6 The Politics of Testing

We are going to use the insights that we have already gained in this subject and take them into a new area which we have not focused upon before: experimental tests. It is very important to realise what actually occurs during experiments in science. So many of our public and political disputes are disputes among experts concerning the environment, technological development or policy, which rest on claims that this or that test or inquiry has established a certain truth. We have to understand how testing works in real "science" so that we will be in a better position to understand how the world of scientific and technological controversy works .

In these terms an experiment may be described as the use of some sort of hardware, instrumentation or measuring device to observe "facts" and in particular to carry out tests. We have already realised that human observation in general is produced in the light of background knowledge, beliefs and values. Observation in experimentation is no different. To a great extent the theories that scientists believe in, the aims and values that they hold, are not just inside the scientists minds, inside their conceptual grids, and these theories, aims and values have already been built into the hardware that they use to observe the "world". The point I'm driving at is that when we test with hardware and experimental apparatus we do not get away from 'theory-loading' or away from the fact that perception is coloured and shaped. In a sense the situation is all the more so because the human actors involved have externalised their beliefs in the hardware, and they then view or measure the world through their hardware.

In the story of method, testing, of course plays an essential part. Method depends on critical testing at very important points. If you have an hypothesis (according to the method story) or some kind of suggested theory and you are to determine whether you believe it or not, you put that hypothesis to a trial of truth. This is what testing is portrayed as in the method story. The good thing about such a trial is that if it is properly carried out there is no human subjective or social intervention in the outcome, because the sole judge and jury is nature or the facts. If testing is a trial against nature, then scientists believe as a result that they can produce untainted knowledge of the world.

However, we are going to argue that all testing is dependent upon human value judgements. It is not just during testing performed by biased selfinterested people that value judgements arise, for no test can be conducted in science without human value judgements being an essential part of the outcome. Human beliefs, values and aims are necessary to determine the outcome of a test. This may sound ridiculous, but I am going to try to persuade you that this is the case, and that it does not make science 'bad' or 'wrong', for it is simply the way that it is, and this needs to be clearly understood.

When I say that the outcome of a test is a human value judgement, I mean it is a judgment that emerges in a relevant group of judges. It is an outcome of social, human interactions amongst the members of that group, and it is a political result in the sense that all the members of the relevant group exert their power and influence toward rendering a final judgement. Testing is a human social institution. The outcomes of tests are social political events in that institution. Do not misinterpret what I am about to say. I am not saying that tests do not occur; or that they are not real; I am saying that tests are not trials simply and directly against objective nature. Tests are social and political events, as are their outcomes.

We will now examine a simple case of a test and its politics. In general what occurs when a test is pursued in science? Obviously, some hypothesis or guess at a law or theory is being put on the rack ie: being trialed or tested. We already know that hypotheses are not simply arrived at by observing the facts of nature in an unbiased way. Laviosier did not arrive at his oxygen hypothesis just by looking, but by viewing certain procedures in tandem with certain newly rejigged concepts. His 'discovery' of oxygen was just the linking of certain material practices with certain (revised) bits of his conceptual grid. Oxygen is not just a concept, nor is it a simple object known in commonsense everyday experience, it is that linkage, that seeing and understanding of those practices through these revised concepts.

When you conduct a test you are actually testing a prediction which arises from the theory or hypothesis. [Fig. 1] A prediction is a statement about what will happen which is derived from your hypothesis. Usually this means some kind of mathematical or logical manipulation of the hypothesis to produce some observable prediction. You often have to take further steps in order to gain a prediction from an hypothesis, such as making unexamined or half-examined guesses and assumptions. [Fig. 1a] Usually you cannot directly derive your prediction from your hypothesis for you have to feed other things into the derivation, such as guesses, assumptions, information, other 'accepted facts'.

