Exploration and Confirmation: An Historical Perspective

Michael S. Mahoney

1 Looking to History

In organizing this session, Dieter Rombach invited me to provide historical perspective on a debate that he anticipated but that did not materialize. Expected to take strong stands on opposite sides of a putative disjunction between exploratory and confirmatory experimentation, the papers by Vic Basili and Barbara Kitchenham instead agreed that, given the nature of the subject and the current state of the field, both are necessary. Before one has something to confirm, one must explore, seeking patterns that might lend themselves to useful measurement. The question was not whether experiments of one sort or the other should be carried out, but rather how experiments of both sorts might be better designed. Nonetheless, some historical perspective might still be useful, not least because discussions of experimentation usually involve assumptions about its historical origins and development. It is part of the foundation myth of modern science that experiment lies at the heart of the "scientific method" created (or, for those of Platonist leaning, discovered) in the seventeenth century. That "method" often serves as a touchstone for efforts to make a scientific discipline of an enterprise and thus forms the basis for much common wisdom about where and how experiment fits into the process.

However, over the past several decades historians and sociologists of science have subjected scientific experiments, both historical and current, to critical study, revealing the complexity and uncertainty that have attended experimentation from the beginning [1]. What tradition has portrayed as straightforward applications of empirical common sense to a readily accessible phenomenal world turn out on closer examination to have involved a delicate interplay of experience, theory, and inspired guesswork, as well as subtle negotiation between the experimenters and the audience they were seeking to persuade. When first carried out and reported, experiments that we now assign to students as canonical examples of scientific method turn out to have been ambiguous in their results and difficult to replicate, and to have provoked disagreement rather than settling it [2].

Moreover, experiment turned out not to have played the role previously attributed to it in the formative period. In "Mathematical vs. Experimental Traditions in the Development of Physical Science" Thomas S. Kuhn sought to break down the monolithic image conveyed by the notion of the Scientific Revolution in the 16th and 17th centuries, a phenomenon also referred to as "the origins of modern science" [3]. The felt need to encompass all the sciences had led historians of science to extend the notion of "revolution" to areas where it was not appropriate, to expand the period well into the 18th century to cover a "delayed revolution" in chemistry, and – perhaps most pernicious – to impose 19th-century categories on a range of empirical and experimental endeavors concerning heat and electricity. Kuhn argued for a more modest view, restricting revolution to a small group of sciences, namely astronomy,

V. Basili et al. (Eds.): Empirical Software Engineering Issues, LNCS 4336, pp. 43 – 49, 2007.

mechanics, and optics, all of which had mathematical traditions reaching back to classical antiquity. For our purposes, it is only in this area that one can find experiments that might be called confirmatory.

2 Discovering What to Measure

Confirmation presupposes that one knows what to measure, how to measure it, and what the measurements mean [4]. Until one knows enough to isolate a phenomenon by means of a hypothesis that specifies its determinative parameters, one has no choice but to explore its behavior more generally. That exploration may take the form of systematic experimentation, but in many cases it emerges from experience. Experiments in mechanics in the early 17th century provide a good example. Classic among them is Galileo's inclined-plane experiment to confirm the law of falling bodies, to wit, that the distance traversed from rest under natural acceleration varies as the square of the time of fall. However, that experiment rested on the premise that a ball rolling down an inclined plane models the essential quantitative behavior of vertical free fall. To establish that principle, Galileo described a series of experiments with pendulums, showing that, friction and air resistance aside, a bob dropped from a initial height rises again to that same height irrespective of a change of trajectory caused by shortening the length of pendulum by means of a nail on the center line. Behind those experiments lay earlier studies of the bent-arm balance and its relation to the pendulum and the inclined plane. Hence, behind the experiment confirming a quantitative law, and underpinning it, lay a series of exploratory experiments, through which Galileo, guided by considerations from philosophical sources, teased out the parameters of natural acceleration and the means of displaying and measuring their interaction.

