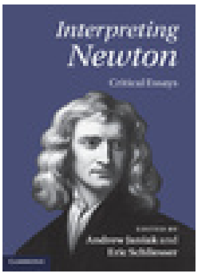

Interpreting Newton

Critical Essays

Edited by Andrew Janiak, Eric Schliesser

Book DOI: http://dx.doi.org/10.1017/CBO9780511994845

Online ISBN: 9780511994845

Hardback ISBN: 9780521766180

Paperback ISBN: 9781107624870

Chapter

15 - How Newton's Principia changed physics pp. 360-395

Chapter DOI: http://dx.doi.org/10.1017/CBO9780511994845.019

Cambridge University Press

15

How Newton's *Principia* changed physics

GEORGE E. SMITH

Newton expressly intended his *Principia* to produce three revolutionary changes in the way physics and astronomy were being conducted:

1. Theorizing in physics should center on identifying fundamental *forces* of nature and characterizing them as quantities related by laws to other measurable quantities.
2. Astronomy should abandon the 1500 year tradition of trying to describe complex orbital motions directly from observations and instead derive them from the forces acting on the orbiting bodies.
3. Physics and astronomy should demand of themselves a much higher standard of evidence in theorizing than just success in deriving observed phenomena from speculative hypotheses.

The *Principia* did indeed ultimately effect all three of these revolutions – the first two obvious to anyone familiar with the subsequent history of physics and orbital astronomy, but the third less obvious. Here accordingly we shall focus on the third, though explaining how the book changed the standards of evidence in physics will involve us with the first two as well. While those two emerged in Newton's thinking only with the *Principia*, the third he had set as a goal more than a decade and a half earlier with the remark, "But truly with the help of philosophical geometers and geometrical philosophers, instead of conjectures and probabilities that are being blazoned everywhere, we shall finally achieve a natural science supported by the greatest evidence" (Newton 1984, p. 87).

One reason why the revolution in evidence is less obvious has been a long-standing, but nonetheless ill-informed misconstrual of the evidential reasoning not only in the *Principia*, but in subsequent research in orbital mechanics as well. The first two sections of the chapter will contrast that construal of the reasoning with the evidence problem Newton saw himself as facing when he started writing the *Principia*. Another reason for the revolution in evidence being less obvious has been the complexity of the *Principia*'s approach to marshalling evidence, involving as it does several distinct elements that are usually discussed, when at all, in isolation from one another. Sections 15.3 through 15.6, forming the main body of the chapter, will lay out those revolutionary elements one by one, indicating how each reflects this evidence problem. A final section will then consider the whole formed by those elements, asking how clearly Newton saw it as a response to the principal worries he had about potential shortcomings in his evidence.

15.1 Introduction: the question

Prima facie, the evidence put forward on issues about orbital and other kinds of motion at the time of Laplace's *Celestial Mechanics* (1799–1805) was of much higher quality than the evidence on those issues at the time of Copernicus or, for that matter, Kepler. Furthermore, of the several advances that were made between Copernicus and Laplace that enabled more decisive evidence to be developed, nothing appears to have been more important than Newton's *Principia*. The central question of this chapter is, How did Newton's *Principia* change the way in which evidence was marshalled in orbital research, and thereby in physics generally?

Prompting that question is a view that empirical science is first and foremost a process of *turning data into evidence*. Evidence is a two-place relation between data and claims that reach beyond them.[1] Data, in and of themselves, are not evidence for one claim more than another; something beyond data is always needed for them to become evidence for anything. In experimental research novel data are often generated, sometimes with masterful artifice, precisely because their likely value as evidence is clear beforehand. Often, however, data are abundantly available in nature, and the problem is one of figuring out what they show about the world. Linguistics provides a clear example of this, for data on the syntax of our native languages are immediately at hand, but we still do not have a fully adequate account of the syntax of any natural language.[2]

In orbital astronomy, too, data have always been accessible in the form of nightly observations of relative positions of objects on the celestial sphere, and efforts to turn those data into evidence go back at least as far as the Babylonians. The introduction of the telescope at the beginning of the seventeenth century provided access to new data, but almost all of the evidence bearing on orbital astronomy until the middle of the eighteenth century came from instrument-aided naked-eye observation. New ways of turning those data into evidence concerning celestial physics emerged between Copernicus and Laplace. Those

1 In deference to Charles Saunders Peirce, who surely would have insisted that evidence involves a third place as well as the two cited, perhaps I should say "two- (or more) place relation."

2 In conversation a few years ago Noam Chomsky and I were unable to figure out which of us first began speaking of science as an endeavor to turn data into evidence, followed immediately by the remark that evidence is a relation and being a datum is not. Regardless of who did, the thought was originally no less his than mine.

new ways are the central concern of this chapter. How precisely did Newton's *Principia* contribute to them?

Seen from that perspective, the most frequently quoted portion of the Preface to the first edition of the *Principia* indicates that Newton saw it as illustrating a new way of turning data into evidence:

> our present work sets forth mathematical principles of natural philosophy. For the whole difficulty of philosophy seems to be to discover the forces of nature from the phenomena of motions and then to demonstrate the other phenomena from these forces. It is to these ends that the general propositions in books 1 and 2 are directed, while in book 3 our explanation of the system of the world illustrates these propositions. For in book 3, by means of propositions demonstrated mathematically in books 1 and 2, we derive from celestial phenomena the gravitational forces by which bodies tend toward the sun and toward the individual planets. Then the motions of the planets, the comets, the moon, and the sea are deduced from these forces by propositions that are also mathematical. If only we could derive the other phenomena of nature from mechanical principles by the same kind of reasoning! For many things lead me to have a suspicion that all phenomena may depend on certain forces by which the particles of bodies, by causes not yet known, either are impelled toward one another and cohere in regular figures, or are repelled from one another and recede. Since these forces are unknown, philosophers have hitherto made trial of nature in vain. But I hope the principles set down here will shed some light on either this mode of philosophizing or some truer one.
>
> (Newton 1999, p. 382)

Newton's approach to turning data from astronomical observations into evidence about forces governing orbital motions and then about those motions themselves was more multi-faceted than is generally appreciated. This chapter aims to lay out his approach and the rationale behind it and then to indicate ways it altered orbital astronomy and physics generally.

There is a commonplace answer to the question of how Newton's *Principia* resulted in exceptionally high quality evidence, an answer that can be extracted from undergraduate textbooks in physics, if not explicitly found in them: What Newton did in the *Principia* was to put forward the law of gravity, together with his three laws of motion, by way of explaining Kepler's so-called laws; and the resulting theory then turned out to explain ever so much more, including the respects in which actual planetary motions deviate from Kepler's laws. In other words, until a small residual discrepancy in the precession of the perihelion of Mercury emerged in the second half of the nineteenth century, Newton's theory turned out to be consistent with all observations, and in that sense passed every test to which it was put. On this view, the high quality of the evidence coming out of the *Principia* lay in the *range* of observations with which the laws it proposed turned out to be in agreement and the *precision* of that agreement.

That view of the *Principia* offers a conception of the enterprise of science in sharp contrast with my "process of turning data into evidence." Science is instead first and foremost a process of coming up with basically correct theories.[3] Once such a theory is in hand, the evidence for it will largely just fall into place as tests of it emerge and it survives them. The task of marshalling evidence itself presents no special challenge save for an occasional need for ingenuity in devising new, more telling tests. Granted this is a stick-figure summary. Even in this form, however, it explains why textbooks in science include so little discussion of details of the evidence.

As a preliminary step toward motivating the view of the evidence for Newton's theory to be presented below, let me offer two objections to the view I have just sketched. First, it distorts history. For example, Kepler's rules for calculating orbits were far from established at the time Newton began drafting the *Principia*. Indeed, they appear never to have been called "laws" before the *Principia*.[4] Furthermore, a number of so-called tests of Newton's theory were not expressly offered as tests of it at the time. A blatant example of this is Cavendish's experiment, which in physics textbooks is usually presented as a decisive test of Newton's law of gravity even though Cavendish himself said that what he was doing was to measure the (mean) density of the Earth.[5]

A second objection lies in Newton's own outspoken dismissal of hypothetico-deductive evidence. As quoted above, Newton claimed in the first edition of the *Principia* to have derived the law of gravity from phenomena of orbital motion; and at the end of the second edition he added that "hypotheses, whether metaphysical or physical, or based on occult qualities, or mechanical, have no place in experimental philosophy" (Newton 1999, p. 943). Important to note here is Newton's lifelong reason for dismissing hypothetico-deductive evidence: "For if the possibility of hypotheses is to be the test of the truth and reality of things, I see not how certainty can be obtained in any science; since numerous hypotheses may be devised, which shall seem to overcome new difficulties" (Newton 1978, p. 106). That Newton was responding to such worries alone gives reason for examining whether he had an alternative approach.

This chapter does not aim to argue against deductivist accounts of evidence. It lays out an alternative account of the logic of the evidential reasoning in gravitation research and the way in which this logic derives from the *Principia*.

3 Through most of this chapter *theory* designates bodies of lawlike relations among quantities. Here, however, it designates more a way of conceptualizing a range of phenomena. Sylvain Bromberger has labeled those two senses of "theory" "theory$_1$" and "theory$_2$" in Bromberger (1992).

4 Curtis Wilson (2000, p. 225) has noted that Kepler's orbital rules appear never to have been called "laws" in print before Leibniz did so shortly after publication of the first edition of Newton's *Principia*.

5 Cavendish (1798) simply assumed the law of gravity throughout. Had he been trying to test it, he would have at least varied the masses of the spheres in his trials.

This alternative, I claim, is more accurate historically and more consistent with Newton. Most of all, however, I want to claim that, on its very face, it is a more tenable account of why the *Principia* had the effects it did on how evidence is developed in physics.

