10

CLOSING THE LOOP
TESTING NEWTONIAN GRAVITY, THEN AND NOW

George E. Smith

10.1 INTRODUCTION

As of 1887, the bicentennial of the first edition of Newton's *Principia*, no scientific theory had been subjected to more extensive and more stringent testing than his theory of gravity. The intervening two centuries of gravity research, in consequence of this, had produced evidence in support of this theory of a quality, as well as an extent, far beyond anything that had emerged in any other area of scientific research. Newton's theory was *the* exemplar of science at its best, or at least at its most successful. Nevertheless, three decades after the bicentennial, Newton's theory was in the process of being replaced by the new theory of gravity in Einstein's general relativity.

The fundamental questions of philosophy of science have always concerned the nature, scope, and limits of the knowledge that can be achieved in science when it is most successful. The need to replace Newton's theory with Einstein's put such questions into an entirely new light, even to the point of prompting some to insist on shudder-quotes around the word 'knowledge' in those questions. Newton himself had expressly pointed out that his theory was open to revision in ongoing research, and hence provisional, if only because it was reached through inductive generalization. The worries about the nature, scope, and limits of scientific knowledge provoked by the switch to Einstein's theory, however, went beyond the mere non-finality of induction. For, if a theory that had been stringently tested for 200 hundred years and had by far the strongest evidence supporting it could still be overturned so abruptly, how can any theoretical claims of science amount to knowledge? Indeed, what does a stringent test of a theory really amount to, and what does strong evidence supporting it really show, if the most stringently tested, most strongly supported theory can fall so quickly?

A large fraction of the struggle of twentieth-century philosophy of science has come from the need to answer questions about the nature, scope, and limits of scientific knowledge in the light of the overthrow of Newton's theory. This essay will offer an answer to those questions, but only for the "knowledge" achieved in gravity research over the last three hundred years; and even then the central concern through most of the essay will be the "knowledge" achieved in that research before Einstein's general relativity. My goal is to spell out the precise respect in which the testing of Newton's theory was stringent in order then to reach conclusions about what the evidence for the theory did and did not show.

Evidence is a two-place relation between data and claims that reach beyond them.[1] Because data are not themselves such a relation, something beyond them is invariably needed to turn data into evidence. More often than not, I claim, the process of turning data into evidence in scientific practice requires theory, with the consequence that the richer the theory available to a field of research, the greater its effectiveness in turning data into evidence. But that need not detain us here. The point is that in order to answer questions about what specific evidence for some claim does or does not show, we must first become clear about what else beyond the cited data is involved in making them evidence for that claim. Put differently, we need to spell out the step-by-step reasoning that make the data in question evidence for that specific claim. Only then will we be in position to critically appraise this evidence. Scientists generally do not take the trouble to make the various assumptions entering their evidential reasoning fully explicit, any more than mathematicians take the trouble to list every step in a proof. Consequently, with the exception of the occasional review article, we cannot rely on the scientific literature to reveal why specific data even constitute evidence for particular claims, much less why certain data constitute much stronger evidence for these claims than other data do. This task must often fall to philosophers.

There is, of course, a standard answer in the philosophic literature to what turns data into evidence for a theoretical claim in science, namely the deduction from that claim of conclusions that can be straightforwardly compared with the data. In the specific case of concern here, the standard view is that what made celestial observations evidence for Newtonian gravity were the predictions derived from Newton's theory that agreed with those observations. In Newton's lifetime, the celestial observations of concern were planet locations relative to the fixed stars as captured to within a few minutes of arc by Kepler's orbital rules; and what made those evidence was the derivation of Kepler's rules from Newton's theory. Over the course of the eighteenth

[1] Charles Saunders Peirce would doubtlessly have insisted that evidence is a three-place relation. The relevant point here is that it is a relation involving more places than *being a datum* involves.

century, however, telescopic observations improved, so that the observations of concern became the departures of the planets and their satellites from Keplerian motion. What made those evidence for Newtonian gravitation were derivations from Newton's theory of gravity of departures from Kepler that were in close agreement with those observed. And what made this evidence so strong by the end of the nineteenth century was the increasing precision with which Newtonian theory could predict the observed locations of the planets and their satellites, to within roughly 2 or 3 arc seconds, two orders of magnitude better than predictions based on Kepler's rules could achieve. Throughout, however, what turned the data from observational astronomy into evidence for Newtonian gravity, on this view, were the derivations from Newton's theory of calculated locations of planets and their satellites that agreed with those data.

This standard hypothetico-deductive view of the evidence for Newton's theory was commonplace before the twentieth century. The shift from Newton's theory to Einstein's in combination with this view of the evidence, however, has led to a very different picture of what the evidence for Newtonian theory really showed. Although those engaged in gravity research over the two centuries following the *Principia* were scarcely aware of it, all along there was this other undiscovered theory that could yield predictions that were no less in agreement with observations than the predictions from Newton's theory. But then, all along, the evidence based on those predictions was being over-valued, for the most it was showing was that Newton's theory was one among many possible theories that could make comparably accurate predictions.

That has led at least some to argue that we are undoubtedly now also over-valuing the evidence for Einstein's theory, and for theories in science generally. Pierre Duhem, an outspoken proponent of the view that all evidence in science can only be hypothetico-deductive, had expressed such a conclusion about the epistemic status of theoretical claims independently of the Einsteinian revolution, but that revolution has lent enormous further weight to his statement of the situation:

> No doubt the physicist has the right to choose between these laws, and generally he will choose; but the motives which will guide his choice will not be of the same kind or be imposed with the same imperious necessity as those which compel him to prefer truth to error....
>
> Thus, every physical law is an approximate law. Consequently, it cannot be, for the strict logician, either true or false; any other law representing the same experiments with the same approximation may lay as just a claim as the first to the title of a true law or, to speak more precisely, of an acceptable law.[2]

[2] Duhem (1982, p. 171).

This standard view is deeply mistaken. To begin with, the evidence for Newton's theory coming out of two centuries of gravity research, as a matter of simple historical fact, was *not* really hypothetico-deductive. The main objective of this essay is to provide a very different view of the evidence that was developed over the course of the history of gravity research. This new view opens the way to a much richer account of the "knowledge" produced by gravity research not only before, but also after the Einsteinian revolution.

Newton's theory was in fact continually tested during the two centuries of gravity research, but the tests did not take the simple form of checking calculated orbital locations against observation. As Newton himself appears to have recognized, the *Principia* left gravity research no way of avoiding a more complicated logic of theory-testing than this. With testing under that more complicated logic, discrepancies between calculated and observed orbital locations did not in themselves provide evidence against Newton's theory—which is a good thing, for over the two centuries in question there was scarcely any period in which all calculated locations fell within the bounds of precision of the observations. The discrepancies instead became a source of evidence far more telling than straightforward agreement between calculation and observation can ever be. Among other things, the evidence showed that Newton's theory of gravity was much more than just one among many possible theories that could, in the manner of curve-fits, make comparably accurate predictions of orbital locations.

I am not the first to challenge the simple hypothetico-deductive view of evidence in gravity research. One of the leading proponents of that view, Carl Hempel, called attention to a difficulty as long ago as 1980 in a much too neglected paper entitled "Provisos: A Problem Concerning the Inferential Function of Scientific Theories."[3] Phrasing Hempel's point in terms of gravity research, when deriving predictions from the theory of gravity, one has to assume that no other forces are at work besides those expressly taken into account in the derivation. This proviso is in no way a part of the theory of gravity, and hence one can legitimately ask, What evidence is there for it? The obvious answer is, the close agreement of the previously derived predictions with observation. Newton himself illustrated this reasoning when he concluded in a corollary in Book 3 of the *Principia* that the fixed stars "produce no sensible effects in the region of our system."[4] But then what is really being tested when astronomers compare their orbital calculations with observations, the theory or the rather brazen claim that every important force has been identified?

[3] Hempel (1988). I thank Michael Friedman for calling my attention to this paper some time ago.

[4] In Corollary 2 of Proposition 14 in Book 3, p. 819. (All references to the *Principia* are to Newton (1999); because this is not the sole available text, I will generally indicate the source by mentioning the combination of the Book and Proposition numbers as well.)

To give a preview of this essay, its answer to what is being tested is that the primary question astronomers addressed when they compared calculations with observations is, *What, if any, further forces are at work?* The preoccupation of their research has not been with testing the theory of gravity, but with identifying further forces at work. To this end, their research presupposes the theory of gravity—or, as I prefer to express it, their research is *predicated* on the theory of gravity. The obvious question is how any test of the theory of gravity can avoid being vacuously circular in this process. The title of the essay, "Closing the Loop," refers to how the element of circularity in the actual testing of the theory of gravity has proved to yield not vacuous, but extraordinarily stringent tests of the theory. To see this, we need first to show how the way in which gravity theory has actually been tested, both then and now, has involved a less direct, more intricate logic than that of simple hypothetico-deductive testing. That will then enable us to show that the actual form of testing has delivered far more powerful evidence than is apparent under the standard hypothesis-testing construal of the tests.

I can summarize the form the evidence took here, though I question how intelligible it will be without reading on. To identify a further force is to identify a physical source giving rise to it—some detail in our solar system that makes a difference that had previously been neglected. Any such physical source is required to be robust. That is, the detail in question must make not only the difference that has shown up as a discrepancy between observation and calculation, but it should make other differences that have theretofore been neglected, differences at least some of which require confirmation. Once confirmed, the detail is incorporated into the calculations. But rather than achieving final agreement with observations, the new calculations, at least over time, yield new, usually smaller, more subtle discrepancies with observation, and the process is repeated, closing the loop. Success in identifying robust physical sources compatible with Newtonian gravity provides evidence in support both of the theory and of all the difference-making details that have been incorporated theretofore in the calculation.

Strikingly, the transition from Newtonian to Einsteinian gravity, as a matter of historical fact, left all the previously identified details of our solar system that make a difference and the differences they were recognized as making intact. In other words, the details that make a difference in our solar system and the differences they make proved more robust in the transition to Einsteinian theory than the Newtonian theory that had provided the basis for identifying them. This collection of difference-making details therefore has the strongest claim to knowledge produced by the two centuries of research predicated on Newtonian theory. But Newtonian theory also has a claim to knowledge, namely as a theory that, while holding only approximately over a restricted domain, still was adequate to establish many details that make a difference and the differences they make within that domain.

Even if intelligible, this summary fails to make clear how forceful much of the evidence actually was. Worse, it violates my whole sense of how the philosophic assessment of evidence in science needs to be done. Evidence in science lies in specifics that are often recondite. To assess the force of any one contribution to the evidence requires an analysis not only of the specifics entering into that contribution, but of those specifics in their historical context. Only in this way can one begin to see the comparative strengths of different contributions to the accumulating body of evidence. My summary of the evidence emerging out of three centuries of gravity research has glossed over all the specifics and has at best merely stick-figured the structure of the history of that evidence. Hopefully, the remainder of the essay does better than that.

The essay consists of three sections. The first explains how the *Principia* dictated a particular logic of theory-testing for the theory of gravity. The second will examine how this logic played out in gravity research ever since then, first before Einstein and then after. The final section will return to the question of what the evidence from three hundred years of gravity research has and has not shown, and with it an answer to the question of the nature, scope, and limits of the knowledge produced by that research.

10.2 THE LOGIC OF THEORY-TESTING DICTATED BY THE *PRINCIPIA*

By gravity research I mean two fields, celestial mechanics and physical geodesy, that now usually lie in the separate academic departments of astronomy and earth science. The central questions in celestial mechanics have from the beginning been, *What are the true motions—orbital, but also in some cases rotational—of the planets, their satellites, and comets, and what forces govern these motions?* Correspondingly, the central questions of physical geodesy have been, *What is the true shape of the Earth, how does gravity at its surface and surrounding it vary, and what distribution of density within the Earth produces this gravitational field?* (The questions of geodesy have also from the beginning been asked of other celestial bodies, but until the age of space-probes research concentrated on the Earth.) Laplace's five-volume *Mécanique céleste* covered both fields, and hence that title originally covered both. They migrated apart during the second half of the nineteenth century, mostly because of differences in the mathematics they employ; but they have now once again become intertwined.

The *Principia* addressed every one of the questions italicized above, and subsequent research on every one of them unfolded from what the *Principia* said. Indeed, the question that led Newton into the *Principia*—the question first put to him by Robert Hooke in 1679 and then by Edmund Halley in 1684—was, what orbital motion occurs under inverse-square central forces?

Table 10.1 Seven Comparably Accurate Ways of Calculating Planetary Orbits as of 1680—All Known to Newton

	Orbital Trajectory	*Location vs. Time*	*Mean Dist. From Sun*
KEPLER	ellipse	area rule	from observations
BOULLIAU	ellipse	a geometric construction	from observations
HORROCKS	ellipse	area rule	via 3/2 power rule
STREETE	ellipse	Boulliau's construction	via 3/2 power rule
		oscillating equant	
WING	ellipse	---------------------	from observations
		a geometric construction	
MERCATOR	ellipse	a geometric construction	from observations

That question needs to be put into historical context. At the time Newton started on the *Principia*, he personally knew of the seven different approaches to calculating planetary motion listed in Table 10.1. The main difference among these seven was the way of locating planets on their orbits versus time: Johannes Kepler and Jeremiah Horrocks following him used the area rule (planets sweep out equal areas about the Sun in equal times), while Ismaël Boulliau and Thomas Streete used a geometrical construction, Vincent Wing initially used equal angular motion around a point oscillating about the empty focus of the ellipse and then switched to his own geometric construction; and Nicholaus Mercator added still another geometric construction. The one other difference among the seven was Horrocks's proposing that the mean distance of the planets from the Sun, which for an ellipse is the same as the semi-major axis, can be more accurately inferred from their periods by means of Kepler's 3/2 power rule—the square of the periods vary as the cubes of the mean distances—than directly from observation. Streete alone followed Horrocks in this, leading Newton to appreciate the importance of the 3/2 power rule when he first learned orbital astronomy from reading Streete.

All seven of the approaches yielded more or less the same level of accuracy, within five or so minutes of arc, where the width of the Moon is 30 minutes of arc. All but Horrocks had published tables for calculating planet locations. Streete's tables were on the whole the most accurate, but none of them were entirely within the accuracy of Tycho Brahe's observations (which served as the primary basis for all the tables), and none of the approaches sufficed for the Moon.[5] As Table 10.1 indicates, the one thing on

[5] For more details about these seven approaches, see Wilson (1989).

which they all agreed was the ellipse, owing mostly to Kepler's comparative success in predicting the transit of Mercury in 1631.[6] That they should all agree on the ellipse was striking for two reasons. First, all the orbits then known were nearly circular. The most elliptical by far, Mercury's, has a minor axis only two percent shorter than its major axis; the next most elliptical, Mars's, a minor axis less than a half a percent shorter; and all the rest, a minor axis less than two-tenths of a percent shorter. Second, the ellipse was something Newton claimed that only he had established;[7] from his point of view all that Kepler and the others had shown was that the trajectories are closely approximated by ellipses. The question that Hooke and Halley put to Newton concerned the true motions, with Hooke expressly challenging the astronomers' ellipse.[8] The issue of astronomical practice lying behind the question was to find some basis for deciding which, if any, of the alternatives for calculating the motions was to be preferred.

10.2.1 Evidence in the *Principia*

Early in his work on the *Principia* Newton had concluded that this issue of astronomical practice was profoundly more difficult than anyone had realized, for none of the approaches gave the true motions. Newton had had Halley register his nine page tract called "De Motu Corporum in Gyrum"—"The Motions of Bodies in Orbit"—with the Royal Society in December 1684. He augmented this tract sometime shortly thereafter with a couple of paragraphs that did not become public until 1893.[9] In one of these paragraphs, now widely known as the "Copernican scholium," he said,

> 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. Each time a planet revolves it traces a fresh orbit, as in the motion of the Moon, and

[6] See ibid., p. 164 and van Helden (1989, pp. 109–111).

[7] In his letter of June 20, 1686 to Halley, Newton remarked, "Kepler knew the Orb to be not circular but oval & guessed it to be Elliptical" (Newton 1959-77, vol. 2, p. 436). Judging from his notebooks while reading the *Principia*, Huygens, too, credited Newton with the ellipse:

> The famous M. Newton has brushed aside all the difficulties together with the Cartesian vortices; he has shown that the planets are retained in their orbits by their gravitation toward the Sun. And that the excentrics necessarily become elliptical (Huygens 1944, p. 143; translation adapted from Koyré 1968, p. 116).

[8] In his letter of January 6, 1679–1680 to Newton, Hooke had remarked, "The Curve truly Calculated will shew the error of those many lame shifts made use of by astronomers to approach the true motions of the planets with their tables" (Newton 1959-77, vol. 2, p. 309).

[9] The paragraphs added to "De Motu" first became public in Ball (1893, pp. 54–56).

> each orbit depends on the combined motions of all the planets, not to mention the actions 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.[10]

The difficulty Newton saw here was not so much planets interacting with one another. The difficulty was that the center of gravity of the system should neither gain nor lose motion, and the problem of simultaneously solving for the motions of six planets and the Sun under that constraint Newton was seeing as intractable.

The quotation puts the challenge Newton saw in undertaking the *Principia* into a light in which it anticipates the statement by Duhem quoted above. The complexity of the true motions was always going to leave room for competing theories if only because the true motions were always going to be beyond precise description, and hence there could always be multiple theories agreeing with observation to any given level of approximation. On my reading, the *Principia* is one sustained response to this evidence problem.

Newton always claimed not to have put forward his law of gravity as a hypothesis, but to have "derived" or "deduced" it from phenomena of planetary motion.[11] At every juncture where he draws a conclusion from phenomena, however, his actual reasoning makes allowances for imprecision in the phenomena themselves, imprecision that he usually flags with the phrase "*quam proxime*." This peculiar double-superlative phrase (which occurs 139 times in the *Principia*) literally means "most nearly to the highest degree possible;" it is probably best translated "very, very nearly" or, as is more customary, "very nearly." I find it best here to stick with the Latin. Newton's "phenomena" amount to descriptions of regularities that hold at least *quam proxime* over a finite body of observations from a limited period of time. Thus, a precise statement of his Phenomenon 5 would be

> The planets swept out equal areas in equal times *quam proxime* with respect to the Sun at least over the period from the 1580s to the 1680s.

[10] Newton (1962, pp. 256 (Latin) and 281 (English)). I have altered the English translation along lines derived from Curtis Wilson.

[11] In the Preface to the first edition Newton spoke of "deriving" (*derivare*) the law from phenomena, but in the famous "*hypothesis non fingo*" paragraph of the General Scholium added at the end of the second edition he instead used the term 'deduce' (*deducere*). In the fourth Rule of Reasoning added in the third edition he used the phrase "gathered (*colligere*) from phenomena by induction." As the text will go on to show, all three of these phrasings are appropriate, pointing to different aspects of Newton's reasoning.

That Newton is not taking his phenomena to hold exactly he expressly notes in the text of Phenomenon 6, but it is also clear from the imperfect agreement between Kepler's 3/2 power rule and the mean distances and periods that he lists in the tables accompanying Phenomena 1, 2, and 4. By contrast, Newton himself never says that he is taking his phenomena to hold over only a limited period of time, but his reasoning makes much more sense if they are taken in this way. In particular, in Propositions 13 through 15 of Book 3 he invokes the law of gravity to give the area and 3/2 power rule the status of laws projecting indefinitely into the past and future; these propositions lose much of their force if the phenomena are instead taken already from the outset as projecting into the indefinite past and future. Construing his phenomena in my restricted way also has the obvious advantage of begging fewer questions.

Newton makes allowances for imprecision not only in his phenomena, but, though it has often gone unnoticed, also in the propositions he invokes to license inferences from them. Every "if-then" proposition that he uses in Book 3 to draw conclusions from phenomena he has taken the trouble to show in Book 1 still holds in the "if *quam proxime*, then *quam proxime*" form illustrated in his corollaries to Proposition 3:

> If a body sweeps out equal areas in equal times *quam proxime* with respect to some body *T*, then the force governing its motion is directed *quam proxime* toward *T*, and conversely.

Newton infers the inverse-square for the planets in two ways, first from the observed 3/2 power relation between the mean distances and the periods and then more precisely from the absence of observed precession of aphelia. In both of these cases the *quam proxime* element of the enabling "if-then" proposition is embedded in a formula he presents relating the exponent in the force rule to other variables. For example, the angle a body sweeps out from one aphelion or apogee to the next in a nearly circular orbit governed by a centripetal force is given by the formula derived in Proposition 45 of Book 1: $360/\sqrt{(3 + n)}$, where n is the exponent of r in the rule of centripetal force. This formula shows not only that the orbit is perfectly stationary if and only if n equals -2, but also that it remains very nearly stationary so long as n is very nearly -2.

Again, Newton nowhere states that he is taking the trouble to establish that his key enabling "if-then" propositions hold in *quam proxime* form, but he does in fact show that they hold in this form. This, by the way, explains why Newton, unlike modern textbooks, never inferred the inverse-square from Kepler's ellipse: he knew that the *quam proxime* form of Book 1 Proposition 11, *if a Keplerian ellipse quam proxime, then inverse-square quam proxime*, is not true.[12]

[12] See Smith (2002, pp. 31–70).