FIGURE 1

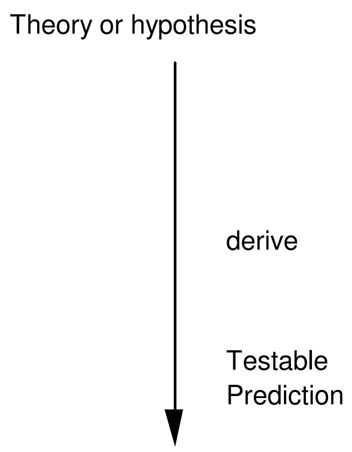
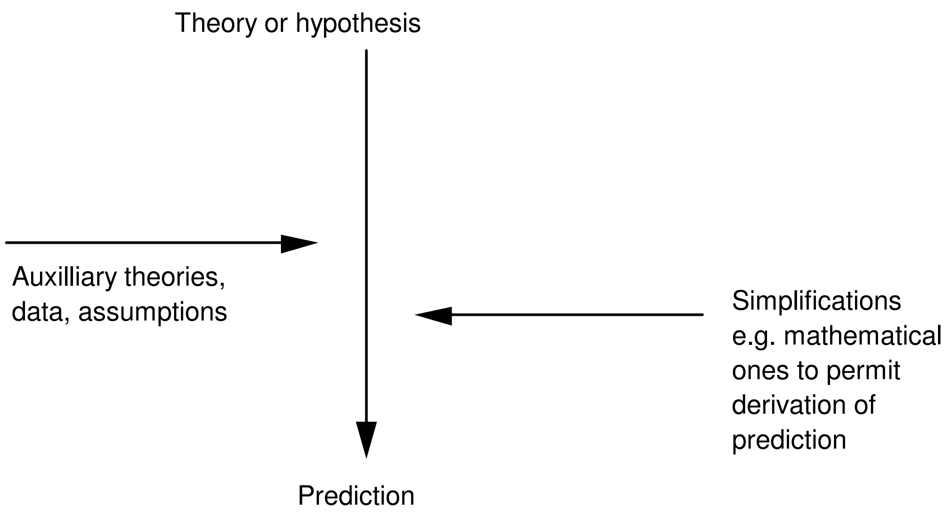


FIGURE 1a



What your experimental set-up produces is usually called 'data' or 'test results'. What we are looking at in a test is the relationship between the data and the prediction. If the data does not 'match the prediction' we have strong grounds for abandoning our hypothesis or questioning our theory. If the data 'agrees with the prediction' then we feel confident about our hypothesis. However, this explanation is too simple and as soon as it becomes a little more complicated the situation correspondingly becomes much more interesting and much more political.

In order to explore this in concrete terms let us consider one of Galileo's famous tests. Galileo, the great 17th century physicist and astronomer helped to establish modern physics as well as the sun-centred system of the universe.