15.2 Complexity and parochialism: the evidential problem

At the time Newton began drafting the *Principia* in 1685, there were several competing approaches to calculating planetary orbits, at least seven of which were known to him. Kepler's approach employed the ellipse and his area rule – *planets sweep out equal areas with respect to the Sun in equal times* – but he did not use his 3/2 power rule – *the semi-major axes of the ellipses vary as the 2/3 power of the orbital periods* – to infer the lengths of the semi-major axes directly from the periods. Instead, he inferred these lengths from observations. Jeremiah Horrocks found that he could improve on Kepler's *Rudolphine Tables* by inferring the semi-major axes directly from the periods, which were known to very high precision (Wilson 1978). The other five approaches employed some alternative to Kepler's area rule for determining where each planet is on its orbit at any given time. Ismaël Boulliau (1657, pp. 29–31) used a geometric construction involving the empty focus. Thomas Streete (1661, pp. 53f. and 39f.) used this same geometric construction, but followed Horrocks in inferring the semi-major axes from the periods. Vincent Wing (1651, p. 44ff.) initially used an oscillating equant – that is, a center of equiangular motion oscillating about the empty focus – and later (1669, pp. 130, 144, 151, 170, 176) switched to his own geometric construction. And Nicolaus Mercator (1676, pp. 163–171) used a still different geometric construction.[6]

Of these different approaches, Kepler's was computationally the most complicated. None of them gave predictions that were consistently within the accuracy of pre-telescopic observations (Wilson 1989). The errors were more or less comparable in all seven – around a third of the apparent width of the Moon. The only thing common to all of them was the ellipse, which is striking because the orbits are actually so near to being circular; the most elliptical of the orbits then known, Mercury's, has a minor axis only two percent shorter than its major axis. Equally striking, the ellipse itself was something that Newton did not consider appropriate to use as evidence for his law of gravity (Smith 2002b).

A question accordingly at the forefront of orbital astronomy in 1679 when Hooke first put the matter to Newton, and still in 1684 when Halley did the same, was which of the different ways of calculating orbital trajectories was to be preferred. The brief tract Newton had registered with the Royal Society in December 1684, "De Motum Corporum in Gyrum," gives an answer to

6 Mercator precedes his account of his new hypothesis with an extensive review of Kepler's area rule and alternatives to it.

the question: "The major planets orbit, therefore, in ellipses having a focus at the center of the Sun, and with their *radii* drawn to the Sun describe areas proportional to the times, exactly as Kepler supposed" (Newton 1967–1981, VI, p. 49). A plausible construal of Newton's reasoning here is that Kepler's (and Horrocks's) calculation rules admit of a physical explanation – they result from an inverse-square centripetal force – while the competing rules of calculation appear unlikely to do so. On this construal, the evidential problem to which Newton was offering a solution was one that had been posed by Kepler: given the imprecision of observation and measurement, any number of distinct curves can fit observations to any given level of precision; only physical considerations can pick out the true curve from among these. That, however, is not the evidential problem Newton addressed in the *Principia*, for by the time he began drafting it a few weeks later, he had come to see ways in which orbital motion poses a far more ramified challenge.

Specifically, Newton had come upon a deep reason why none of the ways of calculating the orbits were yielding results within observational accuracy. In the registered version of "De Motu" he had concluded that there are what we would now call inverse-square centripetal acceleration fields not merely around the Sun, but also around the Earth, Jupiter, and Saturn.[7] He saw no reason why the inverse-square accelerative tendency toward, for instance, Jupiter exhibited by its four known satellites would not extend all the way to the Sun, so that Jupiter and the Sun would be interacting with one another. If they do interact, then this interaction should not produce any change in the motion of the center of gravity of the system. (This is just the principle of inertia applied to a system of interacting bodies.) From this, Newton reached an extraordinary conclusion in an augmented version of the "De Motu" tract that did not become public until 1893:

> By reason of the deviation of the Sun from the center of gravity, the centripetal force does not always tend to that immobile center, and hence the planets neither move exactly in ellipses nor revolve twice in the same orbit. There are as many orbits of a planet as it has revolutions, as in the motion of the Moon, and the orbit of any one planet depends on the combined motions of all the planets, not to mention the action of all these on each other. But to consider simultaneously all these causes of motion and to define these motions by exact laws admitting of easy calculation exceeds, if I am not mistaken, the force of any human mind.
>
> (Newton 1962, pp. 256 and 281)[8]

7 While Newton never employed the term "field," my use of it is not so anachronistic as it might at first seem, for he did speak of centripetal motive forces being "propagated through the surrounding regions" (Def. 8).

8 I have altered the translation along lines derived from Curtis Wilson. This passage did not become public until Ball (1893).

In other words, no calculation scheme like Kepler's or any of the others was ever going to yield exact predictions, not for the comparatively uninteresting reason that observation itself is always imprecise, but for the far more important reason that the true motions are too complex to allow exact computation.

Newton was not the first to decry the complexity of true motions in the world. Both Galileo and Descartes had concluded that motion under air resistance forces (the other topic of Newton's *Principia*) is too complex to allow a science (Galileo 1974, 224; Descartes 1985–1991, III, p. 9f). Newton knew that Descartes had said much the same of the motions of the planets, adding that their trajectories are sure to change from one epoch to another:

> Finally, we must not think that all the centers of the Planets are always situated exactly on the same plane, or that the circles they describe are absolutely perfect; let us instead judge that, as we see occurring in all natural things, they are only approximately so, and also that they are continuously changed by the passing of the ages.
>
> (Descartes 1991, p. 98)

This raised a second worry: not only might Keplerian motion be but one of several comparably accurate approximations to the true trajectories, whose complexity defies exact description; but also, Kepler's and all the other approximations might be mere epochal parochialisms, projected from a few decades of observations that were assumed to be representative, but instead were systematically misleading historical accidents.

As noted earlier, Newton saw his *Principia* as illustrating a new way of doing science. I contend that Newton's new "experimental philosophy" – as he came to call it – was in response to the complexity of the real world and the risk that our straightforward empirical access to it is parochial. That is, it is an approach to developing evidence in the face of, first, a complexity that leaves room for many competing descriptions of observed regularities and, second, a lack of any immediate means of obviating respects in which the observed regularities we invoke as evidence might be misleadingly parochial. In forming this new approach, Newton introduced a number of changes in approach that have persisted at least in the subsequent history of gravitational research, if not in physics generally. It is to these changes that we now turn.

15.3 The Newtonian conception of theory

One change came out of Newton's realization that physics – or, more specifically, mechanics – cannot help but include, within its scope, its own theory of measurement. Newton hints at this in the Preface to the first edition of the *Principia* when he concludes that "*geometry* is founded on mechanical practice and is nothing other than that part of *universal mechanics* which reduces the art of measuring to exact propositions and demonstrations" (Newton 1999,

p. 382). The point emerges more forcefully in the opening section of the book, "Definitions." This section might just as well have been called "Critical Reflections on Measurement." The definitions of *quantity of matter* or *mass*, *quantity of motion* (our *momentum*), and *force*, besides indicating how the terms are going to be used in the *Principia*, emphasize their measures. Indeed, the definitions of *quantity of matter* and *quantity of motion* expressly identify each as "a measure." The discussions following their definitions make clear that both *mass* and *force* are what we now call "theoretical quantities." That is, values for them must be inferred from other measurements and the inferences in question presuppose theoretical claims within mechanics. In the case of *mass* Newton even invokes a pendulum experiment to justify inferring values from *weight* (Newton 1999, pp. 404 and 807).

Following the explicit definitions is the famous Scholium on space and time, the central concern of which is the distinction between "absolute, true, or mathematical space, time, and motion" and "relative, apparent, or common space, time, and motion." The space, time, and motion that we observe fall into the relative, apparent, or common category. Values in the absolute, true, or mathematical category have to be inferred from them. In the paragraph ending the Scholium, Newton remarks:

> It is certainly very difficult to find out the true motions of individual bodies and actually to differentiate them from apparent motions, because the parts of that immovable space in which the bodies truly move make no impression on the senses. Nevertheless, the case is not utterly hopeless. For it is possible to draw evidence partly from apparent motions, which are the differences between true motions, and partly from the forces that are the causes and effects of the true motions . . . But in what follows, a fuller explanation will be given of how to determine true motions from their causes, effects, and apparent differences, and conversely, of how to determine from motions, whether true or apparent, their causes and effects. For this was the purpose for which I composed the following treatise.
>
> (Newton 1999, p. 414f)

Viewed from the perspective of the rest of the treatise, the natural way to interpret what Newton is saying here is that true motions are ones for which all theory-mediated measurements of the relevant forces yield the same values. But then, not just values of *mass*, *force*, and *quantity of motion*, are theory-mediated; so too are values of *velocity*.

Newton may not have been the first to realize that physics must include its own theory of measurement. In one respect the point is obvious, for measurement is itself a physical process and measurements in mechanics involve mechanical processes. Still, Newton does appear to have been the first to appreciate two of its implications. One is that any method of measurement is provisional, subject to replacement by a method that is deemed preferable at some later point. Newton expressly calls attention to this in the case of *time*:

> In astronomy, absolute time is distinguished from relative time by the equation of common time. For natural days, which are commonly considered equal for the purpose of measuring time, are actually unequal. Astronomers correct this inequality in order to measure celestial motions on the basis of a truer time. It is possible that there is no uniform motion by which time may have an exact measure. All motions can be accelerated and retarded, but the flow of absolute time cannot be changed. The duration or perseverance of the existence of things is the same, whether their motions are rapid or slow or null; accordingly, duration is rightly distinguished from its sensible measures and is gathered from them by means of an astronomical equation. Moreover, the need for using this equation in determining when phenomena occur is proved by experience with a pendulum clock and also by eclipses of the satellites of Jupiter.
>
> (Newton 1999, p. 410)

Newton's defense of sidereal time by appealing to the pendulum clock and the eclipses of the satellites of Jupiter is striking because both of these presuppose theories that were first published in the 1670s – the latter including a theoretical redetermination of *simultaneity* in astronomy.[9] The evidence for both the regularity of pendulum clocks and the eclipses of Jupiter's satellites in turn invokes sidereal time. In other words, a confluence of theoretical considerations lies behind the choice of sidereal time. But then, if preferred methods of measurement are subject to change as new theoretical considerations emerge, any lawlike relationship between measured quantities must also be provisional, subject to change as science advances.

The second implication of physics having to include its own theory of measurement that Newton appears to have been the first to appreciate is less spectacular, but no less important. It concerns how theory-mediated measurement can enter into evidence. Insofar as all measurement presupposes theoretical considerations of one sort or another, there is no reason to insist that a theory be firmly established first, before new methods of measurement are derived from it. Huygens in 1659 had used his theoretical laws for the cycloidal and conical pendulums to measure the strength of surface gravity in two different ways, obtaining the same value to four significant figures (Yoder 1988). Huygens, however, seems never to have viewed the stability, convergence, and precision of his measurements as evidence for the theory of uniform gravity from which he derived his pendulum laws. Newton saw not only this, but also that in the long run such stability, convergence, and precision of measurement

9 The former presupposed the theory of the pendulum in Huygens's *Horologium Oscillatorium* of 1673, and the latter, Olaus Römer's determination of the finite speed of light in 1676.

