How science is done (using astronomy as an example):

In simplest terms, each experiment goes round a loop which can be characterized by six stages:

1. Observe: with an observing or data-gathering programme, record or collect the data.

2. Reduce: clean up the data to remove experimental effects, i.e. (in astronomy) flat field it, calibrate it.

3. Analyse: obtain the numbers from the clean data – intensities, positions. Produce from these summary descriptors of the data which enable comparison or modelling – descriptors that lead to reaching the decision which governed the design of the experiment; and which are statistics.

4. Conclude: carry through a process to reach a decision. Test the hypothesis; correlate; model, etc.

5. Reflect: what has been learnt? Is the decision plausible? Is it unexpected? At which experimental stage must re-entry be made to check? What is required to confirm this unexpected result? Or – what was inadequate in the experimental design? How should the next version be defined? Is an extended or new hypothesis suggested? Far too little time is spent here; perhaps the pressure of observing application deadlines and/or the perceived need to publish get the better of us.

6. Experiment design: if the hypothesis is important enough; if the data warrant it; if previous experimental experience suggests it is possible; if technical advances make it feasible – then the next experiment needs to be designed. This may (and usually does) take the form of thinking out an observing proposal, writing and submitting it. It may take the form of re-design of an instrument on a current telescope. It may take the form of a proposal to build a new instrument. It may take the form of designing a new telescope or space mission, a process which, in itself, may occupy much of a research career. The latest such projects involve multi-nation collaborations on scales of billions of dollars. The timescales from initial plans to realization may range to 40 years (e.g. the James Webb Space telescope; the Square Kilometre Array).

And so back to stage 1.

This process is a loop and ‘experiments’ may begin at different points. For instance, we disbelieve someone else’s conclusions based on their published data set. We enter at point (3) or even (4); and we may then go around the data-gathering cycle ourselves as a result. Or we enter at (5), looking at an old result in the light of new and complementary ones from other fields – and proceed to (6) and back to (1) ….

Of course it could be argued that (6) should start the process, but we need some knowledge base before we start designing.

All too often we use (3) to set up the tests at (4). This carries the charge of mingling hypothesis and data.