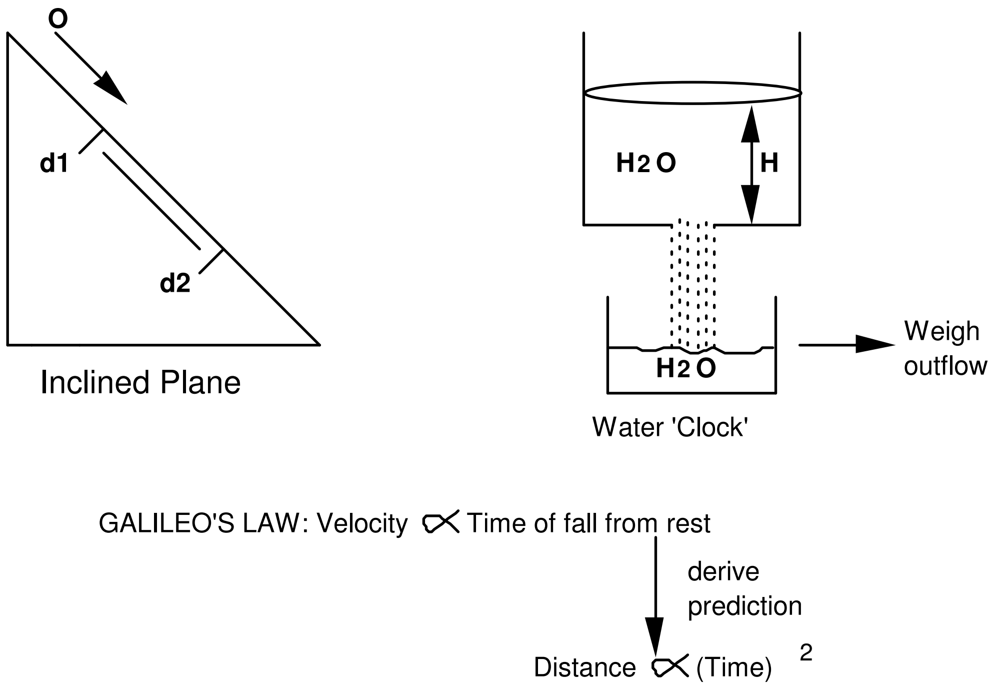
In 1638 Galileo published a book about physics and although a lot of what he proposed is no longer accepted in modern physics, it is very much the first book of what we call today Classical or Newtonian physics. One of the most important parts of Galileo's physics was the mathematical Law of Falling Bodies. This is the mathematical description of how bodies fall under ideal conditions when air resistance and friction are not considered to be important. This sounds elementary but it was the first hypothesis that was established as a law in Classical Physics.

Galileo was testing the claim that if a body is dropped it will travel in the following way: the distance travelled from the start will be proportional to the square of the elapsed time of falling. If the time of fall is doubled the body will travel four times as far; if the time of fall is trebled then it will travel nine times far. The distance [d] fallen from rest (zero) is proportional to the square of the time [t] of the fall (if the proportionality constant k=1, $d=t^2$). Galileo derived this theorem from his more basic hypothesis that velocity is proportional to the time of fall. Thus, when a body falls under ideal conditions its speed increases proportionally to its time of fall. This basic hypothesis was less immediately testable than his derived hypothesis or prediction of $d=t^2$ (Galileo did not derive his testable hypothesis quite the way we would today for we would use calculus of which he had no knowledge; he used Medieval and Greek mathematics.)

Galileo stated that he did experiments but many people have doubted this claim for there was little detail of them in his book. But we know from some of his manuscripts that he actually did perform experiments and how he conducted them. To test $d=t^2$, Galileo built an inclined plane down which he intended to roll the best steel balls that could be found in the 1600s. Presumably these balls were not manufactured from the highest grade steel in outer space, therefore, they were not perfectly spherical, but they were produced to the standards of the best Venetian or Florentine technology of the time.

Galileo intended to roll the steel balls down the inclined plane in order to measure off distances. He compared the distances rolled because he regarded this 'rolling' as essentially 'falling' and he also determined the time, by using a water clock. [Fig. 2] After one opens up the orifice at time zero, water pours through, during the time the ball 'falls' the set distance for the trial, and as the distance is completed, you instantaneously close the orifice. A certain amount of water is available to be weighed and the amount of water is proportional to the length of time the orifice has been open, that is, proportional to the time of fall. To a first approximation this is true, but not according to modern physics or to the physics of 1700, as opposed to 1638.