Strictly speaking, therefore, what Newton deduced from the phenomena he cites were conclusions that held only *quam proxime* over a finite set of observations, as illustrated by the following more precise statement of the proper conclusion to be drawn from the area rule:

> Therefore, the force governing the orbital motion of the planets from the 1580s to the 1680s was directed *quam proxime* toward the Sun.

This deduction was sound, but limited. The most that Newton could have truly *deduced* from phenomena was that his law of gravity held to high approximation over a particular period of time.

Newton gives Rules of Reasoning in the *Principia* for going beyond such limited conclusions. Rule 3 authorizes open-ended projections beyond the finite body of data:

> Rule 3: *Those qualities of bodies that cannot be intended and remitted and that belong to all bodies on which experiments can be made should be regarded as qualities of all bodies universally.*[13]

Rule 4, added in the third edition, authorizes a leap from approximate to exact:

> Rule 4: *In experimental philosophy, propositions gathered from phenomena by induction should be regarded as either exactly or very, very nearly true notwithstanding any contrary hypotheses, until yet other phenomena make such propositions either more exact or liable to exceptions.*[14]

This rule should be followed so that arguments based on induction may not be nullified by hypotheses.

[13] Newton (1999, p. 795). This rule was added in the second edition, with no counterpart to it in the first edition. The phrase "intended and remitted" is Medieval. A body's being hot or cold is an example of a quality that can be intended and remitted. Marilyn McCord Adams summarizes Medieval theories of intension and remission in Chapter 17 of her (1987). For an extended discussion with many helpful references, see Caroti (2004). (I thank Rega Wood for calling my attention to this and several other references on the topic.)

[14] Ibid., p. 796. Reasons for adding this rule emerged during the final stages of the preparation of the second edition when his editor Roger Cotes questioned the legitimacy of Newton's saying "by the third law of motion, every attraction is mutual" in the first corollary of Book 3, Proposition 5 (Cotes to Newton, March 18, 1713; Newton 1959–1977, vol. 5, p. 392). Cotes's challenge led Newton to a forerunner of Rule 4: "he that in experimental Philosophy would except against any of these must draw his objection from some experiment or phenomenon & not from a mere Hypothesis, if the Induction be of any force" (Newton to Cotes, March 31, 1713; ibid., p. 400).

Notice that the main verb in both of these rules is "should be regarded"—in Latin, a form of the verb *habere*, "to hold." These rules are not saying that "propositions gathered from phenomena by induction" *are* exactly or very, very nearly true universally, but that they *should be taken* to be so. Newton was no less aware that he was taking a leap from the approximate to the exact when he concluded at the end of Book 3 Proposition 8 that the law of gravity should be taken as exact than he was when he inductively generalized from the inverse-square centripetal forces toward Jupiter, Saturn, and the Earth to like forces toward Mercury, Mars, and Venus in Proposition 5.

Although his phrasing is slightly different in them, Newton's Rules 1 and 2 have this same thrust: one should admit no more causes than the phenomena require, and "therefore the causes assigned to natural effects of the same kind must be, so far as possible, the same."[15] The principle "same effect, same cause" is notoriously fallible, but that does not make it inappropriate as part of a strategy of *ongoing* research. All four rules thus give instructions for how to proceed in continuing research, instructions that authorize taking claims "gathered from phenomena" to be provisionally established, and the corresponding questions to be provisionally closed, until further phenomena give reasons to reconsider. Newton's leaps—from same effect to same cause, from a finite body of data to universal generalization, and from approximate to exact—were part of a research strategy, a research strategy that I claim was in direct response to the evidence problem posed by his recognition of the inordinate complexity of the true orbital motions. What we need to see is how that research strategy works in response to this problem. What advantage was there in taking the law of gravity to be exact universally? For that matter, what did Newton mean by "propositions gathered from phenomena by induction"?

10.2.2 "Newtonian Idealizations"

Judging from the *Principia*, and I here mean Book 2 as well, Newton imposes two demands in taking a theory derived from *quam proxime* phenomena to be exact. First, the theory must yield specific conditions under which the phenomena from which it was inferred *would* hold exactly without restriction of time. Thus, in Book 3 Proposition 13 he concludes that Kepler's area rule *would* hold exactly in the absence of forces from other orbiting bodies, and in Proposition 14, that the orbits and their aphelia *would* be perfectly stationary, instead of precessing, if there were no perturbing

[15] Ibid., pp. 794f. Before the third edition Rule 2 read, "Therefore the causes of natural effects of the same kind are the same." Newton's rephrasing for the third edition, which first appears as a hand-written correction in his annotated copy of the second edition, shows his deepening appreciation for the thrust of his rules.

forces acting on the orbiting bodies. The subjunctives here are Newton's, not mine. He not only knew, but in these propositions expressly noted, that the claims about the phenomena in question were counterfactual. They nevertheless give the *quam proxime* phenomena in question a preferred status, legitimating their role in the "deduction" of the theory. The area rule was only one of five different ways of locating planets on their orbits, but the deduced theory gives circumstances in which it and none of the others would hold exactly. For some other theory deduced from one of the alternatives to the area rule to compete, not only would it have to be deduced from this alternative, taken as a *quam proxime* phenomenon, but it would then have to yield specific circumstances in which this alternative would hold exactly.[16]

Second, the deduced theory, whether it specifies a full mechanism for effecting forces or not, must at least yield a specific configuration of bodies in which the inferred macroscopic force *would* result exactly from the composition of forces associated with their microphysical parts.[17] Thus, Newton remarks at the end of Book 3 Proposition 8 that he had doubts about whether gravity around the Earth and Sun varies exactly as the inverse-square until he had shown that it *would* do so if they were perfect spheres with spherically symmetric density. In Propositions 19 and 20 a few pages later, however, he argues that the rotation of the Earth makes it (and hence presumably the rotating Sun as well) not a perfect sphere, but a spheroid; and, as he has indicated in the second corollary to Book 1 Proposition 91, gravity does not vary simply and exactly in an inverse-square ratio with the distance to the center about such a spheroid. The text therefore justifies my use of the subjunctive in expressing the claim made at the end of Proposition 8 even though Newton does not himself use it there.

These are not the only such subjunctives that Newton deduces from the theory of gravity. Another is that the planetary orbits *would* be ellipses if no other forces were at work besides the inverse-square forces directed toward the Sun. In the Preface to

[16] Duhem (1982, pp. 190–195), in arguing that Newton could not have "deduced" the theory of gravity from phenomena, complains not only about the theory of gravity entailing that the area rule does not hold exactly, but also that the area rule itself involved idealizing assumptions that guarantee that there must be alternatives to it that describe the phenomenon in question to a comparable level of approximation. The demand that the deduced theory yield circumstances in which the area rule would hold exactly undercuts this complaint by providing means for distinguishing between descriptions of the phenomenon that have claim to being lawlike from descriptions that continue to have only the status of an artifice.

[17] In Book 2 Newton attempts to meet this requirement by laying out the physical mechanism by which the inertia of fluids acting locally on the surface of a cylinder or sphere yields a net resistance force on them. See Smith (2001, pp. 261–272). In the case of gravity, for which he of course offered no mechanism, Newton most likely viewed this compositional requirement as responding to worries about how any net macroscopic force can be claimed to have been exactly characterized in the absence of a mechanism effecting it.

the First Edition Newton speaks of deriving from phenomena the gravitational forces retaining celestial bodies in their orbits and then deducing the motions of the planets, their satellites, comets, and the seas from these forces. When only the principal forces are taken into account, and lesser forces are ignored, the deduced motions themselves should be stated in the subjunctive because of the tacit premise, "if no other forces *were* at work." Subjunctives like these can be contraposed: if the actual orbits are not perfectly stationary Keplerian ellipses, then other forces are at work; and if gravity does not vary in an exact inverse-square ratio everywhere around the Earth—as it does not—then either the Earth is not an exact sphere or its density is not spherically symmetric.

Such contraposed counterfactuals reveal the most important consequence of taking the theory of gravity to be exact: every systematic discrepancy between observation and any result deduced from the theory *ought* to stem from a physical source not taken into consideration while deducing the result—typically, either a further density variation or a further celestial force.

The table shown in Figure 10.1 is among the most striking examples of such a deduced result in the *Principia*, and also historically among the most important. For, as Newton knew and Christiaan Huygens and a few others were quick to realize, this table represents the only result deduced from the theory in the book for which *universal* gravity is indispensable, that is, inverse-square gravity between *every* particle of matter, and not just macroscopic inverse-square gravity between celestial bodies. Newton added the table to Book 3 Proposition 20 in the second edition after Huygens had challenged his theory of gravity by claiming to have empirical evidence that both the variation of gravity with latitude and the extent of the Earth's oblateness are less than Newton's theory implies.[18]

In the table, the column furthest to the right gives the variation in the length of one degree along a meridian for a non-spherical oblate Earth. The non-uniformity of these lengths is a measure of the degree of non-sphericity, and their increase from the equator toward the poles indicates that the Earth is oblate rather than prolate, that is, has a larger radius to the equator than to the poles, rather than vice-versa. The pendulum-length column gives the variation of surface gravity with latitude, expressed in terms of the length of a one-second isochronous pendulum. Huygens had developed this pendulum-measure of the strength of surface gravity in 1659. By the time Newton began work on the *Principia* the standard value for the length of a one-second pendulum in Paris was 3 Paris feet, 8.5 lines (where a line is 1/12 of an inch). Newton's

[18] See Huygens (1690, p. 152), reprinted with original page numbers listed in the text in Huygens (1944, pp. 462–471); and Schliesser and Smith (*forthcoming*).

Latitudo loci.	*Longitudo penduli.*		*Mensura gradus unius in meridiano.*
grad.	*ped.*	*lin.*	*hexapedæ.*
0	3	7,468	56637
5	3	7,482	56642
10	3	7,526	56659
15	3	7,596	56687
20	3	7,692	56724
25	3	7,812	56769
30	3	7,948	56823
35	3	8,099	56882
40	3	8,261	56945
1	3	8,294	56958
2	3	8,327	56971
3	3	8,361	56984
4	3	8,394	56997
45	3	8,428	57010
6	3	8.461	57022
7	3	8.494	57035
8	3	8.528	57048
9	3	8,561	57061
50	3	8,594	57074
55	3	8.756	57137
60	3	8,907	57196
65	3	9,044	57250
70	3	9,162	57295
75	3	9,258	57332
80	3	9,329	57360
85	3	9,372	57377
90	3	9,387	57382

FIGURE 10.1 Newton's Calculated GeodeticValues for a Uniformly Dense Earth

table offers deduced values two significant figures beyond this! The table gives what the figure of the Earth and the variation of surface gravity with latitude *would* be if the Earth were a uniformly dense body of rotating fluid matter held together in equilibrium by Newtonian universal gravity. As Newton notes in Proposition 20 in all three editions, to the extent that the actual shape and variation of gravity differ from the tabulated values, one can conclude that the density of the Earth is not uniform. In the

first edition he even proposed that, should the need arise from future measurements, a linear increase of density from the surface to the center of the Earth should be tried as the next approximation. Thus was the field of physical geodesy born.

In a letter responding to a question his editor Roger Cotes raised about this table, Newton remarked that it has the virtue of being "certain upon a supposition that the earth is uniform."[19] More correctly, the values in the table are certain to more or less the number of significant figures shown under this supposition provided that the theory of gravity holds exactly and no imprecision has been introduced during their calculation. The same point can be made about calculated orbital motion in stationary Keplerian ellipses under the supposition that no forces are at work besides the inverse-square forces directed toward the central bodies of the orbits. The calculated motions provide a first approximation for predicting the true motions. But more importantly, the observed deviations from them are telling us things about the world in the context of the theory of gravity, when it is taken to hold exactly.

Taking the theory of gravity to be exact is thus part of a research strategy that allows the complexities of the true orbital motions to become continuing evidence bearing on the theory. The first step toward seeing this is to appreciate the counterfactual status of calculations from the theory. Taking the theory to be exact allows claims about the world to be deduced that *would* hold exactly under specific circumstances. I call such claims "*Newtonian idealizations*" because they are so central to the *Principia* and its aftermath. There are many different kinds of idealizations in science. This kind, by definition, consists of *approximations that, according to theory, would hold exactly in certain specifiable circumstances*, including provisos of no further forces or density variations.

The purpose of comparing any such deduced idealization with observation is not to test the theory directly, for the calculation is presumed to be representing a counterfactual situation. The purpose is to shift the focus of ongoing research onto systematic discrepancies between the idealizations and observation, asking in a sequence of successive approximations, what further forces or density variations are affecting the actual situation? Newtonian idealizations are special precisely because the theory, taken exactly, implies that systematic discrepancies between calculation and observation must be owing *only* to factors in the world not taken into account in the calculation. Systematic discrepancies between observation and curve-fits, by contrast, can just as well arise from the mathematical framework adopted in fitting curves to observation as they can from further physical factors in the world.

[19] Newton (1959–1977, vol. 5, pp. 242–244).

The theory of gravity and deductions from it, so taken, accordingly become not so much explanations or representations of known phenomena, but instruments in ongoing research, revealing theretofore unspecified systematic discrepancies between, for example, true and idealized orbital motions. I call discrepancies of this sort "second-order phenomena" because they are not something anyone can observe. They are what you get by subtracting observations from idealized calculated results. They are second-order because they categorically presuppose the theory of gravity, taken as holding exactly. They are phenomena because they are systematic and hence constitute regularities that cannot initially be identified because more dominant regularities mask them.

Second-order phenomena provide a continuing source of evidence as the sequence of successive approximations approaches closer to the inordinately complicated true motions that Newton conceded from the outset are probably beyond exact calculation. The Newtonian idealizations of the *Principia* thereby ended up transforming the inordinate complexities of the true motions from an impediment to high quality evidence into a source of extraordinarily high quality evidence. As I have remarked elsewhere, this is one place where the cliché that the exception proves the rule actually makes sense.

10.2.3 Testing the Theory of Gravity

We are now in position to lay out the logic of theory-testing forced on subsequent gravity research by the *Principia*. The theory gets tested through its requiring that every systematic deviation from a Newtonian idealization be *physically significant*. That is, every deviation has to result from some unaccounted for density variation or force.

The unaccounted for force, moreover, does not have to be gravitational. At one point Newton appears to have been thinking that the Earth's magnetism might be producing the half of the mean rate of precession of the lunar apogee that, by his calculation, was not being produced by the Sun's gravity.[20] Suppose that the missing factor of

[20] The factor of 2 discrepancy in Newton's calculation of the mean motion of the lunar apogee is summarized in my note "The Motion of the Lunar Apsis," (Cohen 1999, pp. 257–264). In Book 3, Prop. 6, Corol. 5, Newton says, "[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" (1999, p. 810). And, in a paragraph added in the second edition to the end of the corollaries to Proposition 37, which concern the interaction of the Moon and Earth, he remarks,

> In these computations, I have not considered the magnetic attraction of the Earth, the magnitude of which is very small anyway and is unknown. But if this attraction can ever be determined—and if the measures of degrees on the meridian, and the lengths of isochronous pendulums at various parallels of latitude, and the laws of the motions of the sea, and

2 in the motion of the lunar apogee had turned out not to come from Newton's calculation failing to include higher order terms, as Alexis-Claude Clairaut subsequently discovered, but instead from the Earth's magnetism. This would have greatly complicated the history of gravity research insofar as the magnetic force of the Earth on the Moon would have had to be determined in order to calculate the Moon's motions from forces. But a conclusion that a non-gravitational force is having a notable effect on the Moon's motion in no way would have falsified Newton's theory of gravity, for that theory included no claim to the effect that the only forces affecting celestial motions are forces of gravity. Newton's calculation considered only the forces of the Sun's and the Earth's gravity, but this amounted to an idealization, not to a claim about the world. The factor of 2 discrepancy between Newton's calculation and the known mean motion of the lunar apogee, on its face, was raising a question about what further forces are affecting this motion. Failure to find any such force compatible with the theory of gravity, and only this, would have shown that Newton's theory has to be revised or replaced.

In other words, the test question forced on gravity research by the *Principia* was not whether calculation agrees with observation, but whether robust physical sources can be found for discrepancies between calculation and observation. *It is a failure of this test, and only this test, that falsifies Newton's theory of gravity.* Moreover, judging from his response to the factor of 2 discrepancy in his calculation for the lunar apogee, Newton seems to have fully recognized this—including the need that the physical source be robust and not merely *ad hoc*, for he continued to treat the magnetic force of the Earth on the Moon as unknown instead of inferring it from the missing factor of 2.

Testing of this sort can take two forms. In the "basic" form the requirement is to pin down sources of the discrepancies and confirm that they are robust and physically significant (in the context of the theory) while progressing toward smaller and smaller discrepancies between calculation and observation. In the "ramified" form, the previously discovered physical sources of second-order phenomena become incorporated into the further (still idealized) calculations in such a way that new second-order phenomena presuppose them as well as the theory of gravity; this makes subsequent testing of the theory ever more stringent by constraining the freedom to find physical sources of newly emerging, still smaller deviations. The extent to which the testing of Newtonian gravity was in fact ramified is a *historical* question. Its answer lies in the specifics of the history of gravity research.

> the Moon's parallax, together with the apparent diameters of the Sun and Moon, are ever determined more accurately from phenomena—it will then be possible to undertake all this calculation over again with a higher degree of accuracy. (1999, p. 880)

According to Book 1, Propositions 43 and 44, a superposed inverse-cubic force is just what is needed to produce a precessing orbit, and because of the dipole effect, the force around a magnet does diminish in an inverse-cube ratio of distance.

To the extent that the testing was ramified, the progressive character of the history itself becomes a form of evidence. For, a sequence of successes in pinning down robust physical sources for increasingly subtle second-order phenomena, one after another, under increasingly stringent physical constraints, is evidence that the theory amounts to more than just a curve-fit or mathematical representation that happens to agree with observation to high precision. Such a history is giving increasing reason to think that the theory of gravity really does hold of the physical world, at least to the level of accuracy of the second-order phenomena in question. For if it does not, discrepancies should be emerging for which no robust physical source can be found that is consistent with the theory.

Let me illustrate what I mean by "physically significant." The easiest example is Neptune. William Herschel discovered the seventh planet, Uranus, in 1781. As Figure 10.2[21] displays, over the next decades an anomaly emerged in its motion that by the mid-1840s had reached 120 seconds, that is, 2 minutes, of arc. The perturbations of Uranus by Saturn and Jupiter are larger than this, and hence the systematic pattern shown in the figure would have been masked had those perturbations not been taken into account first.[22] In the mid-1840s Urbain Jean Joseph Leverrier and John Couch Adams independently pursued the hypothesis of an eighth planet as the source of the anomaly, inferring from the pattern where in the sky it should be. Astronomers looked, and discovered Neptune. It is an easy to understand example because Neptune has proved to be robust in many different ways, including not only its being visible in the sky, but its also producing theretofore unnoted small perturbations in the motion of Saturn. (As we shall see, Neptune itself proved more robust than the inferred value of

[21] The figure is from Morando (1995, p. 217).

[22] For details on the determination of the residual discrepancy in the motion of Uranus, see Bouvard (1821) and Adams (1847). This example brings out a need to distinguish between two forms in which the testing logic can be ramified, a strong form and a weak form. In the strong form, further physical consequences of the identified source of a previous second-order phenomenon—that is, consequences beyond its producing that phenomenon—constrain the freedom to postulate physical sources of new second-order phenomena. In the weak form, a second-order phenomenon, and hence its physical source, masks a further second-order phenomenon. Thus, the larger perturbations of Uranus by Jupiter and Saturn would have masked the perturbations by Neptune had they not been taken into account. Because of the differences in the periods of the various perturbations, harmonic analysis of Uranus's deviations from Keplerian motion could, in principle, still have ultimately exposed the unaccounted for discrepancy shown in the figure even without first having to eliminate the perturbations by Jupiter and Saturn. Even with the weak form, however, the logic is generally still properly called "ramified," for whatever the source may be of the further discrepancy, that source will not usually be physically independent of previously determined sources. For example, Neptune also perturbs the motion of Saturn and Jupiter, and they perturb its motion.