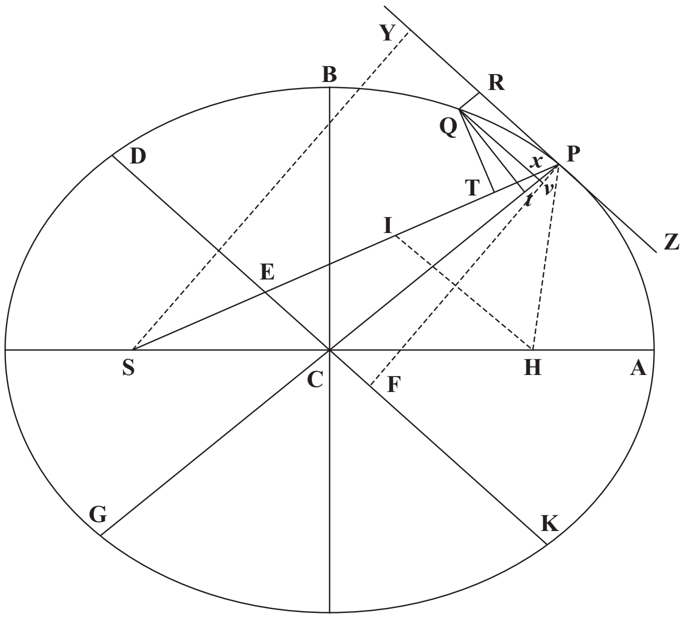

Figure 15.1 The figure accompanying both Propositons 10 and 11 on elliptical orbits in the First Edition

cannot help but be a primary form of evidence for any theory in mechanics. But then, success in theory-mediated measurement should be regarded as evidence for a theory right from the outset, even before any other evidence for it is available.

Freeing quantities like *force* and *time* from any specific way of measuring them allowed them to be considered in the abstract, as mere mathematical quantities separate from any question about physical mechanism. This in turn allowed Newton to introduce a new way of employing mathematical theory in physics. Galileo and Huygens had used mathematics to derive testable consequences from their theories of motion. Newton, by contrast, developed a generic theory of motion under centripetal forces, deriving results not only for inverse-square forces, but also for forces that vary linearly with distance, that vary as the inverse-cube of distance, and finally that vary as any function whatever of distance. Two consecutive propositions of Book I of the *Principia* illustrate this new form of mathematical theory (Figure 15.1):

> *Proposition 10*: Let a body P revolve in an ellipse; it is required to find the law of centripetal force tending toward the center [C] of the ellipse.
>
> *Solution*: The force varies as CP directly.

> *Proposition 11*: Let a body P revolve in an ellipse; it is required to find the law of centripetal force tending toward a focus [S] of the ellipse.
>
> *Solution*: The force varies inversely as the square of SP.

(Keep these two in mind, for I will be making a further point about them in the next section.) In both of these, questions about how such centripetal forces might be physically effected are irrelevant. Book I of the *Principia* consists of more than ninety "if-then" propositions linking motions to forces, forces to motions, and macrophysical forces to microphysical forces composing them; throughout, *force* is treated as a quantity, independent of what brings it about.

Book I is best described as giving a generic mathematical theory of motion under forces directed toward a center *with no regard to how such forces might be physically realized.* Mathematical theory of this sort is a second way in which Newton changed physics.

Late in Book I Newton indicates why he wants a generic mathematical theory in which forces are treated without regard to the question of the physical mechanisms producing them:

> Mathematics requires an investigation of those quantities of forces and their proportions that follow from any conditions that may be supposed. Then, coming down to physics, these proportions must be compared with the phenomena, so that it may be found out which conditions of forces apply to each kind of attracting bodies. And then, finally, it will be possible to argue more securely concerning the physical species, physical causes, and physical proportions of these forces.
>
> (Newton 1999, p. 588)

The passage brings out two points. First, generic mathematical theory for Newton is an instrument for turning data into evidence, more specifically for enabling phenomena to answer theoretical questions about physical forces and processes. The idea is to have generic mathematical theory and phenomena together dictate physical theory. Second, Newton is prepared to leave questions within physical theory open when he can't find phenomena to answer them. In particular, he answered questions about the physical species of celestial orbital forces – they are one in kind with terrestrial gravity – and questions about their physical proportions – the law of gravity; but he found no way of addressing their physical causes. His theory of gravity was rejected by many of the leading figures of his time precisely because it left the question of the cause of gravity open, and therefore offered no explanation of how gravitational forces can act over vast distances. A third way in which Newton's *Principia* ended up changing physics was by making limited physical theory – theory without mechanistic explanations – respectable.

The phrase, *limited physical theory*, describes a methodological innovation. I would be remiss not to point out how revolutionary that innovation also was from a substantive standpoint. Newton was the first to propose that physical theory focus on fundamental kinds of force. In doing so he introduced an intermediate level of theory, between mere description of observed regularities in the manner of Galileo's *Two New Sciences*, on the one hand, and laying out full mechanisms in the manner of Descartes' *Principia*, on the other. Newton's *Principia* showed that this intermediate level of theory, with laws of force but no mechanism, is still sufficient to answer a whole host of questions about observed regularities – especially questions about whether observed regularities, as described, are suitable for playing a role in evidence. This, I take to be the point Newton was making in his remark about his predecessors having "hitherto made trial of nature in vain" in the passage from the Preface to the first edition I quoted in Section 15.1. It is to this intermediate level of theory that we now must turn.

15.4 Evidential reasoning in Newton's *Principia*

The features of Newton's approach identified so far give Newtonian theory what Pierre Duhem called an "abstract, symbolic" character. But as Duhem himself showed, that character need not be peculiar to generic theories. To appreciate the advantage Newton found in insisting on a generic theory of motion under centripetal forces, we need to look in detail at how he reasoned from orbital phenomena to physical proportions of force.

Newton inferred that the force acting on the planets is centripetal from Kepler's area rule, and he inferred that it is inverse-square first from Kepler's 3/2 power rule, and then more strictly from the absence of precession of the orbits. The following three theorems from his generic mathematical theory are the "inference-tickets" licensing those inferences:

> *From Propositions 1–3*: A body sweeps out equal areas in equal times with respect to a second body if and only if the net force on the body is compounded of a centripetal force directed toward the second body and the whole accelerative force acting on the second body.
>
> (Newton 1999, pp. 444–448)
>
> *From Corollaries to Proposition 4*: The periodic times of bodies moving uniformly in circular orbits about a central body vary as the 3/2 power of their distance from the central body if and only if the centripetal forces acting on the orbiting bodies vary inversely as the square of the distances.
>
> (Newton 1999, p. 451)
>
> *From Proposition 45 and its Examples*: The centripetal force acting on a body moving in a nearly circular orbit is inverse-square if and only if the perihelion (or perigee) of its trajectory does not precess.
>
> (Newton 1999, pp. 539–545)

We will discuss below why it was important to Newton to establish not merely the conditionals licensing his inferences from phenomena, but the bi-conditionals as well.

Each of these enabling theorems is richer than it first appears to be. In corollaries to the first Newton adds that an increasing areal velocity with respect to a point entails that the net force is directed forward of that point, and vice versa. In a corollary added in the second edition to the second, Newton points out that the period varies as the distance to the power n (where n need not be an integer) if and only if the centripetal force varies inversely as distance to the power $2n - 1$. And his precession theorem actually gives an algebraic formula tying the rate of precession of a nearly circular orbit to the exponent in the force rule:

> Let θ be the angle at the force center from aphelion to perihelion in a very nearly circular orbit; then the centripetal force varies as $R^{(n-3)}$, where $n = (180/\theta)^2$.

In other words, the enabling theorems show that a real acceleration is a theory-mediated measure of the direction of the force on an orbiting body; the exponent in the power rule relating periods to distances of a collection of bodies moving uniformly in circular orbits is a theory-mediated measure of what we now call the strength of the acceleration field around the central body; and the rate of orbital precession is a theory-mediated measure of the exponent of distance from the center in the centripetal force rule for any one orbiting body.

Still more important, if either clause in any of the three enabling bi-conditionals holds only approximately – Newton's phrase is *quam proxime*, very nearly – then the other clause still holds *quam proxime*. This follows trivially from the algebraic relations in the second and third cases, and Newton expressly points it out in a corollary to the first:

> *From Proposition 3, Corollaries 2 and 3*: The areas with respect to the central body are as the times *quam proxime* if and only if the force retaining the moving body in an orbit around the central body tends toward the central body *quam proxime*.

Thus, for every "if-then" statement that Newton uses to reason from orbital phenomena to conclusions about forces, he takes the trouble to show that the consequent still holds *quam proxime* so long as the antecedent holds *quam proxime*! In effect, then, the logical form of the propositions that serve to license Newton's reasoning from orbital phenomena is not really "if-then", but rather "if *quam proxime*, then *quam proxime*," and hence the premises describing the orbital phenomena in question are required to hold only *quam proxime*. Consequently, Newton is not begging any questions about whether the area rule or some other rule is the proper one for locating planets in their orbits as a function of time, for he knew that the area rule agrees with all the other rules at least *quam proxime*.[10]

That Newton was consciously engaged in such a form of approximative reasoning explains why he did not use Proposition 11, given above, to infer the inverse-square proportion from the Keplerian ellipse. For, he had shown that, when the eccentricity is small, as it is in the case of several of the planets, the contrast between Propositions 11 and 10 – that is, between the proportion implied by ellipses with equal areas about a focus and about the center – becomes a problem. It is demonstrably not true that, if the trajectory is a Keplerian ellipse *quam proxime*, then the exponent in the force rule is -2 *quam proxime*. Current textbooks typically do present Newton as reasoning from the Keplerian ellipse to the inverse-square proportion, and this inference is certainly inviting on its face. Newton, however, was more careful than these textbooks. Nowhere, not even in the original "De Motu" tract, did he make this move. Instead he always relied on the 3/2 power rule and the absence of orbital precession to infer the inverse-square (Smith 2002b).

Newton's not inferring the inverse-square from the Keplerian ellipse, together with his taking the trouble to show that the "if-then" statements he did employ hold in *quam proxime* form, provides the strongest evidence that he was self-consciously engaged in approximative reasoning. There are several other signs of it as well. The phrase, "*quam proxime*," occurs 139 times in the *Principia*. The numerical summaries of the observed relations between periods and mean distances at the beginning of Book III all display some deviation from an exact 3/2 power relation, and Newton openly acknowledges that the Moon is not in perfect accord with the area rule and that its orbit is not stationary. Moreover, Newton had decided before he began writing the *Principia* that the area rule does not hold exactly for the planets, and he had concluded while writing the *Principia*, if not before, that their orbits are not perfectly stationary. Granting that he was engaged in approximative reasoning is thus a way of absolving him of accusations of rank hypocrisy (Lakatos 1978). Finally, it undercuts the complaint made by Duhem and others that the law of gravity cannot be deduced from Keplerian phenomena, taken as premises, because it entails these premises are false: the seeming self-contradictory element of Newton's "deduction" disappears once the reasoning is construed as approximative.