FIGURE 2



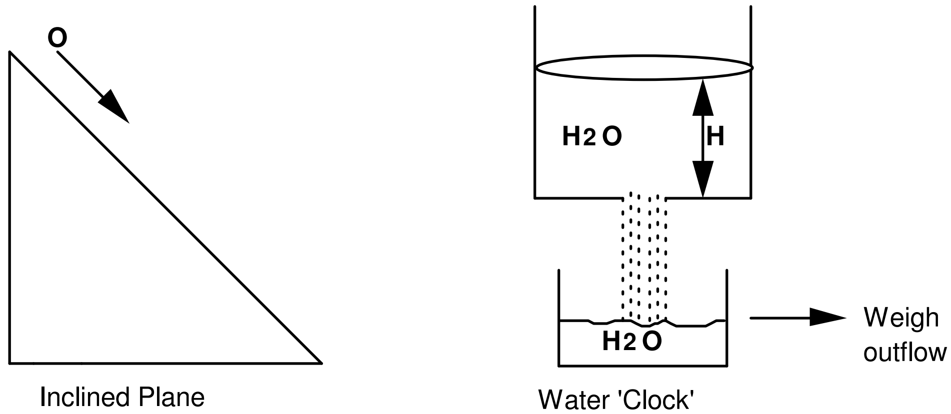
As Newton and others later demonstrated, the amount of water caught, however, is not proportional to the amount of time that the orifice has been open because as the level of water in the top goes down the rate of flow through the hole changes. In fact, it changes continuously as a function of the decreasing height of the head of the water. You need calculus to establish a better calibration of this clock but Galileo did not know this and to a first approximation probably would not have cared. For Galileo the clock is calibrated as follows: the amount of water caught is proportional to the time elapsed. He is utilising a simplified assumption (theory) about how the clock works.

This illustrates that there is often a lot of theory and assumption in the background of the performance of an experiment. How could Galileo perform this experiment unless he assumed that his water clock was accurate. His experiment was dependent upon the mechanics of his clock and ultimately on his theory of the clock. If we change the theory of the clock and change the calibration of this experiment, then results of the experiment are to a certain degree also changed. Therefore, we can say that the water clock is a theory-laden or assumption-laden piece of hardware which was necessary to his experiment.

There are other things involved in this experiment which were essentially auxiliary pieces of hardware that were also theory-laden or assumption-laden. Galileo stated his inclined plane was straight but how do we define 'straight': in geometry it is the shortest distance between two points. Galileo may have said that (a) he did not bother too much about this definition of "straight" or (b) that the path of a light ray is "straight". Of course physics in the 1700s, following Newton, would later state that light could bend a little in the presence of a gravitational field, but for Galileo, a light ray travelled in a 'straight line' through a homogeneous medium.

Galileo made some basic assumptions about his hardware which were necessary for the construction and conduct of his experiment. Some of his assumptions were theory-laden and others were plucked almost out of thin air. For example, Galileo broadly assumed that although air exercises friction, its role is 'not too important' in this experiment although it affected every trial he conducted. Galileo's inclined plane, sphere and clock constituted a whole cluster of theory-laden hardwares wrapped up with background assumptions and auxiliary theories of the conditions of his experiments [Fig. 3]. Galileo could not step into his laboratory and conduct the test unless he was already committed to certain pre-existing assumptions and theories. Galileo's inclined plane experiments were not acts of nature, because his experimental apparatus was a piece of culture; a piece of human technology whereby you might say his assumptions and theoretical commitments and even his aims were manifested as hardware and as modes of practice.

FIGURE 3



Galileo assumes: 1. Inclined plane straight and smooth;
2. Clock works in linear fashion: flow=time;
3. Rolling ball no problem;
4. Air resistance no problem; and
5. Distance measurement accurate 'enough'.

We are now ready to conduct a test [Fig. 4] Here we have some typical data which shows the type of accuracy that Galileo would have found in his own experiments. We have distances laid out in squares; we have predicted times which assume that everything has been calibrated in unit measures; square the predicted times and you gain a match: d is proportional to the square of time. The first three columns contain the predictions, the remaining columns the data we have obtained during the trials. The 'times' and 'times square' columns contain data similar to that which Galileo achieved. To one unit rolled it is close to one unit of time; for four units rolled it is close to 2 units of time; for nine units rolled it is close to 3 units of time. Some are up and some are down for they tend to be 10 to 15% off the predictions.