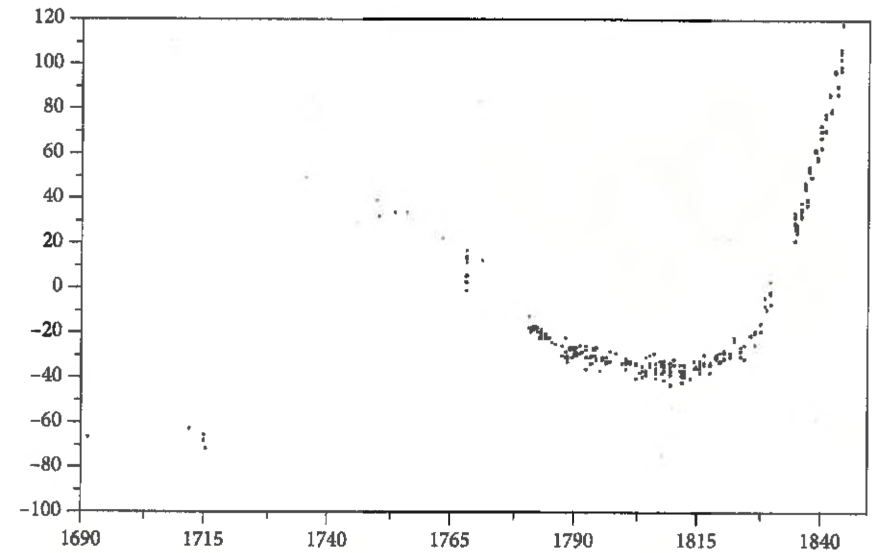

FIGURE 10.2 A Residual Discrepancy in the Motion of Uranus, in seconds of arc. *Source:* In R. Taton and C. Wilson, eds. *Planetary Astronomy from the Renaissance to the Rise of Astrophysics, Part B: The Eighteenth and Nineteenth Centuries.* Cambridge, UK: Cambridge University Press, 211–239. Reproduced by permission of Cambridge University Press.

its mass used in calculating its perturbing effects; the physical source of the anomaly in the motion of Uranus was not merely Neptune, but Neptune with a specific mass moving in a specific region of the heavens.) Though easy to understand, however, Neptune is nevertheless not a typical example of what I mean by "physically significant," for it remains the sole example of a new planet being discovered from second-order phenomena.[23]

A far more typical, and hence instructive example is the "Great Inequality" in the motions of Jupiter and Saturn. Newton had concluded that Jupiter and Saturn must be disturbing one another, and in the second edition of the *Principia* he had proposed that the dominant period of this effect would be the time between their heliocentric conjunctions, roughly nineteen years. By the early 1720s it had become clear that the effect involved a much longer time, leading Halley to represent it as a secular rather than periodic inequality in his posthumously published tables of 1749. Calculating the anomalous motions of Jupiter and Saturn had become a prize problem by the 1740s. In one prize-winning essay, Euler introduced trigonometric series to represent the perturbing effects of the planets on one another and then showed that first-order

[23] Of course, second-order phenomena have more recently been revealing "exo-planets" circumnavigating other stars; see Cole (2006, Part V).

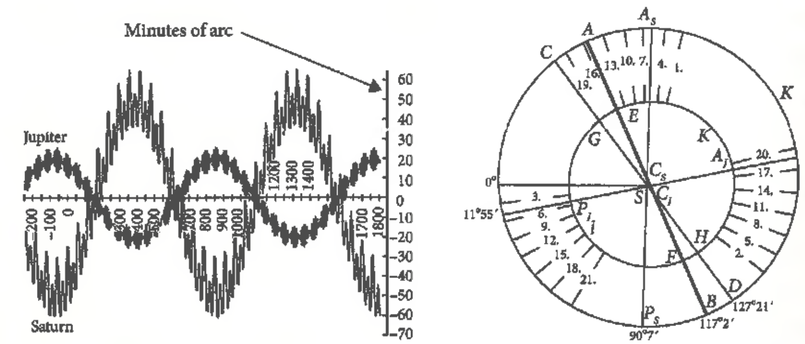

FIGURE 10.3 The Jupiter-Saturn "Great Inequality": The Three Principal Perturbations over Two Millennia and Conjunctions with Respect to the Line AB of the Orbit Centers from 1663 to 2060*Source:* Adapted from Wilson, C. (1985). "The Great Inequality of Jupiter and Saturn: From Kepler to Laplace." *Archive for History of the Exact Sciences* 33: 15–290. Reproduced by permission of Springer.

perturbations do not begin to account for the observed motions.[24] The Great Inequality itself, we need to appreciate, was far from apparent. As Figure 10.3[25] shows, not only is its period far longer than the time covered by then-available observations, but other lesser perturbations were tending to mask the long term trend over the 200 year window, starting in the 1580s, of those observations. Laplace finally found the solution in 1785 by deriving higher-order perturbations in the gravitational interaction of the two planets, revealing a nearly 900-year fluctuation in their motion that peaks at 60 minutes—twice the width of the Moon—in the case of Saturn and almost half that in the case of Jupiter.

The physical source of the Great Inequality is subtle, involving a combination of several factors only one of which I will describe here. If the two planets were in concentric circular orbits, then every time Jupiter approaches conjunction with Saturn, Jupiter would speed up and Saturn would slow down; but after conjunction, the opposite would happen, and the effect would cancel out because of the symmetry of the relative positions before and after conjunction. (That was how Newton had reasoned.) The orbits, however, are not quite circular. In Figure 10.3, the major axis of Saturn (A_s-P_s) runs from 12 o'clock to 6, the major axis of Jupiter (A_j-P_j), from 8:30 to 2:30, and the bold line (*A*-*B*) connects the centers of their elliptical orbits, C_s and C_j. The bold line hence defines the points where the two orbits become nearest to one another, *B* and *F*, and farthest from one another, *A* and *E*. (The line *C*-*D* is the line along which

[24] Leonhard Euler, *Recherches sur la Question des Inégalités du Mouvement de Saturne et de Jupiter*, of 1748, which won the prize in the 1748 Royal Academy of Sciences competition on the topic. The specifics of Euler's first-order calculation are summarized in my note "The Interaction of Jupiter and Saturn" (Cohen 1999, pp. 211–217), in particular, pp. 213ff.

[25] The figure is adapted from Wilson (1985, pp. 29, 35).

the respective planes of the two orbits intersect, adding a further complication I am ignoring here.)

In any conjunction to the right of the bold line, Jupiter is farther away from Saturn after conjunction than before, and hence its force on Saturn is not quite balanced out after conjunction. The opposite is true to the left of the bold line. Conjunctions occur every 19-plus years, with Saturn progressing a little more than two-thirds around its orbit between them. As a consequence, two out of every three conjunctions take place on one side of the bold line, resulting in a net perturbation of both planets over the course of every three conjunctions, that is, every 59 or so years. It takes more or less 450 years of conjunctions before the two out of every three switch to the opposite side of the bold line, and the effect reverses itself. The physical source of the Great Inequality, therefore, involves the respective periods of the two planets and the relative locations of the centers of their orbits, and hence their eccentricities and the angle between their major axes. If, for example, the centers of their orbits were nearer one another, this inequality would be smaller, while if the eccentricities were larger, it would be all the greater.[26]

Notice how I am running counterfactuals off of this physical source. To pin down the physical source of a second-order phenomenon is to identify factors that make a difference and to relate them to other parameters in generalizations that support counterfactuals. Also notice that the discovery of the physical source in this case did not involve a discovery of something new in the world, as it did in the case of Neptune. The angle between the major axes of the two planets and the eccentricities of their orbits were known beforehand—indeed, known to high approximation by Ptolemy. What was discovered was that certain comparatively subtle physical details were making a difference that theretofore had gone unnoticed.

That is how this example is more typical than that of Neptune: what is generally discovered in pinning down the source of a discrepancy is that *certain details among the indefinitely many physical details in the world are making a particular detectable difference that theretofore those details had not been recognized as making.* Showing that

[26] This geometric account of the "Great Inequality" is adapted from ibid., especially pp. 28–33. Wilson's account in turn derives from Airy's account in (1834, pp. 143–155). My account represents a gross simplification—for purposes of illustration—of a far more complicated situation. First, of all, the two orbits are inclined with respect to one another, with the line *CGHD* defining the intersection of the two planes. Hence, there is another asymmetry in each conjunction. Second, as the text goes on to indicate, the forces of Jupiter and Saturn on one another cause perturbations of the eccentricities, inclinations, and location of the perihelia of the two orbits. These changes in orbital elements contribute decisively to the Great Inequality, which is precisely why higher-order perturbational terms have to be taken into account in order to define it. A more complete specification of the physical details contributing to just this one inequality would have taken several pages.

the source is robust usually centers on verifying still further differences that theory entails those details should be making, as illustrated by Neptune's theretofore unnoted perturbations of Saturn.

Those engaged in the "normal science" of gravity research after Newton were not merely testing his theory or tightening the bounds of accuracy to which it was known to hold. In their effort to represent the true motions ever more closely in a sequence of successive approximations, they were discovering which physical details make a difference in those motions and what differences those details make. To those outside the research, this may have seemed mostly just a mop-up process of the sort Thomas Kuhn spoke of in his description of normal science.[27] From the perspective of those engaged in the research, however, the theory was an instrument allowing them to pursue an increasingly deep understanding of the actual motions and the factors in the world to which these motions are sensitive.

The Great Inequality brings out a final point that needs to be made before we turn to how the theory-testing logic played out. I said above that the physical sources of discrepancies between observation and calculations based on the theory of gravity consist of theretofore unrecognized density variations or forces, gravitational or otherwise. I am now saying that the physical source of the Great Inequality involves such geometric details as the eccentricities of their orbits. How do the eccentricities amount to forces? The answer, which lies in the method of perturbations, should prove helpful when we examine the history of gravity research in the next section.

The gravitational force of Jupiter on Saturn depends not only on the mass of Jupiter, but also on (the inverse-square of) the varying distance between them. In the case of a first-order perturbation, the perturbing force is calculated under the assumption that both orbits are Keplerian ellipses, with the calculation giving the deviations from these ellipses resulting from this force. These changes to the orbits, however, entail that this force is only a first approximation to the actual force of Jupiter on Saturn. The whole point of higher-order perturbations is to incorporate changes in the perturbing force that result from the difference between the initially assumed nominal orbits and the perturbed orbits, including their perturbed eccentricities. First-order perturbing forces associated with the nominal eccentricities of the Jupiter and Saturn orbits had been known to have only small effects ever since Euler's calculation of these orbits in the 1740s. Laplace's discovery four decades later was that higher-order perturbations defined in terms of the nominal Keplerian eccentricities have a spectacularly larger effect. The difference the eccentricities of Jupiter and Saturn make to the Great Inequality lies in the further contributions to the force between them that arise

[27] "Mopping-up operations are what engage most scientists throughout their careers" (Kuhn 1970, p. 24).

from departures of these planets from Keplerian motion. The physical source of the Great Inequality discovered by Laplace did indeed therefore amount to configurational details contributing to the force between Jupiter and Saturn, specifically a contribution arising from perturbational consequences of the nominal eccentricities, but still a force, and one unaccounted for in Euler's calculation.

This point holds generally. The details that make a difference to orbits amount almost always to details that in Newtonian theory affect the net forces acting on orbiting bodies. Such details can include incremental corrections to the various attracting masses employed in the calculations, or for that matter, to the values of the elements defining the nominal, unperturbed orbits.

10.3 THREE CENTURIES OF GRAVITY RESEARCH

My claim, then, is that the primary aim in comparing calculation with observation in gravity research is to discover which details of the physical world in point of fact make a difference and what differences they make, both in orbital motions and in gravitational fields surrounding the Earth and other central bodies. This aim is achieved by exposing systematic discrepancies between calculation and observation and then establishing their physical sources. Evidence accrues to the theory of gravity through success in establishing those sources. An inability to find a robust physical source for a systematic discrepancy between calculation and observation is the sole form of evidence giving grounds for modifying or replacing the theory. In the case of Newton's theory, this meant inability to find and confirm a density variation or a force associated with some configurational detail. That this has been the logic of theory-testing in gravity research seems glaringly obvious once one looks carefully at all that research and its history. Why, then, has this logic not been apparent to everyone all along?

It did become apparent to those working on orbits after the discovery of Neptune, if only because of the renown gained by Leverrier and others who were involved in that discovery. In particular, it became apparent that any discrepancy between calculated and observed orbital motion should be taken *prima facie* not as challenging the theory of gravity, but as telling us something about the world we did not know before. Correspondingly, the payoff from comparing calculated motions with observation came not from confirming Newton's theory once again, but from exposing new second-order phenomena from which we could learn new things about the world.

Even then, however, the primary practical purpose to which the tables of planetary motion prepared by Leverrier from the 1850s through the 1870s and by Simon Newcomb and his colleagues over the last two decades of that century was to allow easy prediction of planet locations. This had always been the point of such tables, even when the "predictions" were astrological retrodictions. In his own mind the most important goal of Newcomb's effort was to identify "discordances" between calculation and observation, but the work was carried out under the auspices of the Naval Observatory, and the Navy were looking for the most accurate predictions that they could obtain. Not surprisingly, therefore, those outside the research field tended to see only the goal of prediction, and this, together with the commonplace view that scientific hypotheses are tested by deriving predictions from them, gave reason to think that Newton's theory was being tested directly by these predictions rather than indirectly through the pursuit of physical sources for the discordances.

Book 3 of the *Principia* had offered several examples of using the theory of gravity to discover new things about the world. Those who picked up the theory in the quarter century after Newton died, however, were far more concerned with verifying it in the face of intense Continental opposition to it that had been voiced by Huygens, Leibniz, and Johann Bernoulli. Maupertuis, Clairaut, Euler, and d'Alembert viewed Newton's theory as a hypothesis to be tested by deriving observable consequences from it. Clairaut's initial reaction to being unable to derive more than half of the observed mean motion of the lunar apogee from the theory, for example, was to add a $1/r^4$ term to the law of gravity.[28]

The situation did change following Clairaut's discovery that higher-order perturbational terms yield the motion of the lunar apogee. We find Euler thus remarking in the introduction to his second essay on Jupiter and Saturn, submitted to the Royal Academy's prize competition of 1752:

> ... 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.... One can now maintain boldly that the two planets Jupiter and Saturn attract each other mutually in the inverse ratio of the squares of their distance, and that all irregularities that can be discovered in their movement are infallibly caused by this mutual action.... And if the calculations that one claims to have drawn from this theory are not found to be in good agreement with the observations, one will be always justified in doubting the correctness of the calculations, rather than the truth of the theory.[29]

Notice, however, that Euler's focus is not on what discrepancies might be telling us about the world, but on what they are telling us about how to calculate orbital motions. Euler does not single out the distinctive way in which the eccentricity of the lunar

[28] See Waff (1995, p. 39) and Wilson (2002, pp. 213–215).
[29] Euler (1769, pp. 4f), translation from Wilson (1980, p. 144).

orbit affects the precession of the apogee that was revealed by Clairaut's higher-order terms. The first place I know of where the focus is instead clearly on what a discrepancy is telling us about the world is George Bidell Airy's discussion of Laplace's success with Jupiter and Saturn in his *Gravitation* of 1834; and even there what Airy is doing is helping the uninitiated to understand the qualitative physics captured in Laplace's recondite perturbational analysis.

One can find cases before Neptune in which the focus was on the physical sources of second-order phenomena. Among the earliest was d'Alembert in 1749 seizing on the recently discovered 18 year nutation of the Earth to infer the mass of the Moon, a quantity with which Newton had struggled. Two other examples that spring to mind are the effort to establish the variation of density below the surface of the Earth from measurements of surface gravity and the extent of the Earth's oblateness during the later part of the eighteenth century, and Clairaut's introduction of the gravitational effects of Jupiter and Saturn to compensate for the small irregularity in the 75-year average period in his calculation predicting the return of what came to be known as Halley's comet. Furthermore, not only was the success of such calculations being celebrated as evidence supporting Newton's theory; failures were being taken not as reasons to reject Newton's theory out of hand, but as sources of insight about the world. For example, one extreme suggestion in response to the anomaly in the motion of Uranus before the discovery of Neptune was that perhaps the inverse-square relation of gravity ceases to hold to the same exactitude beyond Saturn.[30] Even so, the idea that such anomalies are something that gravity research ought to be pursuing did not come fully to the forefront in the literature until after the discovery of Neptune.

The key to determining the logic of theory-testing is always to ask what sorts of test results were needed in actual practice to falsify the theory. In gravity research from the *Principia* onward that has required discrepancies for which no robust physical source can be found that is consistent with the theory. Discrepancies by themselves did not falsify the theory. Indeed, in the three centuries following the publication of Newton's *Principia* there has been only one brief period in which there was no widely recognized discrepancy between theory and observation, namely from the time after Laplace published his *Théorie de Jupiter et Saturne* and *Sur l'equatione séculaire de la lune* in 1788 to the emergence of the anomaly in the motion of Uranus three decades later. Only one of these widely recognized discrepancies, the anomaly in the precession of the perihelion of Mercury that emerged in the second half of the nineteenth century, falsified the theory of gravity presupposed in the discrepancy.

[30] See Morando (1995, p. 216).

One cannot, however, simply turn to publications of those engaged in gravity research for clear statements that discrepancies by themselves do not falsify the theory. That is not because the individuals in question did not understand what they were doing, for their practice conformed with this logic from at least the 1740s on. They just did not take the trouble of expressly analyzing and announcing the logic under which the theory of gravity was being tested. As a result, careful philosophical analysis of their evidential reasoning, ideally case by case, over the course of the history of the research, and only that, can reveal the logic. Needless to say, going case by case through the many episodes of the last three centuries of gravity research is a project beyond any one person, and hence beyond the scope of this essay. What follows instead are some summary remarks about the three hundred years of the research together with closer examination of a few cases—in particular, a highly instructive success that established the fluctuating rotation of the Earth in Section 10.3.2 and the notable failure with the perihelion of Mercury in Section 10.3.3.

10.3.1 Complications: The First Two Hundred Years

The failure of those engaged in gravity research to take the trouble to make the logic of their theory-testing fully explicit was not the only reason why that logic was not all along more apparent. Complications intrinsic to the research probably contributed even more to masking the theory-testing logic. Seen in retrospect, such complications fall fairly clearly into the three categories in which I will treat them here. As seen at the time, however, each discrepancy between observation and calculation tended to pose its own distinctive complications. The history of gravity research consequently became a history of finding ways to deal with the complications, discrepancy by discrepancy, for they, and not the subtleties of testing the theory of gravity, dominated day-to-day activity in the field. This focus on the challenges posed by individual discrepancies makes the big picture difficult to see.

One important complication was that discrepancies between observation and Newtonian idealizations almost always underdetermine their physical source. For example, the anomaly in the motion of Uranus shown earlier in Figure 10.2 was not enough to determine both the mass of the hypothesized eighth planet and its distance from Uranus. As a result, the comparatively elliptical orbits that Leverrier and Adams independently proposed, with their eccentricities in excess of 0.1, were both far off the nearly perfect circular orbit Neptune turned out to have once its orbital elements were obtained by using earlier observations in which it was mistaken for a star; and the values of mass proposed by Leverrier and Adams were much larger than the mass obtained once Neptune's satellite Triton was discovered in 1846. The second-order phenomenon displayed in Figure 10.2 was sufficient to locate Neptune mostly because of

the timing of the observations, which happened to bracket the conjunction of Uranus and Neptune in 1821.[31] That the source of this second-order phenomenon was a further planet was made clear more by its appearance as a disk once sighted telescopically than it was by the reasoning of Leverrier and Adams.

Underdetermination of the physical source was the dominant impediment to research in physical geodesy before the twentieth century. As Georg Kreisel finally proved rigorously in 1948, the deviations of surface gravity and the figure of the Earth from the table of Newton's shown earlier in Figure 10.1 are not enough to determine the variation of density inside the Earth.[32] The deviation of surface gravity from Newton's ideal variation does imply the value of $(C\text{-}A)/Ma^2$ for a spheroidal Earth and hence a correction to the difference $(C\text{-}A)$ in the Earth's moments of inertia implied by Newton's table, given the mass of the Earth M and its equatorial radius a. The lunar-solar precession similarly determines the ratio C/A of the two moments of inertia, and hence a correction to the polar moment of inertia implied by Newton's table.[33] But many different density distributions inside the Earth can result in the same values for the Earth's moments of inertia and mass. As a consequence, from Newton until the twentieth century, hypotheses were put forward for the (spheroidally symmetric) density variation inside the Earth and tested against one another as best one could, but without any of them coming remotely close to being established.

This example is typical. When the second-order phenomenon, together with other established constraints, is not sufficient to determine its physical source, hypotheses and hypothesis testing become the norm in pursuing the source. This reinforces any tendency to think that the underlying logic with which the theory of gravity is being tested consists of direct hypothesis testing as well.

Over its first two hundred years physical geodesy faced the difficulty of finding some way or other to verify any hypothesized variation of density in the interior of the Earth. The only known observable consequences of any such variation were the very second-order phenomena to which the hypothesis was responding, that is, deviations

[31] For details, see ibid., pp. 216–222.

[32] The insufficiency of deviations from Newton's table to establish the density variation inside the Earth was recognized at least as early as Laplace's work on the subject, a summary of which covers a large fraction of the second volume of his *Mécanique céleste*. Kreisel's proof is in his (1949).

[33] The polar moment and the difference between the moments, together with Cavendish's 1798 measurement of the mean density of the Earth, were enough to show that the density is (almost certainly) greater at the center, but this still left the specific distribution open. Thus we see Laplace remarking at the end of the second volume of *Mécanique céleste* (1799):

> *The same phenomena indicate also a decrease in the densities of the strata of the terrestrial spheroid, from the centre to the surface,* without giving the precise law of the variation of the density. . . .

from Newton's table. Of these deviations, only surface gravity was amenable to precise measurement, and it was sensitive to localized variations in density near the surface, in what we now call the Earth's crust. As a result, underdetermination by the primary consequences of density variations was not the only factor standing in the way of verifying any hypothesis; its lack of further observable consequences left it *ad hoc*, and hence anything but robust. A struggle to marshall evidence for hypothesized density variations consequently became the hallmark of physical geodesy until the twentieth century. Anyone engaged in the research, or reading original sources from the period, would have had trouble seeing any form of evidence in physical geodesy except everyday hypothesis testing. Real progress toward a robust density variation, in the form of at least the beginnings of a sequence of successive approximations, would have been needed for the underlying logic with which the theory of gravity was being tested to have become apparent.[34]

> Thus every phenomenon, depending on the figure of the earth, throws light upon the nature of the magnitude of its radius; and we see that all these results agree with each other.
>
> These observations are not, however, sufficient to make known the interior constitution of the earth; but they indicate the most probable hypothesis of a density decreasing from the centre to the surface. Universal gravitation is therefore the true cause of all these phenomena; and if its effects are not so precisely verified in this case, as in the motion of the planets, it arises from the circumstance, that the inequalities of the attractive forces of the planets, depending on the small irregularities in their surfaces, and in their internal parts, disappear at great distances; so that we only perceive the simple phenomenon of the mutual attractions of these bodies toward their centres of gravity. (Laplace 1966, vol. 2, p. 931; original italics)

Laplace himself proposed a density variation in Volume 5 of *Mécanique céleste* (1825), with a ratio of 2.8 to 11 of the density at the surface to the density at the center. See Jeffreys (1924, p. 197).