In Galileo's case, however, the question is how the study of motion became experimental at all. It had not been up to that time. Galileo presented his results as the second of "two new sciences" (the first was the strength of materials), noting at the start of his exposition that:

We set forth a very new science concerning a very old subject. Perhaps nothing in nature is more ancient than Motion, and volumes neither few nor small have been written by Philosophers about it. Nevertheless, I have discovered several essential properties that are worth knowing but that hitherto have been neither observed nor demonstrated. Some more obvious things have been noted, for example that the natural motion of falling bodies is continually accelerated. But according to what proportion its acceleration occurs has so far not been established; no one, as far as I know, has demonstrated that the distances traversed in equal times by a body falling from rest stand in same relation to one another as do the odd numbers starting from unity. It has been observed that missiles, or projectiles, trace out some sort of curved line, but no one has established that it is a parabola. That it is, and several other things no less worth knowing, have been demonstrated by me, and, what is more important, they open the way to a most broad and excellent science, for which these our labors will be the starting point from which minds sharper than mine will penetrate into the deepest recesses [5].

Although Galileo published his Two New Sciences in 1638, he was reporting on work carried out during the period 1592-1609, when he taught mathematics and related subjects at the University of Padua and acted as engineering consultant to the Arsenal of Venice. Writing the treatise in the form of a dialogue, he set the action in the Arsenal, where, in his opening words, “everyday practice provides speculative minds with a large field for philosophizing.” His remark points to a larger background against which the emergence of the new subject and its mathematical and experimental methods must be viewed.

As Galileo proposed, his new science of motion was the starting point for the development of mechanics as the mathematical theory of abstract machines, which culminated in Newton’s definition of the subject as “the science of motions resulting from any forces whatsoever, and of the forces required to produce any motions, accurately proposed and demonstrated”[6]. Quite apart from the conceptual issues involved, the very combination of machines, motion, and mathematics brought together subjects that had been traditionally pursued in quite separate realms. Machines were the business of artisans and engineers. Theories of motion were the business of natural philosophers. From Antiquity through the Renaissance, the two groups had nothing to do with one another. Before machines could become the subject of mathematical theory, they had to come to the attention of the philosophers, that is, they had to become part of the philosophical agenda. That process began with the engineers, who during the Renaissance increasingly aspired to learned status, which meant putting their practice on some sort of theoretical basis, that is, expressing what they knew how to do in the form of general principles. In the new literature on machines that appeared over the course of the 16th century, one can see such principles emerging in the form of what I have called elsewhere “maxims of engineering experience” [7]. Though expressed in various ways, they come down to such things as:

```
You can't build a perpetual motion machine.
You can't get more out of a machine than you put into
it; what you gain in force, you give up in distance.
What holds an object at rest is just about enough to
get it moving.
Things, whether solid or liquid, don't go uphill by
themselves.
When you press on water or some other liquid, it pushes
out equally in all directions.
```

Over the course of the 17th century in the work of Galileo, Descartes, Huygens, Newton, and others, these maxims acquired mathematical form, not just as equations expressing laws but also as structures of analytical relations. At the hands of Torricelli, for example, the fourth maxim took the form that two heavy bodies joined together cannot move on their own unless their common center of gravity descends. Combining that principle with Galileo’s work on the pendulum and on natural acceleration, Huygens reformulated it as the principle that the center of gravity of a system of bodies will rise to the same height from which it has fallen, irrespective of the individual paths of the bodies. Expressed mathematically, the principle is an early form of the conservation of kinetic energy expressed as mv^2.

The laws of classical mechanics have been tested and retested experimentally since the 17th century, and every student in high school physics is invited to confirm them in a carefully designed laboratory exercise. We do not invite them to look beyond the exercise to appreciate the centuries of exploratory experience in machine building and the construction of buildings and waterworks from which the laws of motion took their start. Nor do we have them read the ancient and medieval philosophical literature that wrestled with the nature and measure of motion. Before Galileo ever started rolling balls down an inclined plane and measuring the times and distances traversed, he pretty much knew what he would find. Indeed, the experiment doesn't make sense except in terms of the expected results [8].