FIGURE 4

DATA DO NOT SPEAK FOR THEMSELVES--
EXPERTS STRUGGLE AND NEGOTIATE TO SPEAK THEIR MEANING

GALILEO'S (1638) LAW OF FREELY FALLING BODIES:

Distance fallen from rest is proportional to
the square of the elapsed time of fall:

$D = kT^2$

Here are some "Galileo-like" data -- what do they say?

	Predicted Times:		Test Data:	GAP	
Distance	T	T^2	T	T^2	Diff.
0	0	**0**	0.00	**0.00**	0%
1	1	**1**	.90	**.81**	-19%
4	2	**4**	2.21	**4.88**	+22%
9	3	**9**	2.87	**8.24**	-8%
16	4	**16**	3.93	**15.44**	-4%

Average gap counting all trials equally: 13%

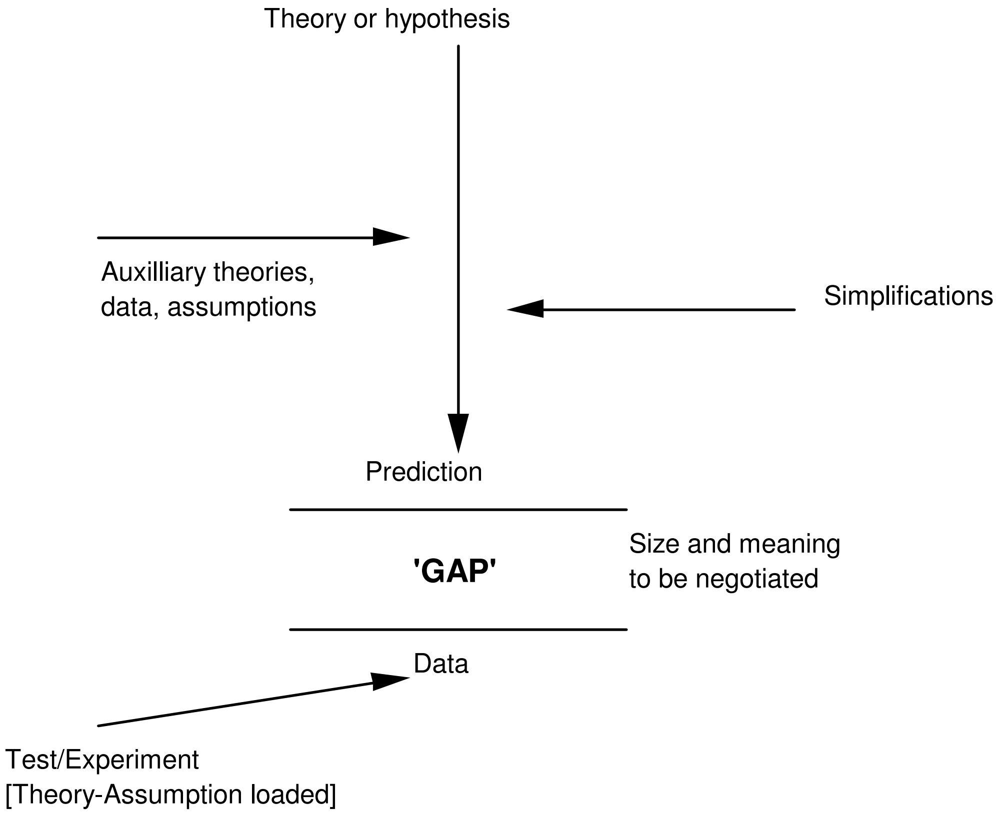
Average gap discounting the 'dodgy' trial at d = 4 :10%

Average gap taking as 'adequate' only the 'good' trials at 'long' distances: 6%

IS GALILEO'S LAW CONFIRMED / FALSIFIED.?
IS THE DATA 'GOOD ENUF' / 'NOT GOOD ENUF'?
WHO HAS THE AUTHORITY / POWER TO DECIDE?
HOW DO THEY DECIDE?
HOW DO THEY ENFORCE/SELL THEIR DECISION?