Of course, what this means is that, strictly speaking, the evidence Newton offers for his law of gravity shows that it is true of the motions of the planets and their satellites, but only *quam proxime*, only to high approximation. Newton is perfectly aware that the orbital evidence does not show that the law holds exactly. Nevertheless, he *takes* the law to hold exactly – or, what in practice amounts to the same thing, to hold unqualifiedly within the limits of

10 See Mercator (1676).

observational accuracy. In the third edition of the *Principia*, he gives a rule of reasoning to authorize this leap from approximate to exact:

> Rule 4: *In experimental philosophy, propositions gathered from phenomena by induction should be taken to be exactly or very nearly true notwithstanding any contrary hypotheses, until yet other phenomena make such propositions either more exact or liable to exceptions.*
>
> (Newton 1999, p. 796)[11]

Note the phrase here, *should be taken to be*. The leap from approximate to exact amounts to a research strategy.[12]

This is a fourth way in which Newton's *Principia* changed physics. On the one hand, because the scope and precision of observation are limited and the real world is complex, evidence in physics can at most show that theoretical claims hold to certain levels of precision over a limited range of observations. On the other hand, when appropriate requirements are met – as expressed in the phrase, *gathered from phenomena* – physicists should nevertheless proceed as if these theoretical claims hold exactly. What we need to do now is to see how this research strategy works.

Newton appears to have required that two conditions be met before he was willing to take the law of gravity as exact. He expressly states in Proposition 8 of Book III that he did not conclude that the inverse-square proportion holds exactly until he had established that it holds exactly around a sphere of uniform (or spherically symmetric) density if it holds with respect to all the particles of matter forming that sphere (Newton 1999, p. 811). In other words, he required there to be *some* configuration for which the macroscopic forces around a body, composed out of forces toward its parts, would accord exactly with the law. Second, he also appears to have required that there be *some* identifiable circumstances in which the phenomena from which the law was inferred *would* hold exactly. As the following quote indicates, the subjunctive here is Newton's, not mine:

> if the Sun were at rest and the remaining planets did not act upon one another, their orbits would be elliptical, having the Sun in their common focus, and they would describe areas proportional to the times.
>
> (Newton 1999, p. 817f.)

11 Newton adds by way of explanation for why the rule is needed, "This rule should be followed so that arguments based on induction may not be nullified by hypotheses."

12 Similarly, the inference that Newton can strictly speaking draw from orbital phenomena has to be restricted as holding only over the period of time for which observations were available, primarily from Tycho Brahe forward. Newton's third rule of reasoning licenses the inference *to be taken* to hold universally into the past and future, just as his fourth rule licenses the inference *to be taken* to hold exactly, or at least to high approximation. For reasons of space I have chosen not to go into his third rule and its strategic role in ongoing research here.

Newton's theory of gravity did not show that the planets revolve in stationary Keplerian orbits, but instead that they decidedly do not. Nevertheless, Kepler's area and 3/2 power rules, and also the absence of orbital precession, are special in one crucial respect. According to the theory, each of these phenomena *would* hold exactly were it not for small gravitational forces directed toward the individual planets. (The deductions of these phenomena are where the converse parts of the enabling bi-conditionals come into play, though here in their exact, not their *quam proxime* form.) Because the actual motions of the planets are exceedingly complex, they can be approximated to any given level of accuracy in an indefinite number of ways. Newton is requiring the approximations from which physical theory is "deduced" to be ones, that according to the theory, *would* hold exactly in specific identifiable conditions. This is a key element of Newton's way of marshalling evidence in the face of complexity.

What Newton has done here is to single out a particular kind of idealization in science: *an approximation that, according to theory, would hold exactly in certain specifiable circumstances*. For want of a better word, I am going to call idealizations of this type "Newtonian" because of the special role they play in his approach to evidence. They include not only the phenomena from which his theory was inferred, but also further phenomena inferred from his theory, such as Kepler's ellipse. Science contains idealizations of all sorts of kinds – mathematical simplifications, schematics of experiments and apparatus, explanatory models, etc. They were commonplace before Newton. Both Galileo and Huygens, for example, had taken the curved surface of the Earth as flat in their treatment of projectile and pendular motion. My point is that Newton singled out and placed great emphasis on one particular kind of idealization: *approximations to the actual world that are deduced from his theory of gravity as holding exactly in specified circumstances*. Idealizations of this sort are not simplifications made in the process of arriving at physical theory; they are offspring of physical theory. Idealizations of this kind and the use to which Newton put them are a fifth way in which his *Principia* changed physics.

As defined here, a Newtonian idealization requires an overarching theory from which the claim of exactness in specified circumstances is inferred. Thus, Newton's law of universal gravity was not itself a Newtonian idealization, for there was no overarching theory which entailed its exactness in specified circumstances. Equally, Galileo's uniform acceleration in vertical fall in the absence of air resistance was not a Newtonian idealization; Galileo claimed it holds exactly in this circumstance, but he did not infer this claim from an overarching theory. As noted above, the theory of gravity in the *Principia* forms an intermediate level of physical theory between mere description of phenomenal regularities in the manner of Galileo's *Two New Sciences*, on the one hand, and laying out full mechanisms in the manner of Descartes's *Principia*, on the other. Intermediate though it may be, this level of theory is nevertheless sufficient to assign certain phenomenal regularities the status of a Newtonian idealization.

In general, what Newtonian idealizations do is to shift the focus of ongoing research to deviations from the ideal – that is, to discrepancies between theoretically deduced, idealized approximations to the world and the world itself. This shifting of the focus in research to discrepancies between theory and observation is a sixth way in which Newton's *Principia* changed physics – perhaps, the most conspicuous way. It was with Newton that the phrase "exact science" took on its current meaning. Instead of explaining away such discrepancies, as for example Galileo had invariably done, they became a source of continuing evidence that had promise of becoming increasingly discriminating. The complexity of the actual motions thus became not an impediment to high-quality evidence, but historically the source of it! And, having discrepancies between theory and observation became not a negative, but a positive.

Newton generally left research into such discrepancies and what they tell us about the world to future generations. The one exception was the non-Keplerian motion of our Moon. In his solution for the systematic deviation from the area rule known as Tycho's variation, Newton starts from the idealization of the Moon in a circular orbit with the Earth at the center and first determines how the gravitational force of the Sun would distort this orbit, elongating it in a direction perpendicular to the line from Earth to Sun. He then calculates the deviation from the area rule, obtaining seven-eighths of Tycho's value, and he ends by pointing out that including the effects of orbital eccentricity would make the calculation still more accurate. Notice what is happening here: one idealization, a simple circular orbit, is being replaced by another idealization that gives a better approximation, the idealization now known as the "variational" orbit produced by the perturbing effect of the Sun on the simple circular orbit. This process can continue, yielding a sequence of successive idealizations that should achieve increasingly closer agreement with observation. They are nonetheless all idealizations.

Newton gets an even more impressive result for the mean motion of the line of nodes, the 18-year cycle in lunar and solar eclipses known since the Babylonians, obtaining a result within three-tenths of one percent without considering eccentricity. The Moon's orbit is exceedingly complex; no attempt just to describe it geometrically had ever come close to the level of accuracy Kepler and others had achieved for the planets. Newton's announced purpose in making his calculations was to show that the best hope for genuine progress lay not in conventional observational astronomy, but instead in his theory of gravity and a sequence of deduced successive approximations.

Systematic deviations from Keplerian motion and other Newtonian idealizations can be thought of as a kind of phenomena in their own right. Only no one can observe them. They arise from the residual discrepancies between observation and idealizations deduced from theory – that is, from the difference that remains after Newtonian idealizations are subtracted from observation. I prefer to call them "second-order phenomena" for just this reason. They presuppose specific theory, and they cease having any meaning – they cease to exist – without that theory. As remarked above, on Newton's approach the focus in ongoing research shifts from primary phenomena and pursuit of a theory covering them to residual discrepancies between that theory and observation. With this shift, the goal in research becomes one of identifying second-order phenomena and determining what they are telling us about the world. If indeed the actual planetary motions are complex to a degree that exceeds exact mathematical description, then residual discrepancies will always remain. The requirement put on ongoing research is that the increasingly refined, and hence more complicated, idealizations result in continually smaller discrepancies.

Notice that, when the focus shifts in this way, further research is being predicated on the theory, and hence the theory has become an instrument entering constitutively into ongoing research. Earlier I said that Newton's generic mathematical theory was a tool for turning data into evidence – more specifically for turning phenomena into evidence for physical theory. Now we see how his physical theory was no less a tool for turning data into evidence – this time, second-order phenomena into evidence about such things as what other forces are contributing to the complex motions of the planets and their satellites. This is a seventh way in which Newton and his *Principia* changed physics. Before him the primary role of physical theory was to explain observed phenomena. With him, that role became subsidiary, superseded by the role physical theory is to play in ongoing research. This was the change that such contemporaries as Leibniz and Huygens had the most trouble seeing and appreciating.

15.5 Beyond the *Principia*: the logic of the continuing evidence

The *Principia* ends up spotlighting a number of potential second-order phenomena beyond the specific lunar inequalities for which Newton obtained results. There were also the past irregularities in the supposed 75-year return of what we call Halley's comet; the thoroughly confusing, not-yet-characterized departures of Saturn and Jupiter from Keplerian motion; and the as-not-yet-confirmed precession of the perihelia of the planets entailed by Newton's theory of gravity. The discrepancy between Newtonian theory and observation that became historically most important, however, was the precession of the apogee (or perigee) of the Moon. The apogee of the Moon shifts in very complicated ways from one orbit to the next. This precession nevertheless has a well-behaved mean value: on average the apogee moves forward slightly more than 3 degrees per revolution. Newton had used his precession theorem to calculate the effect of the gravitational force of the Sun, obtaining a mean precession of 1 degree, 31 minutes, 28 seconds – essentially half the observed value. A question he never managed to answer was why the Sun's gravity could readily account for

90 percent of the Moon's mean departure from the area rule and more than 99 percent of the mean motion of the nodes, yet only 50 percent of the mean motion of its apogee.