A second point worth noting is that the measurement of the quantity *(C-A)/Ma2* has a more complicated history than the text suggests. Since the advent of artificial satellites about the Earth, it has been measured, to higher and higher accuracy, by the motion of the nodes of the satellite orbits. Before artificial satellites, it was measured by a residual inequality in the motion of the lunar nodes—that is, a discrepancy remaining after the effects of the Sun (and planets) on the motion of the lunar nodes is taken into account. But the theory of the Moon was too imprecise to permit this approach until after the Hill-Brown theory was developed, culminating in Brown's tables and substantial further work by Brown, de Sitter, and others. For more details see the sequence of papers on the subject by Harold Jeffreys (1937, 1941, 1948). Before that, efforts were made to derive the value for this parameter from surface gravity measurements, such as by Bowie (1917), as discussed in Jeffreys (1924, p. 189).

[34] Physical geodesy did offer a way of testing Newtonian universal gravity more directly, namely through Clairaut's equation, which expresses a relationship between the variation of surface gravity with latitude and the extent of flattening of the rotating Earth. This relationship is entailed by universal gravity under assumptions about the symmetry of the density distribution within the Earth. "Gravity anomalies" caused by localized concentrations of mass near the Earth's surface, however, stood in the way of confirming that Clairaut's equation holds for the Earth until well into the twentieth century. All this is discussed in a companion piece to the present essay (Smith, *forthcoming*).

All this changed during the twentieth century when the problem of the density distribution inside the Earth was finally solved by resorting to an entirely new source of data, seismic waves generated by earthquakes. This is one of the great success stories in how to deal with underdetermination in science. It is, however, more a story of research in seismology than one of gravity research. The sequence of successive approximations to the distribution of density within the Earth, while constrained by measurements from physical geodesy, was driven by seismological data. Extracting evidence from those data involved problems of a different kind from any encountered in celestial gravity research. I am therefore going to postpone further discussion of the evidence problem in determining how density varies inside the Earth until another essay,[35] concentrating instead through the rest of this essay on evidence from orbital phenomena.

The fact that second-order phenomena generally underdetermine their physical sources was one complication that obscured the logic of theory-testing in gravity research. Another complication came from the need to confirm the robustness of proposed physical sources. The demand here, as noted above, is for evidence for the proposed source beyond the second-order phenomenon itself, for otherwise the proposed source may be nothing more than an *ad hoc* way of allowing the theory of gravity to accommodate that second-order phenomenon. Simply put, the source must make differences beyond the second-order phenomenon in question. As we shall see in the discussion of the perihelion of Mercury in Section 10.3.3, a failure to find such further consequences of a proposed source leaves it at best a conjecture. Identifying further consequences that can actually be checked, however, can turn into a substantial research project in its own right, if only because what the consequences often consist of are some still further second-order phenomena that no one has yet observed. The challenge of finding accessible consequences has led researchers in celestial mechanics to resort to several different kinds of evidence at one time or another—that is, in the words I am using here, a variety of forms of logic in what amounts to testing the

[35] Success in determining the density distribution within the Earth did not come all that easily even with seismology. Seismic waves propagating through the Earth were still not enough to determine this distribution without imposing some constraints beyond those from gravity measurements. A breakthrough came when improvements in instrumentation allowed the resonant standing waves of vibration of the whole Earth to be determined following large earthquakes; these standing waves provided more direct information about how the density varies with depth. For details, see my essay, "Gaining Access: Using Seismology to Probe the Earth's Insides," another companion piece to the present essay, or my lecture of the same title in Smith (2007).

proposed sources. This variety, together with the justly celebrated ingenuity that the specifics of the tests have required in individual cases, has then tended to obscure the logic with which the theory of gravity itself was being tested in the process.

Let me illustrate the many different kinds of evidence for the robustness of proposed sources with some examples. Straightforward hypothetico-deductive evidence has been the most common form. Two examples already came up in Section 10.2.3: the theretofore unnoticed effect of Neptune on the motion of Saturn and such predictions from Laplace's solution for the "Great Inequality" as what the date and conjunction locations are going to be when the anomalous longitudinal deviations in the motions of Jupiter and Saturn next reach their maximum and begin to decline.

Examples of a slightly different form of evidence, in this case for the robustness of the mass of the Moon that d'Alembert inferred from the 18-year nutational wobble of the Earth, were his success in then deriving the combined 26,000-year lunar-solar precession of the equinoxes and Laplace's subsequent successful theoretical calculations of the tides some twenty years later. These are examples of the logic Clark Glymour labeled "bootstrapping."[36] Another example of the same sort comes from the mass of Venus, which until space flight was inferred primarily from an inequality in the motion of Mars and then supported by the full range of gravitational effects of Venus on Mercury, Earth, and Mars implied by this mass.

Evidence for robustness has also taken forms not so widely discussed in the philosophical literature. In Section 10.3.2, I will describe how the fluctuation in the rotation of the Earth, initially proposed in response to a discrepancy between observation and the Hill-Brown theory of the Moon, was finally established through corresponding subtle discrepancies between observation and theory for Mercury, Venus, and the Earth. This is an example of the "common-origin" evidence explored by Michel Janssen.[37] The stability of measured masses and moments of inertia, and the extent to which different ways of measuring these converge toward the same value, exemplify a form of evidence that William Harper has contrasted with hypothetico-deductive in several places.[38] Newcomb reverted to this form of evidence for the masses of the two inner planets after he discovered that Leverrier had not employed the same consistent masses for them throughout his theories of the orbits of different planets.

Finally, evidence for the robustness of the effects of Jupiter and Saturn on Uranus was provided, even before the discovery of Neptune, by the sharply defined anomaly in the latter's motion shown in Figure 10.2, an anomaly that would have been masked had these effects not been included in the calculations. This is an example of a form of

[36] Glymour (1980).
[37] Janssen (2002).
[38] For example, Harper (2002).

evidence I have discussed elsewhere, calling it evidence accruing to a claim from the success of research predicated on it.[39] Ironically, of all these examples, this last is the one most inclined to go unnoticed; yet, as we shall see in Section 10.3.2, it is the one most closely related to the underlying logic with which the theory of gravity was being tested throughout the two centuries following Newton.

So, gravity research has involved many different kinds of evidence, including hypothesis-testing of one form or another, in addition to the logic of theory-testing that has been at its core. This is true both for the initial evidence establishing physical sources of second-order phenomena and of the further, usually subsequent, evidence providing the basis for assessing their robustness. With so many different forms of evidence coming into play in the day-to-day pursuits of those engaged in gravity research, it is scarcely surprising that the logic entering into testing of the theory of gravity itself over the history of the research stands out only when one steps back and views that history from a more comprehensive perspective.

In addition to the two complications discussed so far is still another that time and again has consumed huge amounts of effort in orbital research: difficulties in isolating systematic discrepancies—that is, second-order phenomena—and making them precise. A quotation from G. W. Hill explaining the need for new work on the Moon in 1877, when he was just starting to devise his new approach, highlights one source of such difficulty:

> The rate of motion of the lunar perigee is capable of being determined from observation with about a thirteenth of the precision of the rate of the mean motion in longitude. Hence if we suppose that the mean motion of the moon, in the century and a quarter which has elapsed since Bradley began to observe, is known within 3", it follows that the motion of the perigee can be got to within about 500,000th of the whole. None of the values hitherto computed from theory agrees as closely as this with the value derived from observation. The question then arises whether the discrepancy should be attributed to the fault of not having carried the approximation far enough, or is indicative of forces acting on the moon which have not yet been considered.[40]

From Newton forward the rate of precession of the Moon's apogee, or perigee, has been the most sensitive measure we have of the exponent -2 in the law of gravity. The lingering discrepancy noted by Hill therefore potentially raised questions beyond the one he posed that we will be returning to later. The point here lies with the question he did pose. I have been talking all along about exact Newtonian idealizations. Owing to the

[39] Smith (2001a).
[40] Hill (1905).

mathematical difficulties of the three-body problem, however, calculations of orbits have always been only approximations to such idealizations.[41]

Consequently, discrepancies between calculation and observation can arise from the necessity of employing inexact mathematics as well as from unaccounted for forces. Much of the most famous effort in the history of celestial mechanics has focused on finding methods of framing the perturbation equations and then carrying out calculations that eliminate, or at least limit, the purely mathematical sources of discrepancies. The lesson was first learned from Clairaut, as Euler noted in the passage quoted at the beginning of Section 10.3. Hill's question about inadequate approximation versus unaccounted for forces has had to be asked every time a second-order phenomenon emerges.

Mathematical imprecision is not the only source of difficulty in isolating systematic discrepancies and making them precise. The calculations also require numerical values obtained from observation, which too of course is never perfectly exact. An example will help here. Over the course of twenty years between the quotation from Hill and Simon Newcomb's *The Elements of the Four Inner Planets and the Fundamental Constants of Astronomy* of 1895, a team of people put a monumental amount of effort into achieving the highest standard of exactness for the four inner planets that they could. The "fundamental constants" referred to in Newcomb's title include such things as the speed of light, the aberration-of-light constant, and the obliquity of the ecliptic. These constants enter into corrections that have to be made to the raw data provided by terrestrial observations of the planets. The values of the constants all have to be derived from observation. So too do the masses of both the four inner and the four outer planets, either from the orbits of their satellites or from some isolable perturbational effect each has on one or more of the other planets. Any inaccuracies in the fundamental constants and the masses ultimately limit the meaningful precision of comparisons between calculated and observed planet locations.

The values of the Keplerian elements of the orbits used in perturbational calculations also have to be inferred from observation. Newcomb employed 62,000 meridional observations of the four inner planets from thirteen observatories stretching back to 1750, plus observations of Mercury and Venus transiting the Sun all the way back to 1631. He derived his values of the elements of the four inner orbits by starting from Leverrier's values, calculating deviations from the observations with these values, and

[41] That the calculations have yielded only approximations to Newtonian idealizations is not merely because of the method of perturbations. Direct numerical integration of the equations of motion is still only an approximation to an exact solution, owing to round-off and truncation error and discretization of continuous motions. Numerical integration can yield only a better approximation.

then using weighted least squares to obtain incremental corrections to reach his final values. Only then was he in a position to use gravity theory to derive the values for all of the perturbational effects on the four planets. Newcomb included terms to the same order in the reciprocal masses as Leverrier had, and, as we learned from Hill, these calculations often involved truncated summation of infinite series, not all of which were guaranteed to converge rapidly.

Newcomb's announced goal was to achieve the first "self-consistent" theories that took into account all available observations of assured quality not only for the orbits of the four inner planets, but in separate endeavors working with Hill, for those of the four outer planets as well. Not only did he succeed in this, but the theories of the eight orbits generated by his and Hill's efforts remained the official basis for planet calculation everywhere except Leverrier's France until 1984, when advanced computers allowed the perturbational approach to be replaced by numerical integration of the equations of motion.

Newcomb's personal goal in all this effort was to pin down residual discrepancies—what he called "discordances"—between theory and observation. His 1895 monograph singles out four such discordances in the motions of the inner planets:

1. A discrepancy in the motion of the perihelion of Mercury that was 29 times greater than the probable error associated with uncertainties in the observations and the values of the masses and fundamental constants.
2. A discrepancy in the motion of the nodes of Venus that was 5 times greater than the probable error.
3. A discrepancy in the motion of the perihelion of Mars that was 3 times greater than the probable error.
4. A discrepancy in the variation in Mercury's eccentricity that was 2 times greater than the probable error.

The Mercury perihelion will be considered in Section 10.3.3 below. As we shall see then, Newcomb's value for the discrepancy in the perihelion of Mars was way too large. The most striking case, however, was the discordance in the nodes of Venus. At least one physicist of note, Harold Jeffreys, questioned general relativity during the late 1910s because it offered nothing to account for the systematic discrepancy in Venus's nodes.[42] When R. L. Duncombe redid the motion of Venus in the early 1950s, this discordance simply disappeared, for reasons he could not find.[43]

[42] See Jeffreys (1916, 1918).

[43] Duncombe (1958), especially p. 44 where he proposes that Newcomb's discordance was "probably attributable to large systematic errors in the older observations," and p. 25 where he raises the possibility that Newcomb neglected a term in his calculations.

An even more striking example of a second-order phenomenon simply disappearing occurred in 1993. For a long time there had been systematic residual discrepancies in the motion of Uranus. During the second half of the twentieth century this had spawned research looking for a tenth planet. Myles Standish, the person who is now in charge of the official calculation of planetary motions at the Jet Propulsion Laboratory in Pasadena, California, recalculated Uranus following the Pioneer spacecraft fly-by of Neptune, which provided a more accurate mass for the combination of Neptune and its satellites. He found, to quote his abstract, "the alleged 'unexplained anomalies in the motion of Uranus' disappear when one properly accounts for the correct value of the mass of Neptune and properly adjusts the orbit of Uranus for the observational data."[44] His article ends with the point I want to emphasize:

> Many professional lives have been dedicated to the long series of meridian circle (transit) observations of the stars and planets throughout the past three centuries. These observations represent some of the most accurate scientific measurements in existence before the advent of electronics. The numerous successes arising from these instruments are certainly most impressive. However, as with all measurements, there is a limit to the accuracy beyond which one cannot expect to extract valid information. There are many cases where that limit has been exceeded; Planet X has surely been such a case.[45]

From the point of view of those engaged in gravity research, Standish showed that a seeming second-order phenomenon was spurious. It was not entirely spurious, however, for its source involved physically meaningful incremental revisions of the mass of Neptune and the orbital elements of Uranus. Table 10.2 lists sources of systematic discrepancies. Discrepancies arising from the first three sources on the observation side and the first and fourth on the calculation side are more truly spurious insofar as identifying their sources reveals nothing new about the world. Inadequate corrections for such sources of systematic observational error as atmospheric refraction and the need to transform observations made on the moving Earth to the frame of the fixed stars are like Standish's inadequately precise values of the mass and orbital elements. Those in the field of orbital research view second-order phenomena arising from them as spurious, yet we do learn something further about the world from identifying their source. The same is true of inadequately precise values of the fundamental constants like the speed of light and the aberration constant that enter into the corrections for known sources of systematic error. The fourth and fifth sources on the left of the table

[44] Standish (1993, p. 2000).
[45] Ibid., p. 2005.

and the second and third on the right have accordingly been italicized to mark them off from the others. They all lead to second-order phenomena that are spurious from one point of view, but not from another.

There is, of course, always the possibility of some yet unidentified source of systematic observation error that needs to be corrected for. As we shall see in a moment, those in the field would not regard second-order phenomena arising from it as spurious, for these phenomena would be telling us something new and important about the world, albeit not necessarily about gravity or other forces in the way that second-order phenomena associated with the last three sources on the right of the table do. At first glance, one might take second-order phenomena arising from unaccounted for higher-order perturbational terms to belong with failure to carry out infinite series solutions far enough, but as we have seen in the case of the Great Inequality and will see again in the next section, ignoring these higher-order terms amounts to ignoring physical details that are giving rise to further forces affecting orbital motions. Second-order phenomena arising from the fifth as well as the sixth sources on the right thus lead to new discoveries about Newtonian forces acting in the world, while the last one on the right, to a discovery about gravity that, to use Newton's words, makes the theory of it "more exact or liable to exceptions."

The first two centuries of gravity research reduced the magnitude of discrepancies between observed and calculated planet locations by more than two orders of magnitude, down to not much more than a second or two of arc in the case of the planetary theories of Newcomb and Hill. The many sources of discrepancies listed in Table 10.2 underscore how extraordinary an achievement this was. More accurate corrected observations contributed to close agreement, but even more so the accumulation of more and more observations over time, allowing statistical methods to compensate for random observational error. The greatest obstacles, however, lay with

Table 10.2 Sources of Discrepancies between Calculation and Observation in Orbital Mechanics

In observations:	***In theoretical calculations:***
1. Simple error—"bad data"	1. Undetected calculation errors
2. Limits of precision	2. *Imprecise orbital elements*
3. Systematic bias in instruments	3. *Imprecise planetary masses*
4. *Imprecise fundamental constants*	4. Insufficiently converged infinite-series calculations
5. *Inadequate corrections for known sources of systematic error, incl.*	5. **Need for higher-order terms**
6. **Not yet identified sources of systematic error**	6. **Forces not taken into account**
	7. ***Gravitation theory wrong***

the three-body problem on the calculational side. Breakthrough after breakthrough was needed in the use of perturbational methods to obtain approximate solutions to the equations of motion of sufficient precision to make comparisons with observation physically meaningful. The history of celestial mechanics is primarily a history of those breakthroughs.[46]

Save for Neptune, therefore, the most celebrated discoveries of the first two centuries of gravity research were not about the world, but about analytical methods. The immediate goal driving effort on these methods was to assure that, any shortcomings in the theory of gravity aside, the calculations could predict planetary motion to within the accuracy of the observations. We should accordingly not be surprised that those working in the field tended to view any test of Newton's theory of gravity to lie solely with the success of these predictions after the mathematical obstacles had been sufficiently overcome. This constant focus on calculational problems ended up contributing more than anything else toward masking the logic with which Newton's theory was really being tested. The fact nevertheless remains that the test of the theory was not success in predicting the observed motions, but success in finding physical sources for discrepancies between the calculations and observation. This becomes glaringly obvious with the discrepancy in the precession of the perihelion of Mercury, to be discussed in Section 10.3.3. First, however, we should look more carefully at an example in which Newton's theory passed the test.

10.3.2 Closing the Loop: An Example

So far I have been concentrating on the difficulties encountered in gravity research following the *Principia*. I have yet to present any historical example in sufficient detail to demonstrate how stringently Newton's theory was actually tested. We turn in this section and the next to two such examples, one comparatively unknown that yielded extraordinarily strong confirming evidence and the other very well known that yielded the first decisive evidence of the inadequacy of the theory. They have more in common than one might expect.

Except for some of the asteroids, which of course were unknown to Newton, the Moon displays by far the most complicated orbital motion in our planetary system. Simply describing that motion to the same level of accuracy with which the motions of the planets were then being described remained a major unsolved problem from Ptolemy until the twentieth century, when the Hill-Brown theory finally met this standard. Newton introduced perturbational methods in the

[46] Curtis Wilson has been writing a history of these breakthroughs in a continuing series of monographs beginning with his (1980).

Principia as part of an argument that the best hope for ever describing the motion of the Moon to the accuracy of the planets was to give up the traditional approach of trying to derive it from observations and turn instead to his theory of gravity to derive it from the combination of the gravitational forces on the Moon from the Sun and Earth.[47]

In the early 1690s, shortly after the first edition of the *Principia*, Halley announced that ancient eclipses indicate some sort of secular—that is, non-periodic—acceleration of the Moon. This seemed to make no sense if the gravity of the Sun and Earth were the only forces acting on the Moon, and hence, on its face, it raised questions about whether Newton's proposed approach was ever going to succeed. Newton himself chose never to mention what came to be known as the "secular acceleration" of the Moon in either of the two subsequent editions of the *Principia*, or even in his short monograph on the Moon.[48]

Following Clairaut's 1749 discovery of how to derive the motion of the lunar apogee from the gravitational action of the Sun, the Moon became the first great success story in the history of perturbation methods, leading to the remark by Euler quoted at the beginning of Section 10.3.[49] The secular acceleration of the Moon, however, remained unaccounted for. By the 1770s it had become the subject of various proposals, including one by Immanuel Kant to which I shall return shortly. Euler himself, from the 1750s forward, saw it as evidence for a fluid in celestial space, keeping alive his hopes for a Cartesian-like vortex theory of planetary motion compatible with the discoveries Newton had made involving inverse-square gravity.[50] Hence it was a major

[47] Newton only hints at this view in the remarks he makes in passing in the Scholium to Book 3, Proposition 35 on the motion of the Moon in the *Principia*. In a letter to Flamsteed (February 16, 1695) pleading with him for more data, Newton offered a strong statement of it:

> For I find this Theory [of the Moon] so very intricate & and the Theory of Gravity so necessary to it, that I am satisfied it will never be perfected but by somebody who understands the Theory of gravity as well or better than I do (Newton, 1959–1977, vol. 4, pp. 86–89).

This statement proved prophetic insofar as the first theory of lunar motion to achieve a level of agreement with observation matching that of the planets was the Hill-Brown theory, and Hill undoubtedly did understand Newton's theory of gravity as well or better than Newton did.

[48] Newton (1702).

[49] The clearest indication of the success of perturbation methods at the time was Tobias Mayer's success in finally achieving a theory of the Moon's motion adequate for navigational purposes at the time. See Mayer (1753).