Galileo stated in his book that he performed the experiment and that the data and the predictions matched up thereby confirming the truth of his hypothesis. Note, however, that there is a discrepancy or gap between his predictions and the data. This is always the case in any experiment. It does not matter what the theory or what super computer crunched out the data. Or what hardware was used, for there is always a gap between your test data and your predictions, they never match perfectly. If they did match we would be living in some sort of mathematical wonderland. [Fig. 5]

FIGURE 5



The next thing to notice is the size of the discrepancy which depends upon further assumptions and interpretations. If we look at Galileo's data then we see that in the first case the data is about 19% off the prediction. It is possible that this was a 'bad trial' and we could then discount this data, which would reduce the gap to 10 or 11% for the overall trials. It is also possible to institute a 'rule' that states the trials that we thought were better should be given more 'weight' than trials which we thought were 'tainted'. We could introduce some complex equations about how to weight the different trials. In the end this amounts to what I call adjusting or negotiating the size of the gap.

Thus, the first thing to remember is that there is always a gap and the second thing is that the size of the gap depends upon decisions and judgments about how to manipulate the data. I am not accusing scientists of universal dishonesty and bias because data is manipulated. I am saying, however, that scientists have always got to make decisions; there may genuinely be a bad trial; or a genuine misfire on the water clock which of course means that as an experimenter you would say that particular trial does not count. If someone suspected you of being a dishonest scientist they would say that you manipulated the data (cooked the data) but I put it to you that every scientist -- in fact anyone who has ever walked into a lab -- knows that you have to do this. There is always a gap between the prediction and the data which depends on data handling techniques and decisions. You may ask who makes these decisions: well, the person performing the experiments and the other experts reading about the experiments in journals for it is a big, social decision-making process.

The third point, which is the most important, is that however big the gap is after making your data handling decisions, **the gap still has no intrinsic meaning**. Galileo states that the gap is 'small enough' for the data to match his predictions. But, using a metaphor, does the gap "speak"; which is really the moral of this story: does the gap speak for itself? Does the data jump from the paper or graph paper and say "we are 'close enough' to the predictions for the predictions to count as confirmed". Obviously not. You can program your computer so that when the gap is below a certain size the computer can say that "the prediction is confirmed because the data is 'close enough' to count". But, you are the person who programmed the computer to allow it to make that "decision" and your colleagues may dispute your decision to program the computer to say what constitutes a "small enough gap".

So who speaks for data? Nature! Divine Illumination! Guess work! No, people speak for data and determine what the data means; but only those people in the relevant scientific community, who have been 'licensed' to argue about the particular kind of data, can speak for it. Those people in the relevant scientific community will argue, wrangle, compete and struggle over what the gap means in this case. This renders the study of experimental science much closer to the study of politics than you could imagine. The results of tests do not speak for themselves, for people always speak their meanings.
The question then becomes, who gets to speak and who exerts the predominant influences in speaking? It was always possible for colleagues or competitors of Galileo to put other 'words' in the mouth of the data and of the gap. It is always possible for other friends or competitors to contest the meaning and significance of the size of the gap.

For example, in the case under discussion, some of Galileo's scientific allies chose to interpret this experiment (the meaning of the gap) differently. Some of his colleagues in Paris repeated his experiments and gained data similar in gap-size to the experiment he had conducted, but they refused to accept the gap as sufficiently narrow to show the data supported the prediction. It is important to understand why his friends would reject the gap as being adequately small. These people in Paris were the closest friends that Galileo had. In 1633 Galileo had been condemned and put under house arrest by the

Catholic Church for teaching that the Copernican sun-centred system of the universe was true. Galileo then completed his work on physics and published his book without using the Copernican hypothesis, without actually stating it. His Parisian friends were also Copernicans who supported Galileo and were horrified that he was placed on trial in 1633. They were on the same scientific side as Galileo and yet they did not agree with the results of his experiments.