Galileo's own efforts in other areas make the point clear. The laws of natural acceleration constituted the second of two new sciences. The first was the strength of materials, and there Galileo had less accumulated experience on which to draw. His experiments in this realm were much more of a cut-and-try variety, as he explored possible various ways to test an idea that the cohesion of bodies might have something to with nature's resistance of the vacuum. His successors in the Academia del Cimento picked up where he left off, pursuing a range of experimental inquiries without much sense of where they might lead.

3 Discovering What Happens

In drawing out the contrast between the two traditions, Kuhn pointed to the quite different treatments of magnetism and electricity in William Gilbert's De Magnete (On the Magnet, 1600), considered one of the classics of early experimental science. In experimenting on the magnet, Gilbert drew on a large inventory of empirical data provided by earlier experimenters and in particular by mariners, whose experience of the variation and declination of the compass suggested systematic lines of inquiry. By contrast, when Gilbert turned to electrostatically charged bodies, he had no similar body of experience on which to draw. Hence, his experiments were based for the most part on analogies to the properties of magnets. In general, Kuhn noted,

> When [Baconian] practitioners, men like Gilbert, Boyle, and Hooke, performed experiments, they seldom aimed to demonstrate what was already known or to determine a detail required for the extension of existing theory. Rather they wished to see how nature would behave under previously unobserved, often previously nonexistent, circumstances. (43)

17^{th} century experiments on and with the vacuum quite nicely reveal the different patterns. The experiments began with an effort to account for the observed phenomenon that even the best suction pumps could not raise water more than about 30 feet. At first, it was generally attributed to nature's avoidance of a vacuum, and Galileo even proposed an experiment to measure the force of the vacuum. However, by the 1640s people began to understand the phenomenon in terms of the balance between the weight of the column of water and the force of the air pressing on the surface of the water on which it was standing. Such an explanation lent itself to experiments varying the density of the fluid and the pressure of the air, gradually confirming the hypothesis. The result was the barometer, both as a scientific phenomenon and as a scientific instrument for measuring air pressure.

The experiments involved a closed glass tube, filled with the liquid, and then inverted and placed in an open basin. The fluid would flow out into the basin until the weight of the column and the weight of the air reached equilibrium. Left behind was an empty space at the top of the tube and the question of what, if anything, it contained. Clearly, light passed through it. How about sound? Could an insect fly in it? What would happen to a small animal placed in it? How about a plant? How about chemical reactions? And so on.

The barometer, whether water or mercury, did not lend itself to experiments on these questions, so several people, in particular Otto von Guericke and Robert Hooke (working for Robert Boyle), devised pumps for evacuating the air from a glass receiver or bell jar, making it possible not only to create something close to a vacuum but to vary the air pressure (as measured now by a barometer in the receiver) with some degree of control. For much of the later 17th century, the vacuum became a major site of experimentation, almost all of which was what we would call "exploratory". People placed things in a vacuum to see what happened, and they recorded their observations. But what the observations meant, how they fit with other observations, and how they might be explained, remained open questions. Place a mouse in the vacuum and reduce the pressure. The mouse dies. Why? Air pressure? Something in the air? Something in the mouse? What? How do we find out? Trace the course of any of the experiments in vacuum, and a century or more will pass before empirical exploration gradually gives way to experimental confirmation.