[50] That Euler had been looking for sometime for some secular acceleration that would provide clear evidence of a celestial etherial fluid can be seen in his correspondence with Mayer in the 1750s, especially in Mayer's letter of November 25, 1753 (Forbes 1971, pp. 76–78). His case that the secular acceleration of the Moon provides that evidence is in the prize-winning essay he submitted to the Royal Academy of Sciences in 1772, "Nouvelle recherches sur le vrai mouvement

breakthrough—the last nail in the coffin of vortex theories, as it were—when Laplace announced that he had found the source of the secular acceleration in 1787, a year after his discovery of the Great Inequality: planetary perturbations of the Earth were slowly reducing the eccentricity of its orbit, resulting in a small difference in the Sun's gravitational force on the Moon that was producing a roughly 10 arc seconds per century change in its motion.[51]

Laplace's discovery eliminated the last then-recognized systematic discrepancy in the orbital motion of the planets and their satellites. As noted earlier, the next quarter century, until the anomaly in Uranus's motion began to emerge, ended up being the one period between Newton and the 1990s when no discrepancies in orbital motion remained and gravity research seemed to have answered all the questions about deviations from Keplerian motion, at least up to the level of observational accuracy.

Laplace, however, had ignored higher-order terms in his solution for the Moon's secular acceleration. When Adams carried out Laplace's analysis to higher order in the 1850s, he discovered that the further perturbational terms cancel roughly half of Laplace's calculated change in lunar motion, leaving half of the secular acceleration unaccounted for.[52] The proposal that Kant had offered a century earlier[53] to no effect then took center stage: tidal friction delays the location of the tidal bulge, as depicted in exaggerated form in Figure 10.4,[54] and this produces a small torque that is slowing the rotation of the Earth, with the loss in its angular momentum transferred to and hence accelerating the Moon.[55] As a qualitative proposal this makes obvious sense.

de la lune," originally published in 1777 as "Réponse à la question proposée par l'académie royale des sciences de Paris, pour l'année 1772" (Euler 1777). Strikingly, Laplace takes the trouble to show that the effect of the resistance from any etherial fluid on the motion of the Moon is too small to be detectable in the theory of the Moon in his *Mécanique céleste* [Laplace 1966, vol. 3, pp. 678–698, especially p. 694].

[51] Richard Dunthorne had announced the number of 10 arc seconds per century as the "acceleration" in 1749—see Dunthorne (1749). This value actually represents the change in angular location over a Julian century, and not the acceleration, which is twice the value. Laplace presented his solution at the end of 1787 and published as Laplace (1786–1788). Laplace's explanation of the secular motion, by the way, removed it from this category, for the changing eccentricity of the Earth's orbit is a periodic phenomenon, though of very long term.

[52] Adams (1853, 1860, 1880).

[53] Namely in Kant's essay *Whether the Earth has Undergone an Alteration of its Axial Rotation* of 1754, a decade or more before Laplace's mathematical theory of tidal motion appeared (Kant 1900).

[54] Figure adapted from one in Stephensen (1997, p. 14).

[55] Invoking the transfer of angular momentum here involves much more than is immediately obvious. Neither angular momentum nor its conservation can be found in Newton's *Principia*. The principle that supplies the assurance that angular momentum is conserved in orbital motion—more accurately, in motion under purely centripetal forces—is Kepler's area rule. For a point mass moving under centripetal forces, the area rule mathematically amounts to a notational variant of conservation of angular momentum. As deployed in the *Principia*, however, the area rule is by no means sufficient to cover the transfer of angular motion between a body in orbit and a central rotating body, the purpose to which conservation of angular momentum is being invoked in the text. Whittaker (1904, p. 60) remarks, "Kepler's law, that the radius from the sun to a planet sweeps out equal areas in equal times, was extended by Newton to all cases of motion under a central force: from this the general theory of conservation of angular momentum has gradually developed." Clifford Truesdell (1968) has examined that process of development in his "Whence the Law of Moment of Momentum" in his *Essays in the History of Mechanics*. A point to emphasize here is that the generalizing of the law of areas into the conservation of angular momentum did not arise from the investigation of orbital motion under gravity, the topic of the present paper. Rather, it came out of work in other areas of mechanics, the motion of rigid bodies and deformable bodies.

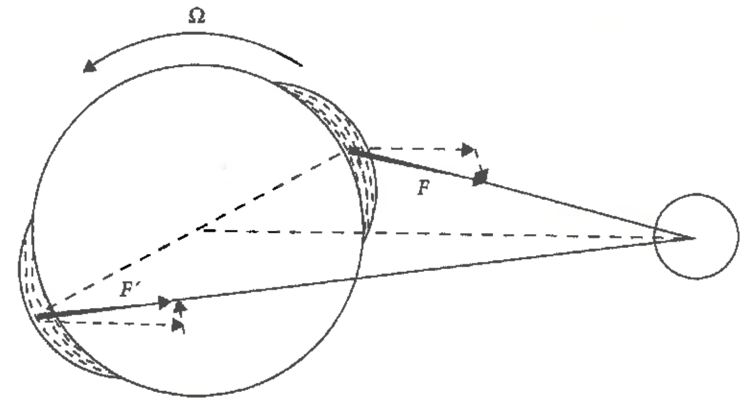

FIGURE 10.4 The Tidal Bulge and Consequent Torque Slowing the Rotation of the Earth. *Source:* From Stephenson (1997). *Historical Eclipses and Earth's Rotation.* Cambridge, UK: Cambridge University Press. Reproduced by permission of Cambridge University Press.

Verifying that tidal forces can indeed produce the fraction of the observed secular acceleration not accounted for by Laplace took six decades, including advances in tidal analysis by George Darwin and others.[56]

The tidal acceleration of the Moon brings out two important points beyond illustrating once more the difficulty of turning discrepancies into evidence in gravity research. First, part of the tidal effect involves a real acceleration in the motion of the Moon, while the other part involves a merely apparent acceleration, arising as it does from a slowing of the Earth—and hence from an imprecision in sidereal time for which corrections have to be introduced in earlier observations. From Hipparchus and Ptolemy forward, if not before, the fundamental problem in orbital astronomy was to distinguish between real and apparent changes of motion. The *Principia*, Newton

[56] See Darwin (1907, vol. 2, especially); the "six decade" reference is to three papers: Taylor (1920), Jeffreys (1920), and Heiskanen (1921). The complexity of this matter has kept it under discussion since then, as one can see from Jeffreys (1976, ch. 8).

had claimed, was offering a way of solving that problem.[57] The 130-year episode of the secular acceleration of the Moon supports that claim.

Second, tidal friction is a non-gravitational force. Newton's theory of gravity, as emphasized earlier, does not require all the forces affecting orbital motion to be gravitational. Euler had seen the secular acceleration of the Moon as evidence that some force beyond Newtonian gravity was at work. He turned out to be right, but instead of giving reason to augment Newton's theory, as Euler had hoped, the extended process of establishing the multiple physical sources of this very small effect ended up providing strong evidence supporting the theory.

The 10 arc-seconds per century secular advance of the Moon had not impeded progress on the lunar orbit during the nineteenth century. The effect was known, even if not the cause, and it was small compared with other still unresolved discrepancies. Peter Andreas Hansen, after he had developed a systematic method for generating all perturbational terms of any given order early in the 1830s, had devised a new theory of the Moon in 1838, publishing tables based on this theory in 1857 that agreed closely with observations from 1750 forward.[58] Making use of some seventeenth century observations upon which he had stumbled while visiting Paris, Newcomb called attention in the early 1870s to a departure of the Moon's motion from Hansen's tables. Newcomb returned to this still further anomaly in the Moon's motion at some length in the first decade of the twentieth century, but he remained unable to sort it out—that is, he remained unable to turn it into a well-demarcated second-order phenomenon, to use my terminology—when he died in 1909.[59]

Part of Newcomb's difficulty in characterizing this anomaly was the need for a still better lunar theory. As Hill noted in the remark quoted in Section 10.3.1 above, slowly converging infinite-series were producing a residue of small inaccuracies in calculation that tended to mask any physically significant discrepancies. After Newcomb persuaded Hill to postpone further work on the Moon and instead develop a new theory of Jupiter and Saturn, the task of completing the theory of the Moon that Hill had begun fell to Ernest Brown.

A couple of decades of effort were needed, but the Hill-Brown theory finally appeared in 1919.[60] It contains in excess of fourteen hundred separate perturbational terms derived

[57] Newton ends his Scholium on space, time, and motion (1999, p. 413), with the statement, "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."

[58] Peter Andreas Hansen (1838, 1857).

[59] Newcomb's publications on the residual inequality in the motion of the Moon extended from 1876 until 1912, three years after he died: Newcomb (1876, 1878, 1903, 1907, 1909, 1912).

[60] Brown (1919). The list of perturbational terms runs from pages 8 to 28.

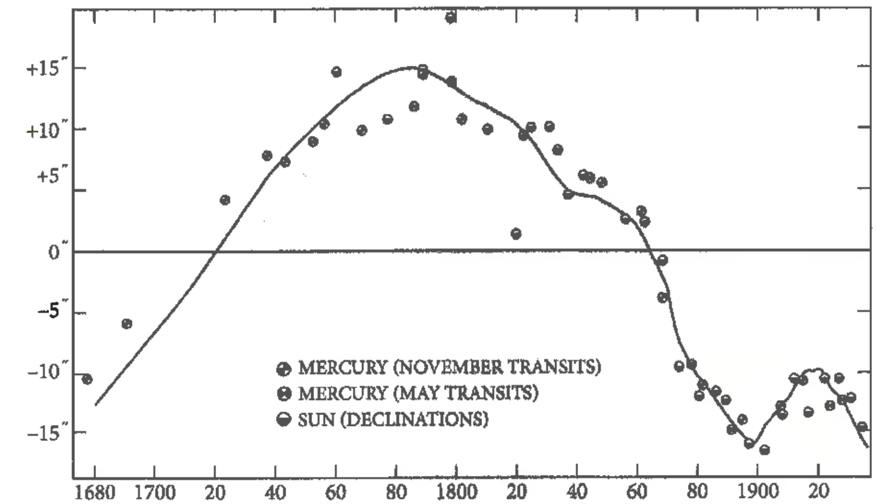

FIGURE 10.5 A Residual Discrepancy in the Motion of the Moon, and in the Motions of Mercury and the Earth. *Source:* From Jones, H.S. (1939). "The Rotation of the Earth, and the Secular Acceleration of the Sun, Moon, and Planets." *Monthly Notices of the Royal Astronomical Society* 99: 541–558. Reproduced by permission of Oxford University Press.

from gravity theory, plus one further term, which Brown called the "Great Empirical term," needed to bring calculations into agreement with observation. Harold Spencer Jones obtained the solid curve shown in Figure 10.5 by removing the Great Empirical term and then taking the difference between the Hill-Brown theory and observation. This curve thus represents a residual discrepancy of as much as 15 seconds of arc between observation and calculation after the fourteen hundred effects of gravity in the theory have been taken into account. The really important point, however, is not the magnitude of the discrepancy, but its clear signature, marked by its predominant sinusoidal form with a period of roughly 225 years. This is the second-order phenomenon that Newcomb had found, but had been unable to fully characterize.

The research question raised by the Great Empirical term was posed by Brown: is its source a *fluctuation* in the rotation of the Earth or is it instead some further unaccounted for force acting on the Moon.[61] Here once again orbital astronomy was confronted with the age-old problem of resolving whether an apparent variation in the motion of an orbiting body is merely apparent or real. Over the two decades after the Hill-Brown theory appeared, Brown, Willem de Sitter, and Harold Spencer Jones published papers on this problem, but only with Spencer Jones's landmark paper of 1939, from which Figure 10.5 is taken, was anyone able to marshal sufficiently compelling evidence to resolve it.[62]

[61] Brown (1926).

[62] Jones (1939). Among the papers leading up to this landmark paper were Jones (1926), de Sitter (1927), Jones (1932).

Spencer Jones summarized his reasoning in the diagram shown in Figure 10.5. If the systematic discrepancy in the Moon's motion described by the solid curve is caused by a further force acting on the Moon, then the effects of this force on Mercury in particular, but Venus and the Earth as well, will not have the same periodicity, and hence the same signature, displayed by the solid curve. If, to the contrary, the cause is a fluctuation in the rotation of the Earth, then Mercury, and Venus and the Sun as well, should each display a discrepancy with this same period.

A fluctuation in the rotation of the Earth, however, amounts to a systematic error in sidereal time. The magnitude of the discrepancy in orbital motions associated with it will therefore be greater the greater the mean daily motion of the orbiting body. So, the magnitude of any discrepancy arising from a fluctuation in the Earth's rotation will be around a factor of three smaller for Mercury than for the Moon, and smaller still again for each of Venus and the Sun. The solid curve in Figure 10.5 reaches a maximum of roughly 15 arc-seconds. Hence the *maximum* of any corresponding discrepancy for Mercury, over the course of the same 225-year period, will be of the order of 5 arc-seconds, and for Venus and the Sun, still less. Just as the many perturbational effects had masked the need for the Great Empirical term for the Moon until Hill-Brown theory revealed it, perturbational effects on Mercury, Venus, and the Earth could all the more easily have all along been masking the need for a corresponding term for them.

What Spencer Jones did in his 1939 paper was to take Newcomb's theory of Mercury, with the precession of its perihelion corrected for the general relativity effect we will examine in Section 10.3.3, and derive deviations from observations of it over the period of time covered in Figure 10.5. He then renormalized these small deviations by the ratio of the Moon's mean daily motion to Mercury's, roughly a factor of 3, to obtain the majority of the dots in Figure 10.5. The rest of the dots represent the same steps, taking deviations from observations of the Sun. The figure thus shows the correlation Spencer Jones exposed between the residual discrepancies of the Moon and the renormalized residual discrepancies of Mercury and the Sun. Though not shown in this figure, he obtained a comparable correlation in the case of Venus, using observations from 1835 forward. Using *same-effect, same-cause* reasoning, he concluded that the results establish a common source for all these discrepancies, namely a fluctuation in the rotation of the Earth over a roughly 225-year period—a still further source of systematic error in astronomical observations over the course of this period.

On the basis of Spencer Jones's findings, sidereal time was replaced by "ephemeris time" in 1950, obtained by taking essentially his curve as the requisite correction to sidereal time. Ephemeris time has subsequently given way to atomic clocks.[63] The

[63] See Seidelmann (1992, Chapter 2: "Time," especially Section 2.5, "The Historical Development of Timekeeping," pp. 73–87), and Audoin and Guinot (2001).

robustness of this fluctuation in the rotation of the Earth has been confirmed by means of these clocks and very-long-baseline interferometry and other measurements, along the way revealing various shorter-period smaller fluctuations of rotation beyond the one corresponding to Brown's Great Empirical term. The physical sources of the different fluctuations in the rotation of the Earth, which by conservation of angular momentum seemingly have to be from fluctuations in its polar moment of inertia, have been a prime topic of research over the last fifty years.[64]

Consider what went into Spencer Jones's finding. The solid curve in the figure represents a subtraction from observation of the 1400-plus perturbational terms of the Hill-Brown theory. Not every one of these terms is large enough to have masked or distorted the clear signature of the curve had it not been included, but a great fraction of them are large enough to do so individually, and clusters of small ones would do so as well. Consider, for example, just the effect of adding the 10 arc-second per century secular advance of the Moon to the curve. Furthermore, each of the perturbational terms in the Hill-Brown theory singles out some detail or other, like the extent of the oblateness of the Earth, that is making a difference in the motion of the Moon.

The number of perturbational terms in Newcomb's theory of Mercury is not so large, but here again a failure to include any one of a large fraction of them would have been enough to mask or undercut the correlation shown in Figure 10.5, and the same is true for Venus and the Sun—that is, the Earth. Newcomb had developed his theory of the four inner planets in combination with one another in order to include higher-order effects of their interactions. Consequently, a large number of details that make a difference to the motions of one or more of these four planets, including details governing the perturbational effects of the outer four planets on them, entered into the correlation Spencer Jones exposed between the residual discrepancies in the motion of the Moon and the motions of the three inner planets. So, we are talking about hundreds of perturbational terms involving Newton's law of gravity, and hence hundreds of details that according to this law make some specific difference in the motion of one or more of the four bodies involved in the correlation.

Needless to say, the point I intend this extended example to illustrate is that evidence from exposing further second-order phenomena and identifying robust physical sources for them *can* be very strong. It is evidence aimed primarily at the question of the physical exactness of the theory of gravity and secondarily at the question of this theory holding only in restricted circumstances and therefore not being unqualifiedly projectible. Over the course of the centuries following Newton, gravity research, in

[64] The standard reference, though now somewhat out of date, is Lambeck (1980). My choice of fifty years stems from the importance of one book in bringing attention to the matter, Munk and MacDonald (1960).

spite of the difficulties it faced, did indeed succeed in producing a sequence of successive approximations to the true orbital motions, and these approximations continually revealed still further second-order phenomena of progressively smaller magnitude.

The new second-order phenomena presupposed not only the theory of gravity, but also previously identified sources of earlier second-order phenomena, and those sources increasingly constrained the pursuit of sources producing the new phenomena. The continuing emergence of still smaller discrepancies with clear signatures and the subsequent success in identifying sources for them under these increasingly tight constraints thus also provided added evidence confirming the robustness of the previously identified sources. The theory became deeply entrenched from the history of its sustained success in exposing more and more subtle details of the physical world that make a difference without having to backtrack and reject sources discovered earlier. In particular, the history of pinning down progressively more constrained physical source after physical source gave increasingly strong reason to think that the theory of gravity really must be capturing the physical world, at least to very, very high approximation.

Logically, as well as historically, the evidence coming out of gravity research on orbits has taken the general form of the "feedback loop" shown in Figure 10.6. The process starts with a calculation of orbital motion that gravity theory entails would hold exactly if no other forces are at work. Comparison with observation then yields a discrepancy with a clear signature—that is, with a sufficiently distinct character that it amounts to a phenomenon. The next step is to find a physical source for this discrepancy, a physical source whose further implications are at least compatible with the prior calculation. This physical source, with all its further consequences, is then incorporated into a new calculation of the orbital motions, closing the loop and reinitiating the process.

Evidence emerges at three points in the diagram. First, a discrepancy with a clear signature, in contrast to random-looking deviations, is evidence that the idealized

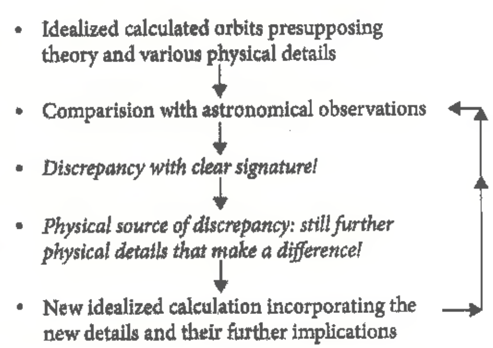

FIGURE 10.6 The Logic of the Evidence from Newtonian Research in Orbital Mechanics: Closing a "Feedback Loop"

calculation is in some respects physically correct. For the calculation is putting us in a position where the empirical world is, so to speak, telling us something. Second, finding a robust physical source compatible with the theory of gravity for such a discrepancy is evidence for that theory. If the theory is false, that should ultimately show up in the form of a discrepancy for which no physical source can be found compatible with it. And third, closing the loop and repeating the process, yielding still smaller discrepancies and increasingly subtle, more highly constrained sources for them, provides evidence for both the theory and the previously identified sources and the details entering into them. The more times the loop is traversed, the tighter the error bands become on both the theory *and* the previously identified details, so that the evidence for the theory-based counterfactual conditionals delineating the differences these details make becomes ever stronger.

One final twist to the Spencer Jones episode may be philosophically the most interesting of all. Least-squares statistical methods are used in celestial mechanics to set the values of parameters such as Keplerian orbital elements and masses. Typically, the discrepancies between observation and a preliminary theory are used to obtain refined values of the parameters in the "final" theory. As a consequence, when the theoretical calculation of the orbits agrees with observation, there is always some ambiguity about the extent to which the success is arising from curve-fitting versus the extent to which it confirms gravitation theory.

Just such ambiguity would have been present had the Hill-Brown theory achieved agreement with observation from the outset and Brown had not needed his Great Empirical term. That discrepancy, however, together with Spencer Jones's success in identifying what has proved to be a robust physical source for it, removed the ambiguity in this case. For, the evidence that the discrepancy itself is physical became evidence that all the elements entering into the calculation from which it emerged are physical too. The evidence this episode provided for gravitation theory and for the physical correctness of all the other perturbational terms of Hill-Brown theory thus became stronger thanks to the need for the Great Empirical term than it would have been had Hill-Brown theory simply succeeded in matching observations without that term! The exception really does "prove" the rule when robust physical sources are found for it. That is what makes the closing-the-loop logic so much more powerful than simple hypothetico-deductive testing.

10.3.3 The Perihelion of Mercury

Over the last forty years of his life Newcomb devoted a large fraction of his attention to two discrepancies between theory and observation, the lunar inequality discussed in the preceding subsection and a comparably small anomaly in the precession of the perihelion of Mercury. Neither of these was resolved at the time he died in 1909. The lunar inequality and its resolution by Spencer Jones is little known outside celestial mechanics, even to the point that the need for a purely empirical term in the legendary Hill-Brown theory of the Moon often goes unmentioned. By contrast, the 43 arc-second per century anomaly in the longitudinal location of the perihelion of Mercury seems more widely known than any other result in the history of orbital mechanics. Much of the reason for this undoubtedly stems from the singular standing of Einstein in twentieth century science and the crucial contribution the anomaly in the perihelion of Mercury made in the development of his theory of general relativity. Still, no small part of the reason for this anomaly being so widely known lies in its being the discrepancy between theory and observation that, to use Popper's term, "falsified" Newtonian gravity. More precisely, this anomaly was the second-order phenomenon that finally revealed that Newton's theory does not in fact hold exactly, but only approximately, and indeed even that only in quite restricted circumstances.