Newton was less than candid about this discrepancy in the *Principia*. In an appendix to the first English translation in 1729, however, John Machin, the orbital astronomer who was closest to Newton during the 1720s, made the problem clear to all:

> But the [mean] motion of the apogee, according to this method, will be found to be no more than 1°37′22″ in the revolution of the moon from apogee to apogee, which (according to observations) ought to be 3°4′7.5″.
>
> So that it seems there is more force necessary to account for the motion of the moon's apogee than what arises from the variation of the moon's gravity to the sun, in its revolution about the earth.
>
> But if the cause of this motion be supposed to arise from the variation of the Moon's gravity to the Earth, as it revolves round in the elliptic epicycle, this difference of force, which is nearly double the former, will be found to be sufficient to account for the motion, but not with the exactness as ought to be expected. Neither is there any method that I have ever yet met with, upon the commonly received principles, which is perfectly sufficient to explain the motion of the moon's apogee.
>
> (Machin 1729, p. 30f.)

Machin went on to concede that, so far as he could see, it is impossible to derive the motion of the apogee and the alteration of the eccentricity "from the laws of centripetal forces."

During the 1740s Euler, Clairaut, and d'Alembert took up the problem, each concluding that solar gravity could account for only half the observed precession. Clairaut went the furthest, for he took the eccentricity of the lunar orbit into consideration. Specifically, he adapted a method of Euler's to derive all the terms for the Sun's effect in which eccentricity occurs to the first power, and still found Newtonian theory giving only half of the observed motion. Based on this, and some discrepancies between recent geodetic measurements and Newtonian theory, he registered a paper with the Royal Academy of Paris proposing that Newton's law of gravity has to be amended with a $1/r^4$ term. This provoked quite a debate within the Royal Academy. As this debate continued, Clairaut decided to derive the terms for the Sun's effect to the next highest order, obtaining terms in eccentricity squared and cubed. When he calculated their effect, he discovered to his surprise that, even though the lunar eccentricity is less than 0.06, these higher-order terms are not negligibly smaller in magnitude than the first-order eccentricity terms. Together with those terms they yield a nearly exact value for the mean motion of the apogee out of the inverse-square effect of the Sun's gravity (Waff 1976). D'Alembert, ever the querulous one, went on to calculate the contribution from terms of the next highest order, confirming that they do not mess up Clairaut's result.

Two points should be made about Clairaut's efforts before turning to the historical importance of his result. First, Clairaut's reasoning when he proposed adding an inverse r^4 term was that the data from planetary orbits which Newton had used in deriving the law of gravity involved distances that were too large to expose the need for the additional term; in other words, Clairaut was saying that the data on which Newton had relied were parochial. Second, the perturbational approach used in Clairaut's calculations – and in virtually all subsequent calculations in celestial mechanics – introduces another layer of complication in the logic of the evidence: Clairaut was deriving not a Newtonian idealization for a specific case of the three-body problem of the Sun, Earth, and Moon, but instead a computational *approximation* to such an idealization. The intractability of the mathematics stood in the way of a rigorous derivation of the exact solution called for in my definition of a Newtonian idealization. Given any remaining discrepancy, then, a question arises about the extent to which it reflects imprecision in the method of calculation versus the need for some refinement, like an unaccounted for force, in the idealized model presupposed in the calculation.

Clairaut's result was much heralded. In a private letter to him, Euler remarked,

> the more I consider this happy discovery, the more important it seems to me, and in my opinion it is the greatest discovery in the Theory of Astronomy, without which it would be absolutely impossible ever to succeed in knowing the perturbations that the planets cause in each other's motions. For it is very certain that it is only since this discovery that one can regard the law of attraction reciprocally proportional to the squares of the distances as solidly established; and on this depends the entire theory of astronomy.[13]

A year later Euler made the same point in print, arguing that they could now be certain that there are inverse-square forces between Jupiter and Saturn causing the confusing irregularities in their motions that had been observed:

> since M. Clairaut has made the important discovery that the movement of the apogee of the Moon is perfectly in accord with the Newtonian hypothesis . . . , there no longer remains the least doubt about this proportion . . . And if the calculations that one claims to have drawn from this theory are not found to be in good agreement with observations, one will be always justified in doubting the correctness of the calculations, rather than the truth of the theory.
>
> (Euler 1752, p. 4f.)

13 Letter from Euler to Clairaut, 29 June 1751, in Bigourdan (1929, p. 38f.); translation from Wilson (1980, p. 143).

The hyperbole in these pronouncements is not so extreme as it may at first appear. As Euler explained in his *Theory of the Moon's Motion* of 1753,[14] the mean motion of the apogee provides an exceptionally sensitive measure of the exponent in the rule of centripetal force. The exponent is exactly -2 if and only if the orbit is perfectly stationary *in the absence of any forces beyond the centripetal force holding it in orbit.* The trouble, of course, is that the Moon's orbit is not stationary, but precesses on average 3 plus degrees per revolution. Even so, as Newton showed in the *Principia*, one can still conclude that the centripetal force on the Moon is inverse-square, *quam proxime.* The question whether it is exactly inverse-square can then be addressed by identifying forces beyond the Earth's gravity and seeing whether any discrepancies remain once the effects of these forces are taken into account. How is one to identify such further forces? By first taking into account the gravitational forces from the Sun and the planets, and then seeing what discrepancies, if any, remain.

This brings me to an eighth way in which Newton's *Principia* changed physics. His approach opened the way to a new form of evidence – *evidence indirectly accruing to a theory from the success of research predicated on it.* The original evidence for Newton's law of gravity showed at most that it holds to high approximation, yet he took it to hold exactly and deduced idealizations from it. This strategy leaves open the question, how exact is the law? We now see that ongoing research on deviations from these idealizations can continue to bring evidence to bear on the law in general and on this question in particular – albeit indirect evidence. Clairaut's result, together with the observed lunar precession, provided direct evidence that, whatever other forces are perturbing the lunar orbit, they are much smaller than the perturbing force from the Sun's gravity. Indirectly, however, it provided evidence that Newton's law of gravity holds to a still higher level of approximation than his original evidence implied. More generally, focusing research on deviations from Newtonian idealizations and demanding progressively smaller discrepancies between observation and the idealized model of the world is a strategy for exposing limitations in this law. Correlatively, because the idealizations are deduced from the law, taken as exact, evidence accrues to the law from continuing success in pinning down robust physical sources of still remaining deviations.

I claim that Newton's new approach to marshalling evidence was a response to the complexity of the world. Those before Newton despaired of any such complexity, concluding it would always limit the quality of evidence that can be achieved in science. What Newton did was to find a way to turn the very complexity into a source of increasingly more telling evidence. This, to me, was the ultimate genius of the *Principia.*

14 Euler (1753, pp. 71–72); translation in Wilson (1980, p. 144).

The logic of this evidence needs to be made clear, for the Clairaut example can be misleading. At first glance, one might think that Clairaut deduced the theoretical mean motion of the lunar apogee, and its close agreement with observation therefore provided hypothetico-deductive evidence for the law of gravity. That is a mistake. For Clairaut to have deduced the motion, he would first have had to assume that no other forces are at work beyond the perturbing force of the Sun – a question that was surely still open. (Newton himself at one point intimated that the missing one and a half degrees in the mean precession of the lunar orbit might be coming from the Earth's magnetic field (Newton 1999, p. 880).) Rather, Clairaut was only deducing the effect of a specific perturbing force entailed by Newton's theory of gravity. More generally, all such calculations of orbital motions in celestial mechanics are merely deducing the effects of the forces specified. When the result of any such calculation matches observation very closely, the appropriate conclusion is that any further perturbing forces either do not exist or are of much lesser consequence. A failure to match observation leaves open the possibility that some other force is making a significant contribution. A less outspoken version of Euler's statements about the strong evidence Clairaut had provided for the inverse-square would still have been valid even if it had turned out that the missing one and a half degrees was from the Earth's magnetic field.

To put the matter differently, the test to which Newton's theory is put by the deviations from his idealizations is more subtle than a simple hypothetico-deductive construal suggests. Newtonian idealizations are by definition ones that, according to the theory of gravity, would hold exactly in specified circumstances. But then any deviation from them must result from some physical departure from those circumstances, an additional celestial force not yet taken into consideration. The implication, in other words, is that any deviation from an idealization must be *physically significant* within the context of the theory – this in contrast, for instance, to being merely a reflection of the mathematical scheme that happened to have been chosen in curve fitting. The test to which Newtonian theory is put in ongoing research centers on the question, *Is every deviation from a Newtonian idealization physically significant?* The evidence that accrues to Newtonian theory comes from pinning down robust physical sources of deviations – a continuing process that ought to result in ever smaller discrepancies between observation and the idealized representation of the world. Whenever all residual discrepancies drop below a then-current level of accuracy of observation, the appropriate conclusion must have a somewhat Popperian flavor: at least for the moment, observation has ceased providing any basis for identifying either further complications in the world or respects in which theory is inadequate.

The conclusion, *any other perturbing forces are of much lesser consequence*, is a variant of a problematic auxiliary assumption, *all forces acting on the planets other than the designated gravitational forces have very small effects*, required in hypothetico-deductive construals of the evidence. Both are variants of Karl Hempel's *the constituent bodies of the system are subject to no forces other than their mutual gravitational attraction* – his paradigmatic example of a "proviso" in his paper, "Provisos: A Problem Concerning the Inferential Function of Scientific Theories" (Hempel 1988). On the account of the logic of the evidential reasoning I am offering, these are not assumptions in the deductions of celestial mechanics at all. The deductions are spelling out the (idealized) consequences of a set of specified forces. The point of the resulting Newtonian idealizations is not as such to test the theory of gravity by making predictions with it, but rather to address the question, are any forces beyond those specified of consequence? Hempel's provisos, instead of being assumptions in the deductions, are conclusions that emerge when the answer to this question is no – that is, when the discrepancies between the calculated and the observed motions are sufficiently small.

Our sense that celestial mechanics over a period of centuries generated extraordinary support for Newton's law of gravity stems not from its having continually yielded predicted motions within observed accuracy (which, in fact, it never really did), but from the success in pinning down – that is, identifying and further confirming – the physical sources of forces responsible for ever more subtle complexities in the observed motions. The extent to which orbital motions are dominated by gravitational forces has been among the most remarkable findings of celestial mechanics.

