is very unlikely that our initial assumption was correct which is that the two means are the same. This comes from the logical structure that we talked about earlier.

As Einstein once commented, a thousand experiments cannot prove my theory, but one experiment can disprove my theory. This is very important understanding in science. A thousand experiments cannot prove a theory, but just a one experiment, if it contradicts the expectation from the theory, then it disproves the theory.

So, these are the things that we learned in this part of the course and in the next part of the course, we will learn how to expose what we have learnt. And what we have learnt so far will form the basis of what we learn in the next stage.