Fifty-six years separated the discovery of the small anomaly in the precession of the perihelion of Mercury and the conclusion that it provided clear evidence of a need to replace Newton's theory. Nothing about it initially made it seem any more of a threat to that theory than any of the other discrepancies we have been discussing. As Steven Weinberg has remarked of it, the lunar discrepancy discussed above, and anomalies in the motions of Encke's and Halley's comets,

> There is nothing in any single disagreement between theory and experiment that stands up and waves a flag and says, "I am an important anomaly." There was no sure way that a scientist looking critically at the data in the latter part of the nineteenth or the first decade of the twentieth century could have concluded that there was anything important about any of these solar system anomalies. It took theory to explain which were the important observations.[65]

It also took decades of pursuing a physical source for the Mercury anomaly to appreciate how recalcitrant it actually was. Its recalcitrance alone, nevertheless, did not expose the limitations of Newton's theory, nor did Einstein's new theory do so just by itself. Rather, it was his theory *in relation to* Newton's that did so. For just that reason, the evidential relationship between these two theories will be of central concern throughout the remainder of this essay.

[65] Weinberg (1993, p. 94). I thank Kenneth Wilson for calling my attention to this passage.

Leverrier had announced an anomaly in the precession of Mercury's perihelion of a little more than 38 arc-seconds per century in 1859, based on observations of eighteen transits of Mercury across the Sun, dating from 1661 to 1848.[66] (Transit observations are especially helpful for exposing small discrepancies because Mercury's alignment with the Earth reduces the margin of error in its heliocentric longitude.) Newcomb published the revised value of 43 arc-seconds per century in 1882, taking into account six additional transits that had occurred between 1848 and 1881.[67] This number held up through the huge study of the orbits of the four inner planets that culminated in Newcomb's 1895 monograph and his subsequent orbital tables.[68] The numbers displayed in Table 10.3 are from Newcomb's 1895 monograph. The comparatively large uncertainties in the values of the discrepancy for Venus and the Earth stem from the fact that the heliocentric longitude of perihelion is itself a quantity inferred from discrepancies in the longitudes of the planet.[69] Our immediate concern is with Mercury; for it Newcomb's uncertainty is a small fraction of the value.

The first thing to notice in Table 10.3 is that the precessions of the perihelia, as observed from the Earth, are all greater than 5000 arc-seconds per century. Almost all of this comes from the 26,000 year wobble of the Earth that produces the precession of the equinoxes. This, in other words, amounts to a systematic observational "error" that requires correction in order to refer the motions to the fixed stars rather than to a wobbling Earth. (G. M. Clemence, in explaining why the precession of Mercury's perihelion was not so straightforward a test for Einstein's theory, cited the wobble "as one

[66] Leverrier (1859). Leverrier's analysis is explained in Tisserand (1889, vol. 4, pp. 516–523). A brief summary can be found in Roseveare (1982, pp. 21–26). A transit of Mercury had also been famously observed by Gassendi in 1631, but Leverrier chose not to include it in his analysis. Also, his equations of condition begin with the transit of 1677, not with that of 1661.

[67] Newcomb (1882). Newcomb also used different masses for the four inner planets and Jupiter than the ones Leverrier had used, as indicated on page 469; in particular, the discovery of Mars's satellites in 1877 had provided a more reliable value for its mass.

[68] Newcomb (1895, 1898, p. 12).

[69] To be specific, discrepancies in the heliocentric longitude of the perihelion show up as discrepancies in heliocentric longitudes of the planet, which during transits are observed longitudes; in the calculations the consequences of precession for heliocentric longitudes arise from the product of longitude of perihelion and the eccentricity of the orbit. In the case of Mercury, with its eccentricity of 0.2056, the 43 arc-second discrepancy in the location of the perihelion was being inferred from a roughly 8.9 arc-second increase per century in the discrepancy in the observed longitudes during Mercury's transits. By contrast, the eccentricity of Venus employed by Newcomb was 0.0068, and hence the 8.9 arc-second increase per century in the observed longitudes for Mercury amounts to less than 0.3 arc-seconds per century for Venus, a magnitude that could scarcely be determined with much precision at the time, transits of the planet notwithstanding.

Table 10.3 Perihelia Advances of the Inner Planets: Newcomb's 1895 Comparison of Theory and Observation

	Mercury	*Venus*	*Earth*	*Mars*
Total observed advance per century by Newcomb	5598".8	5066".1	6176".0	6627.2
Observed advance after removal of general precession	575.1	42.4	1152.3	1603.5
Calculated advance based on Newtonian theory by Newcomb	531.7	31.8	1145.4	1594.5
Discrepancy between observation and Newtonian theory	43.4 ± 2.1	10.6 ± 36.5	6.9 ± 7.8	9.0 ± 3.8
Calculated additional advance from correction by Newcomb	43.37	16.98	10.45	5.55

of the most difficult problems of positional astronomy."[70]) Once the correction is made for it, the precession of the perihelion of Mercury, versus the fixed stars, Newcomb found to be 575.1 arc-seconds per century, versus a value he calculated from the gravitational perturbations of the planets on Mercury of 531.7 arc-seconds. The calculation using the theory of gravity was thus under-predicting the advance of Mercury's perihelion obtained from observation by 43.4 arc-seconds per century.

This difference of 43.4 arc-seconds for Mercury and the smaller values for the other three planets are, of course, not something that anyone could observe. They are second-order phenomena that presuppose not only Newton's law of gravity, but also

[70] Clemence (1947, p. 361). He cites two difficulties besides the one with the Earth's wobble:

> (1) Observations of Mercury are among the most difficult in positional astronomy. They have to be made in the daytime, near noon, under unfavorable conditions of the atmosphere; and they are subject to large systematic and accidental errors arising both from this cause and from the shape of the visible disk of the planet. (2) The planet's path in Newtonian space is not an ellipse but an exceedingly complicated space-curve due to the disturbing effects of all the other planets. The calculation of this curve is a difficult and laborious task, and significantly different results have been obtained by different computers.

the Newtonian gravitational effects of all the other planets on the precession of each of these perihelia. Several proposed sources of the discrepancy in the perihelion of Mercury had been put forward in the years after Leverrier had first called attention to it, including perhaps most famously now an additional small planet "Vulcan" inside the orbit of Mercury.[71] Newcomb reviewed some of the *prima facie* more plausible of these in his 1895 monograph, rejecting all of them, many because they violated constraints imposed by the orbits of the other three inner planets.[72] To give just one example of a different sort, Newcomb rejected the proposal that the 43 arc-second discrepancy was being caused by a non-symmetrical distribution of mass in the Sun on the grounds that the resulting non-sphericity of the equipotential surfaces surrounding the Sun would be seen in the photosphere surrounding it, yet heliometer measurements had ruled out any non-sphericity of the magnitude required to account for the Mercury anomaly.[73]

Newcomb could not leave the discrepancy entirely unresolved, however, for he was obligated to produce tables for the motions of the four inner planets that yielded as accurate predictions as he could achieve, and hence he had to find some way to correct for the discrepancy in the tables. He finally chose to add a fudge to the theories of the four inner planets by adding 0.0000000806 times the centennial mean motion of each to the rate of precession of their perihelia obtained from gravitation theory.[74] As one can see from the last row of Table 10.3, this value was chosen to give him four significant figure agreements with his value for the discrepancy in the precession of the perihelion of Mercury; but it also gave corrections for the perihelia of the other three planets, all of which fell within the bounds of uncertainty for them given in the next to last row of the table. Newcomb was subsequently nonetheless uncomfortable enough with the agreement in the case of Mars to remark:

> Two questions of capital importance require further investigation. First of these is the question of the excess of the observed over the theoretical motion of the perihelion. I have found this excess to be about 6" per century, with a probable error of less than 2". Although the presumption in favor of this result is very strong, the evidence can not be regarded as quite conclusive.[75]

[71] See Roseveare (1982, pp. 26–94) and the popular account by Baum and Sheehan (1997).
[72] Newcomb (1895, ch. 6).
[73] Ibid., pp. 111f.
[74] Newcomb (1898, p. 12).
[75] Ibid., p. 385. The other question concerned the mass of Venus, which was being inferred from a long-period perturbation harmonic in the motion of Mars.

At the time he adopted it Newcomb thought that his fudge might well have a physical basis. In 1894 Asaph Hall had proposed, in effect, to use the precession of Mercury's perihelion rather than the precession of the lunar apogee as the preferred measure of r in Newton's law of gravity.[76] Newcomb cited this in pointing out that his fudge amounted to replacing the exact value of -2 with -2.0000001612—that is, a change in the seventh significant figure after the decimal point. Brown was at work on the Hill-Brown theory of the Moon at the time, but had yet to resolve the residual uncertainty in the motion of the lunar apogee that Hill had emphasized in the quotation given in Section 10.3.1. By 1903, however, Brown had progressed to the point of announcing that the motion of the lunar apogee would not permit anything like that large of a change in the exponent![77] Where Hall and Newcomb required a change in the exponent of 1.6×10^{-7}, Brown limited any change to at most 4×10^{-8}. This left no physically viable hypothesis for reconciling the anomaly in Mercury's perihelion with Newtonian theory. Constraints from gravity research were making it impossible to find a physical source for the 43 arc-seconds.

In November 1915 Einstein announced that the new theory of gravity forming the heart of his extension of special relativity to general relativity gives the missing 43 arc-seconds on the nose.[78] Two years before, while working on an earlier version of general relativity and his new theory of gravity, Einstein had turned to Mercury's perihelion for evidence, but had found that that version of the theory yielded only 18 arc-seconds per century.[79] Much has been written about the development of Einstein's theory from 1913 until November 1915, but our concern is with evidence in gravitation research, and not with Einstein as such, and so suffice it here to say that what made the difference in the results for Mercury in the later theory was its including the curvature of space itself in the presence of a static gravitational field.[80] Until the solar eclipse of 1919, Einstein's success with the perihelion of Mercury was the sole empirical evidence for his theory of general relativity. Indeed, this theory significantly

[76] Hall (1894).

[77] Brown, (1903, p. 530).

[78] A. Einstein, "Erklärung der Perihelbewegung des Merkur aus allgemeinen Relativitätstheorie," in Einstein (1996a, Document 24, pp. 234–242). English translation by Brian Doyle, in Einstein (1996b, pp. 113–116).

[79] Specifically, the earlier theory predicts a relativistic effect on the perihelia that is 5/12 of the value predicted by Einstein's ultimate theory. See "The Einstein-Besso Manuscript on the Motion of the Perihelion of Mercury" in Einstein (1995, pp. 344–359, especially p. 351); and also, immediately following, Document 14, "Einstein and Besso: Manuscript on the Motion of the Perihelion of Mercury," Einstein (1995, pp. 360–473).

[80] I have relied primarily on Norton (1984), Stachel (2002), Janssen (2007), and Janssen and Renn (2007). I am especially grateful to Michel Janssen for correcting a mistake in the embryonic

underpredicted Newcomb's values for the discrepancy in the perihelia of Mars and the Earth. Einstein spoke of this at the end of his November 18, 1915 talk:

> For Earth and Mars, the astronomers assign, respectively, forward motions of 11" and 9" per century, while our formula yields, respectively, 4" and 1" per century. Nevertheless, a small value seems to be proper to these assignments because of the small eccentricities of the orbits of these planets. A more certain confirmation of the perihelion motion will be made by determining the product of the motion with the eccentricity. If we consider these quantities assigned by Newcomb, . . . then we obtain the impression that the advance of the perihelion is, after all, demonstrated really only for Mercury. However, I prefer to relinquish a final decision to the astronomical specialists.[81]

The "astronomical specialists" have fully vindicated Einstein in this regard. Table 10.4 replaces Newcomb's numbers in Table 10.3 with more recent numbers.[82] Myles Standish computed the general relativity effect by numerically integrating the parameterized post-Newtonian equations of motion twice for the entire planetary system over 2000 years, once for general relativity and then for Newtonian gravity.[83] The progressive differences in perihelia locations between the two integrations are shown in Figure 10.7, with the slopes giving the increments in the rates of precession from general relativity. As Table 10.4 indicates, the general relativity effect is now within the error bands of the discrepancy for all four planets. Notice also how excessive Newcomb's fudged corrections were for the other three planets. They provide a good example of why fudged corrections to achieve accurate predictions are no substitute for pinning down physical sources for discrepancies.

version of this paper and then guiding me through unfamiliar terrain, and to Christopher Smeenk for calling my attention to some further inaccuracies.

[81] Einstein (1996b, p. 116).

[82] As remarked in the text, the numbers in the bottom row giving the general relativistic contribution were obtained from Myles Standish in 1999, as were the plots in Figure 10.7. I am grateful to him for his help. All the other numbers for Mercury, Venus, and the Earth are from Duncombe (1956). The observed precession for Mars is from F. E. Ross's (1917) corrections to Newcomb. The other entries for Mars were inferred from this value by using Duncombe's 5026.41" per century value for general precession and a value for the discrepancy between Newtonian theory and observation from Ohanian and Ruffini (1994, p. 406).

[83] The parameterized post-Newtonian formalism is more than just a mathematically more tractable approximation to general relativity; variations of its parameters capture as well "metric theories" of gravity that are alternatives to Einstein's, including Newton's. A presentation of the PPN formalism can be found in Will (1993, ch. 4). For a still more extensive discussion of its history and its use, see Misner, Thorne, and Wheeler (1973, ch. 38 and 39).

Table 10.4 Perihelia Advances of the Inner Planets: 1999 Comparison between Theory and Observation

	Mercury	*Venus*	*Earth*	*Mars*
Total observed advance per century	5601".1	5062".9	6185".3	6626".3
Observed advance after removal of general precession	574.7	36.5	1158.9	1599.9
Calculated advance based on Newtonian theory	531.6	28.1	1153.9	1598.5
Discrepancy between observation and Newtonian theory	43.1 ± 0.45	8.4 ± 4.8	5.0 ± 1.2	1.36
Calculated additional advance from general theory of relativity	42.98	8.61	3.84	1.35
Newcomb's attempted correction for perihelion advance	43.37	16.98	10.45	5.55

The discrepancies between Newtonian theory and observation listed in Table 10.4 for the inner three planets are based on optical observations. Radar-ranging data obtained since 1966 have led to improved orbital elements and planet locations, and with them new values for the discrepancies: for Mercury, 43.1 ± 0.1; for Venus, 8.65; and for the Earth, 3.85.[84] The agreement with general relativity is thereby all the more impressive.

The increased precision gained over the last ninety years has accordingly strengthened the evidence Mercury's perihelion provides for general relativity. One might ask how a number that presupposes Newton's theory of gravity can be the basis for a test of Einstein's theory of gravity. The answer lies in a requirement Einstein imposed throughout his efforts on general relativity: Newton's theory of gravity holds in the static, weak-field limit of Einstein's theory, and consequently the Newtonian calculation of the perturbations of Mercury's perihelion by the other planets is accurate, in comparison with general relativity, to a precision well beyond the level of comparison with observation shown in the table.[85]

From the perspective of general relativity, then, Newtonian gravity is what I call a *limit-case idealization*. This is an entirely different kind of idealization from the "Newtonian" idealizations on which I have been concentrating so far. Newtonian idealizations, by definition, would hold exactly in certain specific circumstances.

[84] Ohanian and Ruffini (1994).

[85] The extent to which Einstein imposed such a constraint on general relativity comes through especially clearly in Norton's (1984) account.

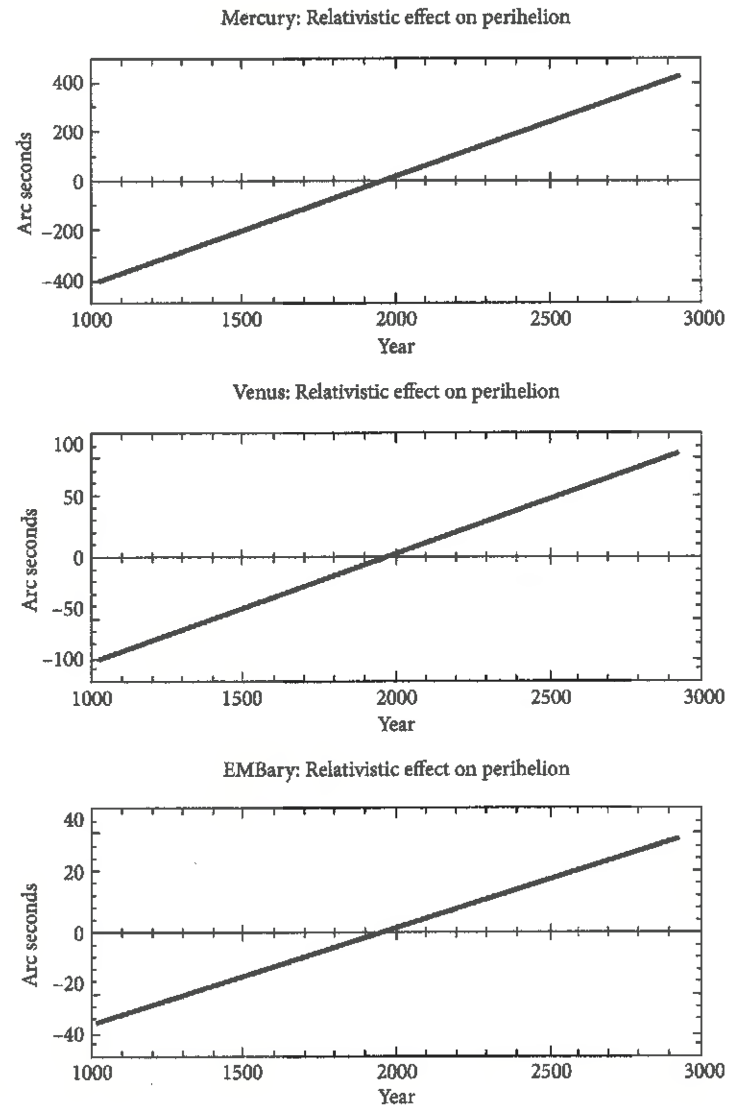

FIGURE 10.7 The Calculated Contribution of General Relativity to the Precession of Perihelia

A limit-case idealization would never hold exactly in any relevant circumstances whatsoever. It must always hold only approximately, though to asymptotically higher approximation as the limit is approached—in the present case, as the gravitational field approaches a static one of zero strength. Einstein's derivation thus did more than show that Newton's theory holds in the limit; it specified circumstances under which that theory is accurate to any given level of approximation.[86]

[86] Though it has gone largely unnoticed, Newton himself went to a good deal of trouble to show that the theory of gravity of Galileo and Huygens—uniform gravity along lines of acceleration

For our purposes, an especially helpful way of viewing the limit-case relationship between Newton's and Einstein's theories of gravity is in terms of the equation defining the orbit in the two-body problem—for example, Mercury orbiting the Sun in the absence of any other gravitational effects on either of them:

$$\frac{d^2u}{d\theta^2}+u=\frac{\mu}{h^2}+[3\frac{\mu}{c^2}u^2+\ldots]$$

where the unknown quantity u is $(1/r)$, that is, the inverse of the radius vector of the orbit; $\mu = G(M + m)$, the product of the sum of the masses and the gravitational constant; h is the angular momentum per unit mass or, equivalently, twice the areal velocity; and c is the speed of light. The bracketed term represents the general relativistic static-field correction to the Newtonian two-body equation that yields the Keplerian ellipse. The dots in this term stand for still smaller contributions from terms in higher powers of c^{-2}.

With the correction, the solution of this equation is a precessing rather than stationary orbit; in particular, in the case of Mercury the fully spelled out term in the brackets representing the correction to first order yields an orbit precessing at a rate of 43 arc-seconds per century.[87] Because the mass of our Sun is not that large, even in the case of Mercury this term is seven orders of magnitude smaller than the second term on the left, and of course it gets still smaller at greater distances from the Sun. The correction term comes from the combined effects of three physical sources: gravitational time dilation, non-linearities in the space-time metric produced by the Sun's mass, and the gravitationally induced slight curvature of space near the Sun.[88] These are the

parallel to one another—is a limit-case of his theory of universal gravity. See Newton (1999, Book 1, Section 10, especially Propositions 49–52), discussed in detail in Smeenk and Smith (in preparation).

87 For a derivation of the equation, an analysis of the comparative magnitudes of its terms, and the solution for the precessing orbit, see Ohanian and Ruffini (1994, pp. 401–408). A parallel derivation can be found in Garavaglia (1987).

For terms up to "the third post-Newtonian order ($\sim 1/c^6$ beyond the Newtonian acceleration)," see Blanchet (2001). Examination of the first-order post-Newtonian term in that paper will show that the two-body equation I have given in the text is itself a simplification for a weak static field. I thank Christopher Smeenk for calling my attention to the Blanchet paper.

88 See Ciufolini and Wheeler (1995, pp. 138–144). As they explain on page 140, this interpretation of the physical sources of the first-order correction term assumes the standard parameterized post-Newtonian gauge. A more detailed analysis of the contribution to perihelion precession (and the bending of light) made by the curvature of space can be found in Rindler (1977, pp. 133–149).

non-Newtonian physical sources of the Newtonian 43 arc-second residual discrepancy in Mercury's perihelion.