A long tradition of carelessly talking about evidence in celestial mechanics as if it were straightforwardly hypothetico-deductive has obscured the extent to which the focus of ongoing research has been on questions about further forces. In saying that Newton's theory of gravity has been an instrument in post-*Principia* research in celestial mechanics, I mean more than just that this theory has been presupposed in instance after instance of evidential reasoning throughout that research. Because the overall pattern has been one of successive approximations, the evidence for the physical sources of the increasingly smaller deviations from the current ideal presupposes not only Newtonian gravity, but also the previously identified sources of the larger deviations from the earlier ideals. In other words, the ongoing evidential reasoning has presupposed the theory of gravity in an increasingly ramified fashion. To question the law of gravity is to throw into question a huge collection of facts (or, if you prefer, quasi-facts) about the world that post-*Principia* research has established. The burden of proof required to discard the law of gravity thus became increasingly large – which is the same thing as saying that the law became increasingly entrenched.

At the beginning of this chapter I posed a question about the much higher quality of evidence after the *Principia* than before it. The continuing evidence in gravitation research, and not Newton's original evidence, is the high-quality evidence in question. Listing all the evidence of this sort that has unfolded over the last 300 years would be a Herculean task. Still, it is instructive to list a few highlights: (1) Clairaut's prediction of the month of return of Halley's comet in 1759 after taking the gravitational forces of Jupiter and Saturn into account; (2) Laplace's 1785 discovery of the 890-year "Great Inequality" in the motions of Jupiter and Saturn; (3) Leverrier and Adams deducing the existence of an eighth planet, Neptune, from residual anomalies in the motion of Uranus (1846); (4) the Hill–Brown theory of the Moon (1919), involving more than 1400 physically significant terms, which finally brought lunar theory to the level of accuracy of the planets and revealed as well that the Earth's rotation is not uniform, and hence that sidereal time is not an exact measure of time.[15]

The key point, however, is that the process of research is continuing, for there will always be discrepancies. The difference now is that they are at levels of significant figures of which Euler and Clairaut, much less Newton, scarcely ever dreamed.

15.6 Parochialism and the continuity of evidence

The glaring omission in my list of highlights is the precession of the perihelion of Mercury, the discrepancy that finally falsified, so to speak, Newton's law of gravity. Newton already knew that most of the apparent precession of Mercury's perihelion is just that – apparent, stemming from the precession of the equinoxes, the 26,000-year wobble of the Earth. He had no way to calculate the true precession implied by his theory of gravity, in part because he had no way to determine the mass of Venus. By the end of the nineteenth century it became clear that Newton's theory was 8 percent slow for the true 225,000-year precession of Mercury's orbit. This 43 arc-seconds per century residual proved recalcitrant: Newtonian theory was unable to provide any physical source for it, and hence it appeared not to make physical sense within the context of that theory. Later, of course, it turned out to be evidence for the new theory of gravity of Einstein's general relativity.

This residual discrepancy in the very slow motion of Mercury's perihelion shows how Newton's response to the complexity of orbital motion was, at the same time, a response to the risk that our observations are somehow parochial. What better way was there to expose any such parochialism than to push his theory for all it is worth until some subtle discrepancy emerges that might shed light on just how it is parochial? The obvious alternative, *contra* Newton's fourth Rule of Reasoning, was to try to obviate parochialism from the outset by proposing a wide range of competing theories compatible with the available data and then identifying cross-roads experiments to choose among them or to falsify them one-by-one. One problem with this alternative was the degree to

15 These four as well as other contributions from continuing research in orbital mechanics are discussed in Smith (forthcoming b).

which the complexity of the world would have limited the quality of evidence in trying to decide early among the competing theories. The more serious problem, however, was the absence of any way of assuring that the list of proposed alternatives would cover respects in which the available data were indeed parochial. The specific respect in which Newton's data are now known to have been parochial was not something anyone imagined at the time.

The residual in the precession of Mercury's perihelion also brings a distinction into sharper focus that was alluded to in the preceding section, the distinction between Newtonian idealizations and curve-fits. At the end of the nineteenth century Simon Newcomb prepared a new set of planetary tables that, together with the theories of the orbits underlying them, remained the basis for orbital predictions until the switch to direct numerical integration of the equations of motion on high-speed computers in the 1980s. Newcomb's tables were based on Newtonian gravity plus an added term in the calculation of the secular precession of the perihelia of the four inner planets. This term, which turned out to amount to a fudge factor, consisted of a constant times the mean motion of the planet in its orbit.[16] It added 43.37 arc-seconds per century in the case of Mercury, and less for the other planets. With this term included, the calculated orbits involved an element of curve-fitting, and hence they were no longer, strictly speaking, Newtonian idealizations. No longer could any systematic discrepancy between the calculated and observed precessions be automatically taken to be symptomatic of some physical source not yet taken into account. For, the discrepancy might instead represent some physically arbitrary feature in the curve-fit. In particular, suppose a new systematic discrepancy were to emerge in the case of Mercury much smaller than the prior 43 arc-seconds, say a discrepancy around 0.4 arc-seconds. Why should that discrepancy automatically be taken as a sign of some yet-to-be-noted physical effect when it could just as well be attributed to the choice of mean speed as the curve-fitting parameter or to the decision not to include terms in mean-speed squared?

Both Newtonian idealizations and curve-fits can be carried out in sequences of successive approximations in response to a complex world. Curve-fits aim at prediction, with the mathematical scheme chosen to reflect a trade-off between accuracy of fit and calculational ease. Least-squares curve-fits have the virtue of minimizing the expected value of the square of the error in prediction, relative to the adopted mathematical scheme, and errors in prediction are expected to be Gaussian. In general, whether the curve-fitting criterion is least-squares or otherwise, the goal is for errors in prediction to have a random character; and, in that regard, curve-fits attempt not to highlight discrepancies, but to achieve prediction within a certain level of precision, in the process sweeping lesser

16 Specifically 0.0000000806 times the centennial mean motion; see Newcomb (1898, p. 12).

discrepancies under the rug. Discrepancies between Newtonian idealizations and observation, by contrast, are not expected (or even desired) to have a random character, for the driving research question is, what further physical factors, if any, need to be taken into consideration?

The distinction between Newtonian idealizations and curve-fits is especially important when worried about the possibility that available data are somehow misleadingly parochial. With each successive approximation in curve-fitting, physical sources of features in the data become progressively more submerged in a welter of choices embedded in the mathematical scheme. As a consequence, there are multiple potential sources for a recalcitrant discrepancy besides some respect in which the accessible data are physically parochial. By contrast, the further a sequence of successive approximations progresses with Newtonian idealizations, the stronger the grounds are for attributing any recalcitrant discrepancy to some physical parochialism. For, the alternative, that the theory has been (by its own standards) radically wrong all along, is countered by the record of success so far in pinning down robust physical sources of discrepancies.

Within two decades of Newcomb's new orbital tables, Einstein put forward his theory of general relativity, and Newcomb's curve-fitting response to the residual in Mercury's perihelion ceased to matter. Einstein's relativity produced a conceptual revolution in physics, but not really a revolution in evidence. For, Newtonian gravity holds as an asymptotic limit of Einstein's, specifically the static weak-field limit. This had two important consequences. First, save for qualifications about levels of precision, all the evidence for Newtonian gravity carried over immediately to Einsteinian gravity. Physics did not have to go back to an earlier time and begin reconstituting evidence. The data that had been evidence for Newtonian gravity were guaranteed to be evidence for Einsteinian gravity as well insofar as, under the conditions of our solar system, Newtonian gravity amounts to an approximate special case of Einsteinian, and evidence for a special case counts as evidence, though often of reduced strength, for the more general theory of which it is a special case. Second, Einstein's theory did not out-and-out nullify the evidential reasoning supporting Newton's theory. That is, it did not entail that the evidence supporting the prior theory was merely illusory, and never truly evidence at all. For, the steps in the original reasoning can still be justified, though the justifications themselves have to be amended to include qualifications – for example, the qualification that a Euclidean metric provides a good approximation so long as the gravitational field is weak. If the reasoning had not remained so justified, then the 43 arc-second residual, taken in itself, could not have provided evidence for Einstein's theory in the manner in which it did, for that residual is a Newtonian second-order phenomenon that presupposes Newtonian gravity; were the evidence for Newtonian gravity illusory, the specific value of 43 arc-seconds would be nothing but an artifact of an illusion.

This point can be put in another way. The transition from Newtonian to Einsteinian gravity certainly did entail a revolutionary discontinuity in conceptual structure. Nevertheless, because Newtonian gravity holds in the static weak-field limit, the transition did not entail any discontinuity in evidence. The residual 43 arc-seconds per century in the precession of the perihelion of mercury is a Newtonian second-order phenomenon that turns out to be physically significant, but only within the context of the less parochial theory. From the point of view of Einstein's theory, Newton's is a limited special case reflecting a systematic bias in the data to which we have ready access, in particular orbital data from within our solar system. From the point of view of Newton's theory, on the other hand, Einstein's is a more general theory – one among an indefinite number of possible more general theories – that a strictly Newtonian phenomenon helped single out and substantiate. The transition from Newton to Einstein yielded the discovery that one specific respect in which the readily available gravitational data within our solar system are parochial is that the fields are so weak and so nearly static.

What we have here is another kind of idealization in physics: a theory that, even though it would never hold exactly in any realizable circumstance, nevertheless holds in a mathematical limit with respect to a more general theory. Let me call these "limit-case" idealizations. They have a different role in the development of evidence from the Newtonian idealizations I have been emphasizing so far. Their most important contribution is to allow evidence to remain continuous and hence cumulative across theory change, especially across theory change involving removal of parochialisms.

Although it has gone largely unnoticed, continuity of evidence is itself a form of evidence. Of course, the continuity of evidence from Newton to Einstein cannot be evidence for Newton's theory itself, or even Einstein's. It is evidence for something more basic that is common to both. In taking the huge inductive leap from inverse-square gravity and the orbits of six planets to universal gravity, Newton was making two tacit, but nonetheless indispensable taxonomic assumptions: (1) gravity marks a distinct natural kind or, to use Newton's phrase, a physical species; and (2) orbital motions of our planets and their satellites represent a pure enough example to typify this species as a whole. Both of these assumptions, at the time Newton made them, could not help but fall largely in the category of wishful thinking.

The research predicated on Newton's law of gravity over the next two centuries succeeded spectacularly in reducing the gap between theory and observation; and this success provided support for these two taxonomic assumptions. All of this success nevertheless came from phenomena within our planetary system over a very short period of astronomical time. Consequently, none of it spoke directly to the possibility that gravity is an accidental feature of our solar system, in much the way that many geological phenomena are mere artifacts of the Earth's history, and not symptomatic of deep physical laws.