A major hurdle facing the use of experiment to confirm hypotheses was the nature of the hypotheses themselves. The new science of the 17th century rested on the premise that the physical world consisted ultimately of small particles of matter moving according to laws of motion expressible as mathematical relationships. Newton summed it up in Query 31 added to the second edition of his Optics in 1710:

> Have not the small Particles of Bodies certain Powers, Virtues, or Forces, by which they act at a distance, not only upon the Rays of Light for reflecting, refracting, and inflecting them, but also upon one another for producing a great Part of the Phenomena of Nature? For it's well known, that Bodies act one upon another by the Attractions of Gravity, Magnetism, and Electricity; and these Instances show the Tenor and Course of Nature, and make it not improbable but that there may be more attractive Powers than these. For Nature is very consonant and conformable to her Self. How these attractions may be performed, I do not here consider. What I call Attraction may be performed by impulse, or by some other means unknown to me. I use that Word here to signify only in general any Force by which Bodies tend towards one another, whatsoever be the Cause. For we must learn from the Phænomena of Nature what Bodies attract one another, and what are the Laws and Properties of the Attraction, before we enquire the Cause by which the Attraction is performed.

What had worked so well in accounting for the system of the planets, uniting their motion and the motion of bodies on earth under the same laws, should now work at the submicroscopic level. And it should work by experimental means:

> There are therefore Agents in Nature able to make the Particles of Bodies stick together by very strong Attractions. And it is the Business of experimental philosophy to find them out.

As promising as that agenda looks in retrospect, it posed a daunting challenge to researchers at the time, not only to come up with explanations of the requisite mechanical sort but also to devise experiments that confirmed them. Nature proved very hard to grasp at the submicroscopic level, where our senses cannot reach without instruments that depend in turn on an understanding of the phenomena being measured. For the 18th and much of the 19th century, experiments aimed at discovering regularities at the macromechanical level that might give clues to behavior at the micromechanical level, while often purporting to demonstrate the working of hypothetical entities such as the “subtle fluids” that explained the behavior of light (ether), electricity (electrical fluid, later ether), heat (caloric), and chemistry (phlogiston). In short, a lot of exploratory experimentation intervened between Newton’s “atomic” chemistry and Dalton’s atomic theory of the elements, and much of the science of the 19th century would be directed toward devising models, experiments, and instruments that tied the directly observable world to an underlying reality of matter in motion. Increasingly precise and decisive in confirming mathematical theory, none of that work would have been possible without the exploratory experiments that preceded it. It was a slow process at first, and there is no reason to think that experimental software engineering should follow a different path of development.

Notes and References

1. See, for example, H.M. Collins, Changing Order: Replication and Induction in Scientific Practice (London, 1985), A. Pickering, Constructing Quarks : A Sociological History of Particle Physics (Chicago, 1984), S. Shapin and S. Schaffer, Leviathan and the Airpump: Hobbes, Boyle, and the experimental life (Princeton, 1984), P. Galison, How Experiments End (Chicago, 1987).
2. Owing in part to his casual description, Newton’s famous experiments with the prism, in particular the “crucial experiment” showing that monochromatic bands could not be broken down further, did not produce the same results in other people’s hands. We tend to overlook how much work goes into preparing student laboratory exercises to make sure that the experiments turn out as we wish them to do in the hands of neophytes. Note too that, when the experiments do not work well, students are invited to explain what they did wrong, not to question the desired result.
3. Reprinted in Thomas S. Kuhn, The Essential Tension: Selected Studies in Scientific Tradition and Change (Chicago, 1977), 31-65.
4. I am here using “measurement” in a broad sense of objective criteria which include qualitative assays, as in the case of chemical reagents and products.

5. Galileo Galilei, Discorsi e dimostrazioni matematiche intorno à due nuove scienze (Leiden, 1638)150, my translation.
6. Isaac Newton, Mathematical Principles of Natural Philosophy, [London, 1687], trans. A. Motte (London, 1728), Preface.
7. M.S. Mahoney, "The Mathematical Realm of Nature", in D.E. Garber et al.(eds.), Cambridge History of Seventeenth-Century Philosophy (Cambridge: Cambridge University Press, 1998), Vol. I, pp. 702-55
8. Several of the experiments for which Galileo is most widely known, such as the dropping of objects from the Tower of Pisa, were never carried out. In most cases, they are aimed at refutation of purported experience, and in some cases they are part of a thought experiment aimed at bringing out logical inconsistency rather than failure to match experience.