The limit-case relationship between Newton's and Einstein's theories has important implications for the history of evidence in gravity research. Conceptually, the transition from Newtonian celestial mechanics to general relativity constituted a revolution in physics in the sense emphasized by Thomas Kuhn: no sequence of incremental conceptual changes can bridge the gap between the two ways of thinking. In this respect the "conceptual readjustment," to use Kuhn's phrase, required by general relativity involved a discontinuous change. Nevertheless, because Newtonian theory holds in the static, weak-field limit, evidence remained continuous across the conceptual divide between it and Einsteinian theory.

Jed Buchwald and I have examined continuity of evidence in theory-change elsewhere, but not for the transition from Newton to Einstein.[89] In the case of this transition, "continuity of evidence" involves four elements, discussed here in ascending order of importance. First, as already noted, Einstein's limit-case reasoning legitimated his use of a Newtonian second-order phenomenon as evidence for his new theory. All that this required was that the net 531 arc-second per century gravitational effects from all the other planets on the precession of the perihelion of Mercury remain the same with Einsteinian gravity instead of Newtonian to a level of precision well beyond that of the observed value of this precession.

The situation here resembles Regnault's use of a constant-volume air thermometer, which presupposed the ideal gas law, in his extended experimental determination of the deviations from the ideal gas law exhibited by different real gases. The sole question is whether any imprecision in doing so is large enough to compromise the empirical conclusion being drawn. Still, the Newton-Einstein case differs from the case of Regnault in one notable respect. Paradoxical though it may seem, the 43 arc-seconds per century turned out to be at one and the same time evidence *against* Newtonian theory and evidence *for* it over a restricted domain. For, Einstein's success with the discrepancy had the effect of further validating the Newtonian derivation of the 531 arc-seconds.

89 Buchwald and Smith (2002). There we consider first the case of the several-step transition from Boyle's law of the 1660s to the virial expansion of current statistical mechanics, a clear case of continuous evidence in which the conceptual change was largely incremental; and second, the case of the transition from ray to wave optics in the first half of the nineteenth century, a clear case of discontinuous evidence that is, perhaps, one of the best candidates for Kuhnian incommensurability in modern physics. Smeenk and I have also examined continuity of evidence achieved by Newton's derivation of Galilean gravity as a limit-case of his universal gravity (Smeenk and Smith in preparation); this case more closely resembles the sort of continuity of evidence displayed in the Newton-Einstein transition, though Newton's main concern was to legitimate his use of key measures developed by Huygens, which presupposed Galilean gravity, as evidence for his claim that terrestrial gravity extends to our Moon.

Second, thanks to the limit-case reasoning, the evidence for Newtonian gravity, as a matter of historical fact, simply carried over to Einsteinian gravity with at most qualifying remarks about degrees of precision. Gravity research did not have to go back and restart from the eighteenth century forward. This is the element in continuity of evidence that Buchwald and I emphasized in the paper mentioned above. It implies a point raised at the very beginning of this paper: all along there was this other undiscovered theory that could yield predictions of planetary motions that were no less in agreement with observations than the predictions from Newton's theory. But it implies more than just this. *Each perturbational term in the Newtonian calculations of the orbits throughout the prior history of celestial mechanics has a counterpart in an Einsteinian calculation of these orbits that agrees with it to well within the limits of accuracy of observation.* Therefore, not only did all the then-available evidence for Newton's theory immediately become evidence for Einstein's (though not conversely),[90] but further, none of the evidence for Newton's theory could count as evidence against Einstein's.

Third, even though Einstein's theory limited, for example, the precision of orbital precession as a measure of the exponent in the inverse-square law, it did not invalidate or nullify the reasoning underlying the evidence for Newtonian gravity. That is, it did not entail that the purported evidence for Newtonian gravity was never really evidence at all, but only a coincidence creating a misleading illusion of evidence, as might be achieved by a mere curve-fit. In particular, Newton and those following him viewed the evidence as supporting the projectibility—in Nelson Goodman's sense of projecting open-endedly beyond known data—of his law of gravity and the many more specific generalizations that they derived from it. Under Einstein's limit-case derivation, that law and those generalizations remain projectible, but now as holding only to high approximation over a restricted domain. Gravity research is thus still being predicated on Newtonian theory in areas like geophysics in which its degree of approximation far exceeds the accuracy of observation, and for that matter in much of the research on "dark matter," which was initially proposed as a source for a Newtonian second-order phenomenon in the outer parts of galaxies. The one difference now is that the exactness of Newtonian theory is no longer being tested during the process of this research, but instead at most whether it is approximate in the right way.

[90] Lest anyone think that I am running roughshod over Kuhn's claim of incommensurability between Newton's theory and Einstein's, Kuhn fully granted me not only the point made in the text in our discussions during the last year of his life of his unfinished manuscript on incommensurability, but also the further point that the main reason Einstein insisted on being able to derive Newton's theory as a limit-case was to enable all the evidence for it to carry over immediately to his theory. Kuhn's reason for granting these two claims was simple: they are historical claims about practice in celestial mechanics, and as such they hold independently of any reconceptualization entering the process.

Fourth, and most important, the myriad of configurational details that Newtonian theory had revealed as making a difference in the actual motions of the planets and their satellites are all still there, making the same differences that they did before Einstein. The advent of Einsteinian theory has merely continued the process of identifying which details make a difference and what differences they make, adding still more physical details that make a difference, some of them ones that Newtonian theory could never have revealed. Thus, for example, Neptune is still the physical source of the systematic discrepancy displayed in Figure 10.2, and the relative location of the centers of the orbits of Jupiter and Saturn remains the crucial factor governing the "Great Inequality" in their motions. Notice that these details are not themselves forces in Newtonian mechanics, but sources affecting forces. This is true in general of the details I am referring to, which explains why they can remain in place even though Einsteinian theory itself reconceptualized Newtonian gravitational forces.

The claim that the details that make a difference remained in place, as I have stated it, is a straightforward historical claim about continuing practice in the history of orbital mechanics, notwithstanding any reconceptualizations of those details following Einstein. One can readily see from the paragraph before last why it is warranted. General relativity has not undercut the projectibility of Newton's law and the more specific generalizations derived from it over the domain in which they were historically taken to hold. Therefore, the various counterfactual conditionals licensed by that law and those generalizations throughout the first 200 years of that history remain valid, even after qualifications are introduced about the precision with which they ultimately hold. But those counterfactual conditionals themselves simply identify which configurational details of the world make a difference and what differences they make.

In a controversial passage of *The Structure of Scientific Revolutions* Kuhn remarked, "Though the world does not change with a change of paradigm, the scientist afterward works in a different world. . . . I am convinced that we must learn to make sense of statements that at least resemble these."[91] The fact that all the physical details revealed as making important differences in the motions of the planets and satellites by Newtonian theory remained in place after Einstein, with new details subsequently being added to them, points to a very important sense in which physicists engaged in gravity research continued to work in the same world. It was only because of this that Einstein's new theory turned the small anomaly in Mercury's perihelion into evidence showing, as I noted at the beginning of this section, that Newton's theory does not hold exactly, but only approximately, and that only in restricted circumstances.

[91] Kuhn (1970, p. 121).

However much it may appear otherwise, I do not intend these last remarks as a polemical rejoinder to Kuhn. The remark is not even inconsistent with the point Kuhn was trying to make in the passage quoted, namely that many aspects of the way in which the world is described within the framework of general relativity are thoroughly foreign to the Newtonian framework. Still, my remark is prompted in large part by objections Steven Weinberg has lodged against Kuhn.[92] Weinberg, though perhaps more outspoken than others, is typical of physicists who complain that philosophers and historians of science have more of a problem with the relation between theory and the physical world than they should. In conjunction with this, they think that philosophers and historians have overreacted to the Einsteinian revolution and the realization that Newtonian science, after two hundred years, turned out not to be the final word after all. Maybe what lies behind these intuitions of physicists is the extent to which the aggregate of physical details that Newtonian gravity has revealed as making a difference has simply continued to grow in the twentieth century. I cannot think of any better reason for *taking* gravity theory—first Newtonian and then Einsteinian—to be exact—or very, very nearly so—than the sustained success it has exhibited across now *three* centuries in revealing which physical details make a difference and what differences they make.

10.4 ON THE KNOWLEDGE ACHIEVED IN GRAVITY RESEARCH

Having conclusive reasons for taking a theory of gravity to be exactly or very, very nearly true as part of a research strategy is one thing; knowing it to be exactly or very, very nearly true is quite another. As Wittgenstein reminds us, we must never lose sight of the expression, "I thought I knew," as in "many physicists in the middle of the nineteenth century thought they knew that Newton's law of gravity was going to be the final word on the matter." The questions posed at the beginning of this essay concern the nature, scope, and limits of the knowledge achieved across three hundred years of gravity research, with special emphasis on any claim to knowledge achieved over the first two hundred years, before general relativity. Questions about knowledge concern the claims propositions have to permanence, in the case at hand in the face of continuing research over the course of the twentieth century and into the future.[93] What, if

[92] Weinberg (2001, pp. 187–206).

[93] Weinberg's criticism of Kuhn (Weinberg 2001, p. 198), includes just such a claim of permanence:

> It is important to keep straight what does and what does not change in scientific revolutions, a distinction that is not made in *Structure*. There is a "hard" part of modern physical theories ("hard" meaning not difficult, but durable, like bones in paleontology or potsherds in archeology) that usually consists of the equations themselves, together with some understandings about what the symbols mean operationally and about the sorts of phenomena to which

anything, has the evidence that has come out of gravity research shown about claims of propositions to finality?

Again as noted at the beginning of this essay, relativity theory put these questions into a new, troubling light. Suppose that the only change that had ever been required of Newton's theory was a small increment in the exponent of *r* of the sort that Hall and Newcomb entertained. That would scarcely have undermined all claims Newtonian theory had to being knowledge, for we have always recognized that empirical evidence can never show relations among quantities to hold to anything beyond high approximation. Newton himself added the phrase "or *quam proxime*" after "exactly" in his Fourth Rule of Reasoning. General relativity, however, has given strong empirical grounds for concluding that some of the most fundamental assumptions underlying Newtonian gravity are false—most notably, that one can at least in principle always resolve questions of simultaneity and always distinguish free fall under gravity from inertial motion.[94] Still, one cannot just say that general relativity has shown Newtonian theory to be false and hence to have no claim at all to being knowledge. For, the limit-case relationship between Newtonian theory and general relativity described in

> they apply. Then there is a "soft" part; it is the vision of reality that we use to explain to ourselves why the equations work. The soft part does change; we no longer believe in Maxwell's ether, and we know that there is more to nature than Newton's particles and forces.
>
> The changes in the soft part of scientific theories also produce changes in our understanding of the conditions under which the hard part is a good approximation. But after our theories reach their mature forms, their hard parts represent permanent accomplishments. If you have bought one of those T-shirts with Maxwell's equations on the front, you may have to worry about its going out of style, but not about its becoming false. We will go on teaching Maxwellian electrodynamics as long as there are scientists. I can't see any sense in which the increase in scope and accuracy of the hard parts of our theories is *not* a cumulative approach to the truth.

In response to a letter to the editor objecting to this passage, Weinberg (ibid., p. 207) subsequently added:

> For one thing I don't see the difficulty in describing an approximate theory as "true" or "false." As Professor [Alex] Levine has guessed, I don't regard approximations as mere useful fictions....
>
> But approximate theories are not merely approximately true. The can make a statement that, though it refers to an approximation, is nevertheless precisely true. For instance, although Maxwell's equations give only an approximate account of electric and magnetic fields, it is precisely true that the error introduced by using Maxwell's equations to calculate these fields can be made as small as one likes by considering fields that are sufficiently weak and slowly varying. This is part of the reason Maxwell's equations are a permanent part of physical science.

These statements by Weinberg did much to prompt Part 3 of the present essay.

[94] I owe Rob DiSalle for making clear to me that these two claims are the ones that most deserve to be cited.

Section 10.3.3 has left it with a very different status than, for example, post-Lavoisier chemistry left phlogiston theory.

We cannot, of course, simply assume that general relativity is going to be the final theory of gravity, but that does not preclude asking what it tells us about the force of the evidence that had accumulated for Newton's theory before Einstein.

One thing general relativity has driven home is that the evidence in support of Newton's theory was all along misleading in one crucial respect: the phenomena involved in it were not actually representative of gravitational phenomena generally. Inductive reasoning invariably assumes that the available data are representative of the entire body of possible data. Newton, recognizing this, took the caution in his third Rule of Reasoning of saying that "qualities of bodies... that belong to all bodies on which experiments can be made should be regarded as qualities of all bodies universally," where the phrase "experiments can be made" extends as well to the motions of planets and their satellites in our solar system. General relativity, however, and the twentieth century evidence supporting it entail that these motions are not truly representative of motions governed by gravity generally. They are what they are because the gravitational field surrounding the Sun is virtually stationary and unusually weak. Drawing conclusions from misleadingly parochial evidence that one has no way of recognizing to be parochial at the time is an unavoidable hazard in scientific research. The evidence for Newtonian theory over the first two hundred years of gravity research fell prey to this hazard.

Another thing that general relativity has driven home—forcefully, I should add—is that the entire body of evidence accumulating in support of Newtonian theory was all along evidence that could support any number of then-unformulated alternatives to that theory. I say "forcefully" because the point is stronger than Duhem's point from the beginning of this essay. Section 10.3.2 above has explained how the evidence for Newton's theory was much stronger than just agreement between calculation and observation. That account of the evidence underscores how compelling the reasons were for Einstein's requiring Newtonian theory to be a limit-case of any new theory of gravity. Thanks to this limit-case relationship, it is not merely that all the predictions of planetary motion derived from Newtonian theory and confirmed by observation could in principle have been derived as well from general relativity. All the counterfactual conditionals derived from Newtonian theory and supported by the evidence for them could have been derived from general relativity and supported by the very same evidence. This is true not only of general relativity, but of some of the preliminary theories Einstein considered before he reached it and, more importantly, of various theories that were subsequently proposed as alternatives to it, such as Dicke's theory.[95]

[95] Derivation of the Dicke-Brans-Jordan theory of gravity can be found in Misner, Thorne, and Wheeler (1973, ch. 39), and Weinberg (1972, pp. 157–160). Will (1993), much of which

That all these theories can meet the limit-case requirement shows that the evidence that had accumulated for Newton's theory before Einstein, even under my strong interpretation of it, did not, and indeed could not, foreclose the possibility of evidence ultimately emerging that would require some other theory compatible with the accumulated evidence to replace it in the future.

Some important evidence for Einstein's theory of gravity has come from orbital motions within our solar system. In addition to the precession of the perihelion of Mercury,[96] for example, is the precession of the other three inner planets. (Clear observational support for the general relativistic effect on the perihelion of Venus finally emerged in the 1950s.[97]) Yet much of the evidence for Einstein's theory has come from sources other than orbital motions in our solar system, for Einstein's theory has implications, like the gravitational bending of light and black holes, that Newton's theory does not. The development of evidence bearing on general relativity is an ongoing activity on which I am not qualified to speak. I am therefore going to restrict questions about the knowledge achieved in the continuing discussion to evidence from sources that dominated the first two hundred years of gravity research, that is, orbital motions and gravitational fields in our solar system.

The questions of this section, then, concern the nature, scope, and limits of the "knowledge" achieved in gravity research from three hundred years of using data from our solar system to test Newton's theory of gravity. What, if any, claims to permanence have come out of the evidence provided by these tests? Even though much of it predates general relativity, the implications of this evidence still have to be considered in the light cast on it by Einstein's theory—that is, in the light of the parochialism of this evidence and its inability to rule out yet to be formulated alternative theories. I shall proceed in three steps: first, to re-examine the structure of the evidence in the light of Einstein, then to consider what the evidence shows, and finally to assess what claims to permanent knowledge it has yielded.

10.4.1 The Structure of the Evidence

Figure 10.6 offered a schematic depiction of the structure of the evidence from such historical examples as Spencer Jones's showing that the source of the "Great Empirical

can be viewed as laying out the evidence for general relativity versus alternatives to it in which Newtonian gravity is still a limit case, summarizes Dicke's framework for testing general relativity in chapter 2.

[96] The perihelion of Mercury has continued to be a source of evidence bearing on Einstein's theory versus alternatives to it; see Will (1993, pp. 180–183).

[97] See Duncombe (1956), which ends with the remark, "For the first time we have observational results accurate enough to detect the relativity effect in the motion of the perihelion of Venus."

Term" in Hill-Brown lunar theory is a small, long-term fluctuation in the Earth's rotation. The remark accompanying that figure in Section 10.3.2 was that evidence arises at three places in the diagram: (1) the emergence of discrepancies with a clear signature between calculation and observation, that is, second-order phenomena; (2) the identification of robust physical sources for these discrepancies, that is, configurational details that make the specific differences in question and other confirmable differences as well; and (3) the closing of the loop in which the sequence repeats itself, each time presupposing prior second-order phenomena and their identified sources, but with newly emerging discrepancies typically smaller and smaller, and hence with the agreement between calculation and observation ever tighter. Schematic though it is, that figure, so understood, does depict the general structure of the evidence generated by the *successful* testing of Newton's theory in classical celestial mechanics and physical geodesy. It does not, however, depict the structure of the evidence arising from the second-order phenomenon of the 43 arc-seconds per century additional advance of the perihelion of Mercury. This was a case in which Newton's theory *failed* the test: the physical sources identified for this discrepancy were not compatible with that theory.

I propose that Figure 10.8, properly understood, covers all the testing of gravitation theory—Newtonian and Einsteinian—that has arisen from research on orbital motions in our solar system and the gravitational field surrounding the Earth. The added step termed "theory change when deemed necessary" is enclosed in parentheses because it becomes pertinent only when the physical source responsible for the discrepancy is not one that the prior theory recognizes as capable of yielding it. Adding the new source to the calculation is not enough; the theoretical principles underlying the calculation must undergo some change as well. As already noted, once the general relativistic changes were introduced into the calculation of the orbital motions, not only did the anomaly in Mercury's perihelion disappear, but the many second-order phenomena and the Newtonian physical sources of them presupposed in the derivation of this anomaly remained intact. More generally, introducing the general relativistic changes into the calculations for both orbital motions and the Earth's gravitational field, as a matter of historical fact, left all the prior Newtonian second-order phenomena and the sources of them intact. Although more must be said, this is basically what justifies the implication of Figure 10.8 that, so far as the structure of the evidence from orbital motions is concerned, the feedback loop itself remained intact in the transition from Newtonian to Einsteinian theory.

One respect in which Figure 10.8 threatens to mislead lies in the vague term "theory-change." The phrase, "until yet other phenomena make such propositions either more exact or liable to exceptions" in Newton's fourth Rule of Reasoning indicates that he allowed for the possibility of ongoing research forcing theory-change. But the sort of theory-change his phrase anticipated might more aptly be called "theory-revision."

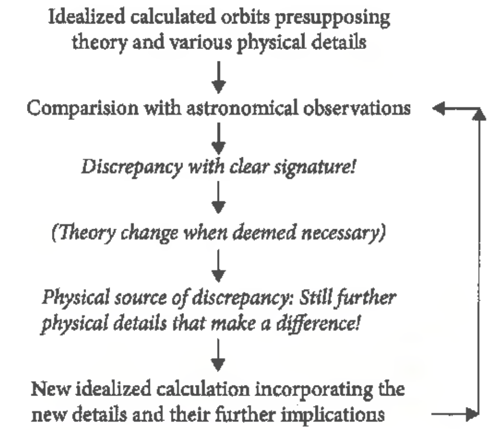

FIGURE 10.8 The Logic of the Evidence from Orbital Research after General Relativity: The Loop Still Closed

The small change in the exponent of r in the law of gravity considered by Hall and Newcomb would have been an example of making the theory more exact had it proved to be the source of the anomaly in Mercury's perihelion. Clairaut's proposal of adding a $1/r^4$ term to the law to handle the then missing factor of 2 in the precession of the lunar perigee (and a phantom geodetic discrepancy in the combination of the Earth's oblateness and surface gravity at the equator[98]) would also have been an example of making the theory more exact, though it might as well have been regarded as arising from exceptions to the law of gravity when the distance from the Earth is small.[99]

In neither of these examples would the proposed theory-revision have opened the way to possible physical sources of discrepancies beyond the unaccounted for forces and density variations allowed by Newton's theory; and neither would have rejected fundamental assumptions underlying that theory. In doing both, Einstein's theory amounted not to a revision of Newton's theory, but to a replacement. The term "theory-change" in Figure 10.8 was chosen to cover both the sort of revisions Newton had anticipated and the sort of theory replacement that occurred with Einstein. As we have seen, the parameterized post-Newtonian framework allows relativistic corrections to be added to equations and other relationships that were originally derived from Newton's law of gravity. From the point of view of the calculations of orbital

[98] At the time he proposed the $1/r^4$ term, Clairaut was also struggling to reconcile the measured value of surface gravity at the equator with the anomalously large degree of oblateness implied by geodetic measurements there and in Lapland. Inaccuracies in the Lapland measurements were greatly exaggerating the degree of oblateness.