Besides revealing the weak-field parochialism of our planetary system, general relativity has enabled data from the universe at large to become evidence bearing on gravitation theory. A prerequisite for continuity of evidence in theory change is that the taxonomy underlying the old theory remain essentially intact within the new theory. The fact that evidence remained continuous from Newtonian to Einsteinian gravity has accordingly provided much stronger support than ever before for the claim that gravity marks a distinct physical species.

Surprising though it may be, limit-case idealizations are something else that Newton introduced in the *Principia* and hence a ninth way in which this book changed physics. Newton, of course, had concluded that the data supporting the theory of uniform gravity acting along parallel lines developed by Galileo and Huygens were parochial, coming as they all did from the narrowly confined region near the surface of the Earth. This theory and the evidence for it were nonetheless important to Newton for a series of reasons: (1) he expressly invokes results by Galileo confirming that the acceleration of gravity is independent of weight as evidence that mass is proportional to weight (Newton 1999, p. 806); (2) he similarly invokes Galileo's vertical fall and parabolic projection and Huygens's pendulum results as evidence for his first two laws of motion, and indeed Huygens's pendulum measurements of surface gravity offered the best evidence for those laws (Newton 1999, p. 424); and (3) Huygens's measured value of surface gravity, which presupposed uniform gravity acting along parallel lines toward a flat surface, provided crucial evidence that terrestrial gravity extends to the Moon (Newton 1999, pp. 803–805).

The way in which Newton chose to treat uniform gravity as a limit-case of universal gravity will surprise anyone not thoroughly familiar with the *Principia*. Newton does not argue that Galilean gravity is simply an approximation to inverse-square gravity over small distances – that is, distances over which the variation in the acceleration of gravity is too small to matter. Instead, he treats it as a limit-case of gravity that varies linearly with distance from the center of a spherical Earth, specifically the limit at the surface as the Earth's curvature approaches zero. The statement of this limit-case idealization occurs in Section 10 of Book I of the *Principia*, in a corollary to one of the propositions on hypocycloidal pendulums – that is, pendulums the arc of which is defined by the trajectory of a circle rolling not on a flat plane, but on the underside of a spherical surface (Figure 15.2):

> *Prop. 52, Cor. 2.* Hence also follows what Wren and Huygens discovered about the common cycloid. For if the diameter of the globe is increased indefinitely, its spherical surface will be changed into a plane, and the centripetal force will act uniformly along lines perpendicular to this plane, and our cycloid will turn into the common cycloid. But in that case the length of the arc of the cycloid between that plane and the describing point will come out equal to four times the versed sine of half of the arc

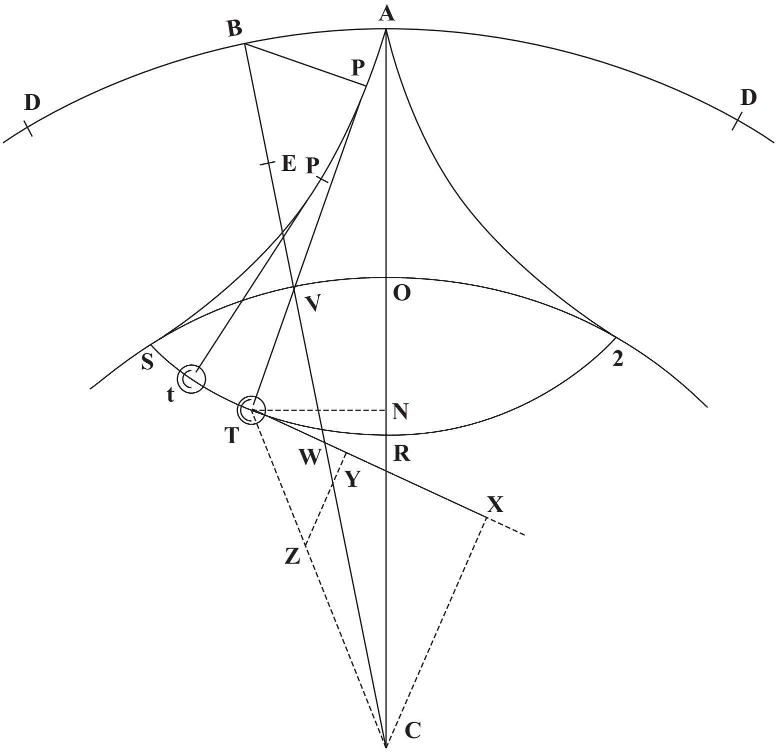

Figure 15.2 The figure accompanying Propositions 51 and 52 on isochronous motion of pendulums

of the wheel between that same plane and the describing point, as Wren discovered; *and a pendulum between two cycloids of this sort will oscillate in a similar and equal cycloid in equal times, as Huygens demonstrated. But also the descent of heavy bodies during the time of one oscillation will be the descent Huygens indicated.*

Moreover, the propositions that we have demonstrated fit the true constitution of the earth, insofar as wheels, moving in the earth's great circles, describe cycloids outside this globe by the motions of nails fastened in their perimeters; and pendulums suspended lower down in mines and caverns must oscillate in cycloids within globes in order that all their oscillations may be isochronous. For gravity (as will be shown in Book 3) decreases in going upward from the surface of the earth as the square of the distance from the earth's center, and in going downward from the surface is as the distance from that center.

(Newton 1999, p. 555f., emphasis added)

As Newton shows later in Book I, his universal gravity entails that, below and up to the surface of a uniformly dense sphere, the net gravitational force varies linearly with distance, while above the surface it varies as the inverse-square (Newton 1999, pp. 593–597 and 617f.).

Galileo's and Huygens's results can be shown to hold to high approximation in inverse-square gravity so long as the vertical distances are small.[17] What then does Newton gain with his limit-case idealization? In Huygens's measurement, the strength of surface gravity is inferred, via his law of the cycloidal pendulum, from the measured period, which he had shown does not vary with the length of the arc of the bob. This isochronism was crucial to Huygens's measurement beyond its being explicit in the law. Thanks to isochronism, no attention needed to be given to the length of the bob's arc and whether it was varying during the measurement of the period. Isochronism was accordingly a key factor in the claimed precision of Huygens's measurement, a precision important to Newton. Now, hypocycloidal pendulums are isochronous under gravity that varies linearly with distance from the center (Prop. 51), but not under inverse-square gravity! Therefore, what Newton's specific limit-case idealization enabled him to show was that the logic underlying Huygens's measurement is not nullified when uniform gravity acting along lines parallel to one another is replaced by his universal gravity. (Notice that this is precisely what Newton said in the portion of the quotation I italicized above.) Remarkably, the *Principia* thus actually goes to the trouble of confirming continuity of evidence in the transition from Galilean to Newtonian gravity.

15.7 Newton or Newtonian?

Employing limit-case idealizations to maintain continuity of evidence across theory change is the ninth and last of my ways in which the *Principia* changed how evidence is developed in physics. Table 15.1 recapitulates the nine ways for the convenience of the reader.

I see these not as nine distinct ways, but as nine aspects of a single change: *a new approach in which theory is first and foremost an instrument for developing evidence, and evidence of increasingly telling quality is then brought to bear on it indirectly through the research predicated on it.* More important than this summary description, however, is the degree to which these nine elements mesh with one another to form a coherent whole. I claim that they gain this unity from their being a response to Newton's conclusion that the true motions of the planets are hopelessly complex and his worry that the data to which we have ready access may be misleadingly parochial. I have trouble imagining a more reasonable response to the complexity of the true motions and the likely

17 Doing so amounts to treating uniform gravity acting along parallel lines as a mere curve-fit approximation to Newton's universal gravity.

Table 15.1 *Nine aspects of how Newton's* Principia *changed physics*

1.	Physics has to include its own theory of measurement
2.	Develop generic mathematical theory to provide "inference-tickets"
3.	Restrict physical theory to principles that phenomena dictate
4.	Leap from approximative evidential reasoning to exact theory
5.	Idealizations that would hold exactly in specified circumstances
6.	Shift focus of ongoing research to deviations from such idealizations
7.	Physical theory becomes an instrument for turning data into evidence
8.	Evidence accrues to a theory from success of research predicated on it
9.	Limit-case idealizations enable continuity of evidence across theory change

parochialism of our observational situation than this one. One can scarcely say of those who have traced the path initiated by the *Principia* that they have "made trial of nature in vain."

Two and a half years after the *Principia* was first published, Huygens published a response to Newton's theory of gravity, *Discourse on the Cause of Gravity*, bound together with his *Treatise on Light*. In the Preface to the latter he offers a wonderfully succinct statement of the then prevailing view about how evidence is to be developed in empirical science:

> One finds in this subject a kind of demonstration which does not carry with it so high a degree of certainty as that employed in geometry; and which differs distinctly from the method employed by geometers in that they prove their propositions by well-established and incontrovertible principles, while here principles are tested by the inferences which are derivable from them. The nature of the subject permits of no other treatment. It is possible, however, in this way to establish a probability which is little short of certainty. This is the case when the consequences of the assumed principles are in perfect accord with the observed phenomena, and especially when these verifications are numerous; but above all when one employs the hypothesis to predict new phenomena and finds his expectations realized.
>
> (Huygens 1888–1950, XIX, p. 454)[18]

Newton's famous pronouncement in the General Scholium that he added at the end of the second edition of the *Principia* twenty-three years later was presumably, at least in part, a response to this statement by Huygens:

> I have not as yet been able to deduce from phenomena the reason for these properties of gravity, and I do not feign hypotheses. For whatever is not deduced from the phenomena must be called a hypothesis; and

18 The English translation is from Matthews (1989, p. 126). The hypothesis which Huygens had most in mind was the longitudinal wave theory of light.

> hypotheses, whether metaphysical or physical, or based on occult qualities, or mechanical, have no place in experimental philosophy. In this experimental philosophy, propositions are deduced from the phenomena and are made general by induction.
>
> (Newton 1999, p. 943)

Newton, however, says nothing more about his approach and why it may be better. In particular, nowhere in the *Principia* does he invoke the complexity of the orbital motions to argue either that too many disparate hypotheses can meet Huygens's requirements or that a hypothetico-deductive approach offers less promise of bringing to light the physical sources of small discrepancies between theory and observation. Indeed, nowhere in the *Principia* does he even intimate that his alternative approach involves remotely the logical intricacy that I have attributed to it. A natural question, then, is whether the approach I have laid out is better called Newtonian rather than Newton's. How much of it did Newton himself see?