As we previously stated, people have to give meanings to the gap and there are only certain basic things they can do when they are involved in a debate about the meaning of a gap between prediction and data in the experiment. You can claim that the available results have a small enough gap to confirm the predictions, or you can say that the gap is too large and the predictions are not confirmed by the data. People who say that the gap is 'small enough' usually have some other agenda; they basically agree with the research and want it to be established as a basis for going on with some other branch of research. This is what Galileo was doing with his inclined plane experiments, stating that his experiment can be used as the basis for research in other areas. Usually when people take the other line, saying the gap is 'too big', they have an agenda which, for whatever reason, aims to stop the line of inquiry involved in the experiment.

Sometimes people will look at the gap and think it is a little too big and that therefore, further research is necessary to make the gap smaller. This further research aimed at 'closing the gap' can focus on the derivation of the prediction: it can be rejigged and other assumptions or auxiliary facts used. Or it can focus on the data--do research on how the experimental apparatus functions. For example, in this case try to improve the theory of the water clock to improve the data so the gap closes. Or perhaps pay more attention to air resistance through a theory of air resistance that would allow an adjustment of the data and so close the gap. The size and meaning of the gap is open to discussion, and further lines of research may be advocated as strategies for attempting to close the gap between the data (experiment) and the prediction (derivation) direction.

Galileo's case needed the gap to be small enough so that his larger project of establishing the Copernican world system would be accepted. Galileo wanted his physics to be true, and the most important way to do this was to get his test results accepted as basic principles. Galileo wanted his physics to work so that it could support his astronomy. His Parisian friends were more cautious in their acknowledged acceptance of the Copernican system. They were not known as Copernicans except in a select circle; therefore, they would not willingly follow Galileo's theories and possibly arouse the wrath of the Church or the scientific community. Hence, they recommended that Galileo do some more work on closing the gap of these experiments. There was perhaps another thing on their agenda: there might be glory in these experiments for themselves, if they can first claim the gap is too big, and then succeed in 'closing' it through their own researches. A lot of science proceeds in this manner. It's as though this is an issue in party politics -- not one party against another -- but in the party room. People are jockeying for positions in their scientific and public careers.

A test and its results are an occasion for negotiation, argument, debate and further research in the relevant research community. It is a debate about the size of the gap, and the manner of narrowing it if it is judged too large. (Or perhaps, if the gap is judged too large, it can be a debate about rejecting the entire experiment, the entire derivation or both.) No test is definitive in and of itself. No test speaks for itself. A test is only definitive if a preponderant part of the relevant community of researchers reaches an agreement about the size and meaning of the gap. A preponderant part does not mean 51%, but it means the people who impose their will on a sufficiently wide domain within the scientific community through a process of argument, struggle, and negotiation.

This view of testing and experiment is fundamental to understanding science as a social and political enterprise. For at the heart of science there always lies this kind of small scale political, community struggle. So when experts spout their results of this or that investigation or test, we are now able to read through all the rhetoric and understand that what are put forward as the test 'results' are themselves not some lessons taught directly by nature or lessons taught directly by some super objective reality, but rather the outcomes of small-scale social processes of negotiation and contestation amongst those very experts. If accepted test results are 'facts'; that is, widely accepted verbal/symbolic reports about natural states of affairs, then those 'facts' are also social constructs, the products of processes of interpretation, judgment, negotiation and conflict in the expert communities licensed to issue such reports.

The traditional method story does nothing to enlighten us about all this--but note that the traditional method story can be used rhetorically by the contending parties in the dispute to beat each other and so help establish one or another interpretation of the size and meaning of the gap. Galileo's colleagues and enemies could have accused him of some lapse of proper method, or vice versa to help explain why their own view of the gap should be accepted. Be that as it may, there is still some life in the method story. It has been revitalised in this century by Sir Karl Popper, in a version that many people believe is a good account of how science works. In the next chapter we look at this last gasp of the method story and see why, based on what we already have learned, it does not and cannot be a workable account of the reality of scientific practice.