While this question is clearly of historical interest, especially to Newton scholars, it is not of central importance to this chapter. The goal of this chapter has been to lay out a picture of the general logical structure of the evidence across the history of research in Newtonian gravity and to trace key constituents of this logical structure to Newton's *Principia*. The data entering into this evidence extends from Tycho Brahe's efforts a century earlier well into the twentieth century in the case of orbital motion. As crucial to the history of research in Newtonian gravity as Newton and his *Principia* were, this research was carried out by a large community that stretched across many generations. The individuals forming that community focused far more heavily on specific, narrow questions in evidence, and not on the general logic of the evidential reasoning across the entire history. Consequently, although what those individuals said and thought is relevant, it is of limited weight when judging the adequacy of the picture of the logic presented in this chapter. One should think of this chapter as emulating the perspective of a review article, unusual only in the scope of time covered and the limited attention given to specific items of evidence. The decisive issue in judging the picture of the overall logic presented here should be its coherence.

That said, let me return to the question of how much of my proposed "Newtonian" approach to evidence Newton himself saw. My guess is, all of its key constituents and, at least on occasion, their *potential* for coming together to form a whole. Much of my reason for saying this is "autobiographical" and hence not of much moment for others: I came to see this logic from repeatedly working through the *Principia* while teaching it cover-to-cover. On a less personal note, each of the items listed in Table 15.1 has been tied in this chapter to specific passages in the *Principia*. I chose Clairaut's work to illustrate continuing indirect evidence from success in using the theory as a tool in ongoing research, but I could almost as well have used Newton's quantitative results on other lunar inequalities. Those results, moreover, are not the only place where Newton proposes to proceed by successive approximation to increasingly refined idealizations – what I. B. Cohen (1980, pp. 3–154) called the "Newtonian style."

Arguing that the *Principia* provides explicit basis for each of the nine items in my list is one thing; arguing that Newton saw much or all of the logical structure I claim they form is another. The argument for that has to come from more subtle features of his text, such as his precise phrasing of his Rules of Reasoning and his careful use of the subjunctive when discussing whether the orbits of the planets are Keplerian and stationary. There are also features of the *Principia* that make totally good sense if he was paying attention to nuances in the logic I have proposed, but are difficult to explain otherwise. The most notable of these are his refusal to infer the inverse-square variation from the Keplerian ellipse and his treatment of Galilean uniform gravity as a limiting case of gravity that varies linearly with distance. Finally, in a similar spirit, my picture of the logic of evidence in the *Principia* absolves Newton of stupidity (or dishonesty) in claiming to have derived the law of gravity from phenomena.

Saying that Newton saw much of the logic I have described does not mean that there was some moment when he had a clear, comprehensive vision of the whole picture and thereafter consciously fashioned the *Principia* accordingly. He appears to have been fully aware from early on that his inferences from phenomena involve "if *quam proxime*, then *quam proxime*" reasoning. Propositions establishing this *quam proxime* form of relevant "if-then" statements occur in the very first draft of Book I, and even the registered version of the "De Motu" tract shows signs of his knowing that the *quam proxime* form of "if a Keplerian ellipse, then inverse-square" does not hold (Smith 2002b, p. 40f.). Not so clear is when, and how fully, Newton saw that deviations from what I have called Newtonian idealizations can provide an evidential basis for a sequence of successive approximations in ongoing research. When he first calls attention to the intractable complexity of planetary motion, in the augmented version of "De Motu," he presents the vagaries as an obstacle in determining the proper Keplerian orbits and seems resigned to never being able to do anything constructive with the deviations from Keplerian motion (Newton 1962, p. 281). In the initial draft of what became Book III of the *Principia* the inequalities in the lunar orbit are treated only qualitatively, and the intent seems merely to be to eliminate the apparent counterexample the Moon offers to Kepler's rules (Newton 1934, p. 577). My suggestion, then, is that Newton saw the possibility of using the deviations as the basis for successive approximations when his quantitative results on the lunar inequalities emerged, between the first draft of Book III in 1685 and the final draft in late 1686 or early 1687.

Newton had good reasons to be cautious about putting too much of the evidential burden for universal gravity on success in pinning down the physical sources of deviations from Keplerian motion. In contrast to the limited quantitative results he achieved on the restricted form of the three-body problem involving the Sun, Earth, and Moon, he obtained no quantitative results at all on the three-body problem posed by the Sun, Jupiter, and Saturn. Worse, the factor of two error in his derivation of the precession of the lunar apogee raised the distinct possibility that the Earth's magnetism was contributing to this effect. Newton tells us that the magnetic force, "in receding from the magnet, decreases not as the square but almost as the cube of the distance, as far as I have been able to tell from rough observations" (Newton 1999, p. 810);[19] and he knew that a superposed inverse-cube centripetal force is precisely what is needed to make an orbit precess (Newton 1999, pp. 535–539). If, however, non-gravitational forces have any significant effect on the motions of the planets or their satellites, the prospects for developing continuing evidence for universal gravity out of the vagaries of the motions is not so straightforward. For, laws of these non-gravitational forces would first have to be established, independently of those vagaries, and even with those laws in place, problems would potentially remain in specifying conditions for their applicability to specific celestial motions – for example, what is the fractional iron content of the Moon?

From his work on the tides Newton knew how much more difficult quantitative analysis becomes when non-gravitational forces are involved. If they are not virtually negligible in orbital motions, then the process of pinning down physical sources of deviations would likely be long and maybe tortuous, and the evidence accruing to universal gravity would be of reduced strength.

Furthermore, we should not lose sight of the limits of observational accuracy in astronomy during Newton's lifetime. The need to correct observations for the effects of solar parallax and atmospheric refraction had long been recognized, but the precise magnitudes of those corrections remained under dispute throughout Newton's lifetime, and no consensus had been reached on the need for a further speed-of-light correction at the time the *Principia* was published.[20] The need for still further corrections for the aberration of light and the nutation of the Earth emerged shortly after Newton died (Bradley 1728 and 1748). Thus, the prospects for increasingly precise observation of the sort needed to support successive approximations beyond the first level of refinements became much clearer only after Newton.

A prominent physicist responded to the account of the evidence for Newtonian gravity given above by remarking, "Newton was lucky."[21] That is surely correct on two counts. He was lucky that a relationship as mathematically

19 Newton was not in error here, for the dipole effect of a magnet gives rise to an inverse-cube variation.

20 Cassini still insisted that the irregularity in the timing of the eclipse of Io came from an inequality in its orbit; see Halley (1694).

21 Kenneth G. Wilson, in conversation.

simple as his law of gravity remained intact across two centuries of pursuit of ever greater precision.[22] And he was lucky in the degree to which gravity dominates celestial motions, making the task of marshalling evidence out of those motions far easier than it would otherwise have been.[23] Newton had no basis for expecting either of these eventualities to work out remotely as well as they did. Gravitation research has been successful, however, not merely because the empirical world happened to cooperate, but also because it has followed an approach that enabled continuing evidence to be brought to bear from increasingly subtle complexities in the motions. Its following that approach was not a matter of luck. In his research in optics Newton conducted experiment after experiment, with only slight variations, in order to address possible loopholes in the experiments that he ultimately published.[24] In the *Principia* Newton shows a similar constant concern for evidential loopholes that

22 The algebraic simplicity of the law as Newton formulated it was an automatic consequence of his inferring the law from phenomena by means of approximative reasoning. For, this blocked him from incorporating any feature into the law unless the phenomena dictated it, and insofar as the original phenomena amounted to first-order approximations to the real motions, nothing in them was going to dictate further complications. But that gave all the more reason to expect that a need for complications might well emerge as research went beyond those first-order approximations. For example, the law does not include time as a variable – something that might at least have raised questions early on.

23 A letter Newton wrote to Leibniz in 1693 shows that he anticipated this possibility:

> For since celestial motions are more regular than if they arose from vortices and observe other laws, so much so that vortices contribute not to the regulation but to the disturbance of the motions of planets and comets; *and since all phenomena of the heavens and of the sea follow precisely, so far as I am aware, from nothing but gravity acting in accordance with the laws described by me*; and since nature is very simple, I have myself concluded that all other causes are to be rejected and that the heavens are to be stripped as far as may be of all matter, lest the motions of planets and comets be hindered or rendered irregular.
>
> (Newton 1959–1977, III, p. 287; emphasis added)

Perhaps Newton is here being disingenuous with Leibniz, who had published his own vortex theory of Keplerian motion four years earlier, but he knew perfectly well that tidal phenomena do not all follow precisely from the laws described by him, and his suggestion that celestial phenomena follow *precisely* was at best wishful thinking. (Two years before this letter Newton had asked Flamsteed for observations of Jupiter and Saturn over a fifteen-year period, presumably because he wanted to answer the question of how precisely their motions follow from the law of gravity.) The evidence that gravity is the overwhelmingly dominant force in celestial motions was incomparably stronger a century after this letter to Leibniz, when Laplace was setting to work on his *Celestial Mechanics*. The extraordinary quality of evidence achieved in gravitation research over the two centuries following the *Principia* would have been far more difficult to attain if non-gravitational forces were more prominent in celestial motions.

24 To quote Alan Shapiro (2002, p. 230), "Sometimes, as in the *Optical Lectures*, the large number of experiments with slight variations to establish various points may seem tedious, but Newton attempted to leave no room for objections."

might arise from the gap between complex motions of the actual world and mathematical representations of them. Each of the items in my list of ways in which the *Principia* changed physics surfaces in a context in which explicit attention is given to this gap. So, regardless of how clearly Newton ever saw the total package formed by the items listed in Table 15.1, the mutual coherence they acquire from their forming a response to a specific evidential challenge truly is owing to him.

A second prominent physicist offered a different response to my account: "[Smith] makes very clear that Newton's celestial mechanics was something truly novel, namely that it displays the currently used method of doing mathematical physics."[25] No comment on my efforts on Newton has ever pleased me more.

Acknowledgements

This is the essay form of a talk prepared in 2001 in response to invitations from the Physics Departments of Tufts University, the University of Colorado, and the University of California at Santa Cruz. The talk was also given at the Department for History and Philosophy of Science of Indiana University and the Dibner Institute for the History of Science and Technology at MIT in 2001; as a History of Science and Technology Colloquium at the University of Minnesota and before an audience of physicists and philosophers at the University of Western Ontario in 2002; and at the New York Academy of Sciences in 2004. I received too many helpful questions and suggestions on those occasions to cite them all here.

Overlap with my subsequent much longer and philosophically more ambitious essay, "Closing the Loop: Testing Newtonian Gravity, Then and Now," had originally convinced me not to publish the essay version of the talk, especially after the editors of the present volume requested the long essay. When, at the last moment, Cambridge University Press decided that the requested essay was too long for the volume, I substituted this one.

25 Markus Fierz, in a letter to Silvan S. Schweber.