[99] Notice, by the way, that Newton's law would have held in the long-distance limit-case of Clairaut's proposal.

motions, the transition to Einstein thus appears to amount only to the same sort of theory revision as Clairaut's adding a small $1/r^4$ term. But that is wrong. The correction terms represent physical sources, like the gravitation-induced curvature of space, that fall entirely outside Newton's theory; and they are derived from principles that contradict fundamental assumptions underlying that theory. We have no reason to think that Newton ever anticipated any possibility of so radical a theory-change.

Of course, Einstein's new theory was not put forward as a response to the discrepancy in Mercury's perihelion. This is another respect in which Figure 10.8 threatens to mislead. Granted, Einstein appears to have viewed that discrepancy as a crucial test while he was developing his theory, and for a couple of years it provided the sole empirical evidence for it versus Newton's. Einstein's new theory nevertheless grew out of far more fundamental concerns. Special relativity solved the problem of reconciling Newtonian mechanics with Maxwell's equations. General relativity began as a response to the problem of extending special relativity to frames of reference accelerating with respect to one another. It was Einstein's genius, so to speak, to link this problem, which seemed largely mathematical, to the physics of gravity. The radical departure from Newtonian theory had far more to do with these considerations than with the recalcitrant problem then posed by Mercury's perihelion.

Thanks to the limit-case relationship, the change from Newton's to Einstein's theory, radical though it was, nevertheless conformed with the structure of evidence schematized in Figure 10.8. All the specific Newtonian second-order phenomena that had emerged in the earlier course of research in celestial mechanics remained intact across the change, along with the specific physical details that had been identified as their sources.

Two historical contingencies lay behind this. First, the level of precision with which those second-order phenomena had been defined was too gross for the numerical differences between Newton's and Einstein's theory to have been detectable in them. Second, all the second-order phenomena that had emerged before the perihelia anomalies happened to have been associated with weak gravitational fields that were virtually static. Such contingencies do not alter the historical fact that Einstein's new theory conserved the earlier second-order phenomena and their physical sources and thereby continued the loop-structure of Figure 10.8. We might call a replacement of one theory by another that does this a *restricted, asymptotically conservative theory-change*, where the word 'restricted' serves to remind us that contingent factors entered into the new theory's conserving the prior evidence.

Yet another respect in which Figure 10.8 threatens to mislead stems from its suggesting that all the evidence developed in gravity research fits into the loop. This was not true even of the evidence for Newton's theory. During the second half of the nineteenth century physicists finally began conducting Cavendish-like experiments to test

Newton's law of gravity in which the attracting masses and the distances from them to the attracted masses were systematically varied.[100] (Cavendish had varied neither in the experiments he reported in 1798; rather he had presupposed the law of gravity to measure, in effect, the constant of proportionality in it, G.) The evidence for Newton's law from those experiments does not fit within the structure shown in Figure 10.8. Still, virtually all the evidence for Newton's law before the anomaly in the perihelion of Mercury became pressing did fit that structure.

More importantly, very little of the evidence for Einstein's theory of gravity that has been developed over the ninety years since he proposed it fits into that structure. This is because very little of that evidence involves orbital motion in our solar system or the gravitational field surrounding the Earth.[101] Some of the evidence for the new theory nevertheless does fit into the structure. In particular, in the wake of the success with Mercury, the loop soon became traversed again with the success achieved with the perihelia of the other three inner planets and with the tiny discrepancies in the longitudes of Mercury, Venus, and the Earth that enabled Spencer Jones to identify the long-term fluctuation in the Earth's rotation. Furthermore, the increased precision in the agreement between calculation and observation in the case of Mercury's perihelion has entered into the evidence comparing Einstein's theory with such alternatives to it as the one proposed by Dicke.[102] So, the loop shown in Figure 10.8 has continued to be traversed during the twentieth century. Indeed, in some cases the loop has continued to be traversed employing only Newton's theory—for example, in the recent systematic discrepancy in the motion of the object Sedna in the Kuiper belt and speculation that it might be revealing a "hidden planet."[103]

So much for respects in which Figure 10.8 might prove misleading if not properly understood. The point is that the evidence from orbital motions in our solar system and the gravity field surrounding the Earth continued to close the loop both with and following the switch to Einstein's theory. Still more important, *it did not have to be this way*. Einstein imposed the requirement that Newton's theory hold as some form of limit-case. Two centuries of the loop structure had given him overwhelming reason to impose that requirement and thereby conserve as much of existing celestial mechanics as he could. He nevertheless had no guarantee that the limit-case relationship between the old and the new theory would leave in place *all* the specific second-order phenomena that had emerged during those two centuries and the physical sources that had been identified for them. The new theory might have required some backtracking

[100] See Smith (*forthcoming*).

[101] The standard reference reviewing the evidence for general relativity is Will (1993), but see also Ciufolini and Wheeler (1995, Chapter 3: "Tests of Einstein Geometrodynamics," pp. 87–184).

[102] See, in particular, the discussion in Ciufolini and Wheeler (1995, pp. 144–146).

[103] See, for example, Lykawka and Mukai (2008).

in which some of these second-order phenomena or their physical sources became reconstituted in ways not strictly in keeping with their original form.

For that matter, Einstein might have been unable to find *any* new theory that simultaneously met the limit-case requirement and yielded the 43 arc-seconds per century anomaly. An empirically successful new theory might have required compromise or abandonment of the limit-case requirement. The loop structure shown in Figure 10.8 would then simply have ceased once Newton's theory was replaced. That the loop structure stayed in place is—to use a phrase Newton used at a critical point in the *Principia*—"worthy of note."[104]

Just as before, evidence arises at three places in Figure 10.8: the emergence of discrepancies with clear signatures; the identification of robust sources for these discrepancies; and the repeated closing of the loop, yielding progressively tighter agreement between calculation and observation. The one key difference is that non-Newtonian sources of discrepancies, like the curvature of space, are now providing evidence for the previously identified Newtonian sources presupposed by the discrepancies.

Again, it did not have to be this way. Einstein's theory of gravity might have implied that Newtonian calculations had misrepresented the effects of certain physical sources, or even worse, that some physical sources do not make anything like the differences Newtonian theory implies they make. Instead, the non-Newtonian sources have provided all the more evidence for earlier Newtonian sources and the differences Newtonian theory claims they make. This is what I meant when I said in Section 10.3.3 that the 43 arc-seconds per century turned out to be evidence *for* Newtonian theory as well as *against* it.

Figure 10.8 requires one last clarification. It is intended to depict the structure of the historical evidence, that is, the logical structure as the evidence actually happened to have unfolded historically. The diagram has second-order phenomena emerging before their sources are identified and smaller discrepancies hidden until the physical sources of larger ones have been incorporated into the calculation. But this is not the only form of evidence that orbital motions have yielded for gravitation theory. One can take the idealized calculation at any stage and conduct "quasi-experiments" with it by dropping or modifying some physical factor in it. Standish's dropping the general relativistic contributions to the perihelion precessions in one of the calculations that yielded the curves shown in Figure 10.7 is an example of this. Direct comparison of

[104] In the Scholium following Book 1, Proposition 77, where Newton was referring to the somewhat remarkable conclusion that there are only two laws of centripetal attraction that are the same for attraction toward whole spheres and toward the individual particles of matter forming them, namely inverse-square and linear. Understatement though it may be, this is the only place in the *Principia* where Newton offers any such remark.

the results of that calculation with observation over the last 350 years would then have yielded second-order phenomena very close to the curves in the figure, thereby providing evidence in support of the general relativistic corrections.

This example generalizes. One can always tweak the idealized calculation by removing or modifying some physical factor, whether that factor did or did not historically enter the calculation in response to some specific discrepancy. Discrepancies with clear signatures will not automatically emerge when the altered calculation is compared with observation. But when they do, they provide evidence bearing on what differences the physical factor that was removed or modified actually makes. Any number of tests of the theory can be carried out by means of this strategy, revealing any number of second-order phenomena that happened not to have emerged historically. Both the tests and the emergent phenomena are nevertheless empirical, and not merely calculational, for the key step has the "counterfactual" calculations being compared with observation. Figure 10.8 should not be taken as dismissing such extra-historical evidence or denying its value. It too is evidence from orbital motions in our solar system.[105]

10.4.2 Evidence for What?

An important point needs to be made before we turn to questions about what the evidence has shown. Throughout the essay, but especially so in the preceding section, I have been speaking of second-order phenomena arising out of comparisons between calculations based on gravity theory and observation. It goes without saying that the observation being referred to has not been theory-independent. Raw optical observations have been corrected originally for parallax and atmospheric refraction and subsequently for the finite-speed and aberration of light, not to mention the theoretical considerations entering into the determination of time. Moreover, the reductions of the observations to determine such things as areal velocities and perihelia locations have presupposed aspects of Keplerian orbital theory. Nowhere have I been taking observation to be theory-neutral. Nor, as Table 10.2 in Section 10.3.1 illustrates, have I been taking the theoretical calculations to be observation-neutral. The values used for the masses and orbital elements in the calculations have been derived from observations. But now the important point: *before the advent of relativity theory, none of the astronomical observations with which calculations were compared in any way presupposed Newton's theory of gravity.* The assumptions underlying the observational processes, the corrections made to the raw observations, and the steps in the reduction of them to their summary form were all independent of the theory of gravity.

[105] I am indebted to Christopher Smeenk and Sheldon Smith for not letting me lose sight of this point.

The gravitation-theory-neutrality of solar-system observations ceased with the advent of general relativity. Observations now call for relativistic corrections that depend, at least in part, on what the theory says about gravity.[106] Observation in physical geodesy has never been gravitation-theory-neutral, for pendulum measurements of surface gravity involved assumptions about gravity and corrections for altitude were being made from shortly after the time of Newton. Again, it goes without saying—or at least it should—that such gravity-theory incursions into observation do not automatically undercut or even seriously compromise the evidence arising from comparisons of calculation and observation. Any such incursion, however, does call for careful logical analysis to determine the precise force of the resulting evidence. In asking "Evidence for what?"—I am mostly concerned with the evidence before general relativity. Even the post-relativistic evidence from optical observations bearing on the precession of perihelia, as summarized in Table 10.4, and hence on Spencer Jones's fluctuation in the Earth's rotation, is free of any incursion of gravity theory into observation. So, the discussion in this section can proceed without having to worry about how such incursions have affected evidence elsewhere bearing on Einstein's theory.

That said, let me proceed in two steps, asking what the evidence from orbital motions in our solar system has shown first before general relativity and then after it. As noted earlier, the most immediate bearing the evidence before the twentieth century had was on the accuracy of Newtonian theory. For example, because of the factor of 2 Newton could not account for in the precession of the lunar perigee, the most he could say was that the exponent of r deviates from -2 by no more than 2 parts in 243. Clairaut in 1749 reduced this bound by more than an order of magnitude. Hill claimed in the passage quoted earlier that the improvements in lunar theory from Clairaut until 1870 still left the bound greater than 1 part in 500,000. By the first decade of the twentieth century Brown's efforts on the Hill-Brown theory had reduced it to 4 parts in 100 million.

More generally, though the statement has to be vague, the 1 to 2 arc-second agreement between calculated and observed planet locations achieved by Newcomb and Hill, aside from the precession of Mercury's perihelion, confirmed the very high precision of Newtonian theory.[107] Even more so, it confirmed success in identifying all the factors that have effects on orbital motions of sufficient magnitude to show up at this level of precision. The evidence before the anomaly in the perihelion of Mercury, however, gave no informative way of characterizing the domain over which Newtonian theory achieved this level of accuracy. Nothing but a systematic discrepancy or observations beyond those from the solar system could have provided evidence that Newtonian theory does not hold to high precision everywhere.

A major caveat is needed here: the evidence confirmed the high accuracy of *those aspects of Newtonian theory that entered indispensably into the calculations*. One element of Newtonian theory that made no detectable difference in the calculations of orbital motions before the twentieth century was the claim that the strength of the gravitational field surrounding each body is proportional to its mass. This is because the masses of celestial bodies were inferred from the inverse-square centripetal acceleration fields surrounding them, and these masses then entered the calculations only as *GM*—that is, as the inverse-square centripetal acceleration fields from which they were inferred. In other words, in the calculations of orbits the masses of the celestial bodies served only as placeholders for the acceleration fields from which they were inferred. Consequently, as I have discussed at length in a companion essay, before the twentieth century the evidence from orbital motions did not confirm that the mass of the attracting body even has to be present in the numerator of the law of gravity, much less that its exponent there is 1, at least *quam proxime*.[108]

[106] See Brumberg (1991), Chapter 6: "Relativistic Reduction of Astrometric Measurements," and Chapter 7: "Relativistic Effects in Geodynamics," for details on the incursion of Einstein's theory of gravity into observation.

[107] Secondary sources often claim that perturbation approaches with planetary orbits achieved a 1 arc-second level of agreement with observation with Newcomb, Hill, and those following in their footsteps. The actual situation, however, is more complicated than this. Hans G. Hertz (1953) compared Hill's theory of Jupiter with observations that Hill had not used, inferring a correction to Saturn's mass in the process. With this change the theory was in agreement with observations to better than 1 arc-second most of the time, but occasionally the deviations were a little larger, though none ever exceeded 3 arc-seconds. At roughly the same time G. M. Clemence (1949, 1961) developed a theory of Mars that aimed to include all perturbational terms having an effect as large as 0.01 arc-seconds. At the end of this effort he expressed a worry that inaccuracies in Saturn could still be making a difference in his theory: "It is known that Hill's tables of Saturn fail to represent observations by several seconds of arc" (Clemence 1961, p. 266). He nevertheless expressed confidence that his theory for Mars generally agrees with observation to within a fraction of an arc-second.

[108] Smith (*forthcoming*) reviews the history of the evidence that the law must include the mass of the attracting body—not to be confused with the claim that weight is proportional to mass, and hence the law must include the mass of the attracted body. The only difference the attracting masses made in the orbital calculations, and even that has to be qualified, were in the two-body corrections to Kepler's third law and consequently to the major axes of the nominal Keplerian orbits. The largest difference any of these corrections had on any calculated longitude, 12.4 arc-seconds, was for Jupiter (when the Earth is at maximum elongation as seen from Jupiter), and no reliable theory for its motion was confirmed to agree with observation to this level before the twentieth century confirmation of Hill's theory. The next largest difference, for Saturn, is only 2.15 arc-seconds. The only other sources of empirical evidence for the mass of the attracting body belonging in the numerator of Newton's law were from physical geodesy—where the universality of gravity, particle-to-particle, is tested—and from Cavendish-type experiments in which the mass of the attracting body is varied systematically. Difficulties with evidence in physical geodesy prevented it from being decisive until the twentieth century. Hence, the only decisive evidence came from the several efforts on Cavendish-like experiments, starting in the second half of the nineteenth century and continuing into the twentieth. All this is described in my companion essay.

I can put this point in another way. When I say that the most immediate bearing the evidence had was on the accuracy of Newtonian theory, I do not necessarily mean the accuracy of Newton's law of gravity in all respects. What actually entered into the orbital calculations were various principles deduced from the law of gravity that concern specifics of orbital motions and are accordingly less general than it is.[109] The evidence from orbital motions in our solar system, making exception for the anomaly in the perihelion of Mercury, generally confirmed the accuracy of those principles. To be more specific than this, we would have to single out the principles and analyze the bounds of precision that the evidence determined for each of them individually. I have used the phrase "the accuracy of Newtonian theory" in an effort to capture the results of doing this in a summary fashion even though what the evidence really showed about the accuracy of the individual principles surely varied from principle to principle. This is important because these individual principles deduced from the law of gravity, and not the law itself, are what underpin the counterfactual claims about the differences any configurational detail makes.

In addition to questions of accuracy, the evidence before Einstein increasingly supported the projectibility of Newtonian theory, at least across the macroscopic domain of our solar system. Projectibility here involves two claims: (1) the principles derived from the law of gravity that were used in the calculations apply not just to the actual specific situations, but to a range of "possible" variants of them; and (2) the parameters entering into those principles, and no other parameters but these, delineate the kinds of changes that can make a difference to the orbital motions in our solar system—and hence they the delineate the relevant "possible" variants of the actual motions. Again I will speak of "the projectibility of the theory" in a summary fashion even though the issue concerns the evidence for the projectibility of different individual principles deduced from Newton's law.

Three aspects of the evidence before Einstein supported projectibility of the theory, thus understood. First, the emergence of further discrepancies with clear signatures provided evidence that the clear signatures of prior discrepancies were not mere coincidences. This is especially so when the further discrepancies would have been masked had the sources of prior discrepancies not been incorporated into the calculations. Second, successes in identifying robust physical sources for discrepancies provided evidence that the tighter agreement with observation then achieved by incorporating

[109] My choice of the term 'principles' here is intended to mirror Newton's use of the term for the 200 or more numbered "if-then" Propositions and the corollaries to them in Books 1 and 2 of the *Principia*. As Duhem stressed, one great virtue of a generic theory is that ever so many such more specific "if-then" claims can be derived from its fundamental principles, taken as axioms. Carl Hempel (1965, p. 436) emphasizes the lawlike character of such specific derived principles.

those sources into the calculation represents something more than a mere numerical representation of the data, as in a curve-fit; the parameters involved gained increasing claim to being ones that physically make a difference. And third, the overall sequence of successive approximations—addressing a variety of discrepancies involving different orbiting bodies, all with a single unifying theory and without having to backtrack and modify earlier stages—provided evidence that the principles and the parameters employed in the calculations are capable of characterizing the full range of physically possible variants of the actual motions.

Again a caveat is needed, though this time only a minor one. We now have clear reason to conclude that some orbital motions in our solar system are chaotic,[110] including the overall orbital motions of the planets in the long term.[111] "Chaos" here means that the calculated motions are infinitely sensitive to initial conditions. In the case of orbital motions in our solar system such sensitivity amounts to a *mathematical* fact about the equations of motion derived from gravitation theory for the set of orbits in question. Consequently, there can in principle be some systematic discrepancies between calculation and observation that stem not from unaccounted for forces, but from unavoidable limits in precision in specifying initial conditions. So, not every discrepancy that can in principle arise with the orbital motions in our solar system automatically fits into the scheme of evidence we are considering. Still, the time-scale of exponential divergence in the orbital motions that provided the evidence before Einstein is millions of years.[112] Chaotic motion has therefore been of no consequence in the evidence of concern here.

Beyond questions of accuracy and projectibility, the evidence before Einstein provided strong confirmation that the specific factors identified as the sources of the various second-order phenomena either really are the physical sources of these phenomena, or are at least close proxies for them. As a corollary, the evidence thus confirmed what amounted to wishful thinking for Newton: inverse-square centripetal acceleration fields dominate motions in our solar system. The evidence for the physical sources has been especially strong when ancillary consequences of a

[110] In particular, the motions of at least some asteroids. See Murray and Dermott (1999, pp. 456–466), and the motion of Pluto, ibid., pp. 466–468.

[111] Sussman and Wisdom (1992); and more recently Murray and Dermott (1999, pp. 469–471). Sussman and Wisdom (p. 62) make a comment at the end of their analysis that strongly supports the central claim I make in this essay about the underlying logic of theory-testing in celestial mechanics: "To positively conclude that the chaos observed in these long-term planetary integrations is not a result of numerical artifacts requires an unambiguous identification of a physical mechanism and an analytic evaluation to determine that the mechanism actually accounts for the observed chaos."

[112] Ibid.

proposed source—consequences that show it to be robust—have led to newly emerging second-order phenomena for which robust sources were then found. An example of this cited earlier was the discrepancy in the motion of Uranus that emerged only when the gravitational effects of Saturn on it were taken into account. The gravitational reach of Saturn had originally been put forward as the source of irregularities in the motion of Jupiter; the unmasking of the discrepancy in the motion of Uranus provided support for Saturn's reach long before the 900-year "Great Inequality" could be properly confirmed by observation.

The claim that the evidence strongly confirmed the proposed physical sources requires a word of clarification, if only to lessen the appearance of invoking a perspective of omniscience. To say that certain physical factors are the source of some phenomenon means only that those factors made the difference between this phenomenon occurring and its not occurring at all, or at least its occurring very differently from the way in which it does. But this amounts to asserting a counterfactual conditional: if the values of the parameters characterizing those factors had differed notably from what they are, then the phenomenon would not have occurred either at all or at least in the way it does. For example, if Jupiter and Saturn had instead been in concentric circular orbits, then their departures from Keplerian motion would not have exhibited anything remotely of the magnitude of the "Great Inequality." Evidence in support of claims that proposed physical sources make some specific differences thus amounts to evidence for certain counterfactual conditionals—or, if you prefer, for the comparatively specific principles derived from the law of gravity that lie behind those counterfactuals.

So much for what the evidence showed before the recalcitrant anomaly in Mercury's perihelion and the switch to Einstein's theory of gravity. What did the evidence, both prior and subsequent, show after this switch?

Consider first what it has shown about the accuracy of Newtonian theory. The three-hundred-year body of evidence gives clear grounds for concluding that Newton's law of gravity is nowhere exact. Instead, it is subject to a lower bound of imprecision that for orbital motions varies locally with the magnitude of GM/r^2c^2, where GM is the strength of the centripetal acceleration field governing the orbit. In other words, where comparison with observation gave upper bounds on the inaccuracy of Newton's theory of gravity, the switch to Einstein's theory and the evidence for it has given lower bounds.

Within our solar system, however, the lower bounds in question are generally well below the level of accuracy of the observations with which orbital calculations were being compared over the first three centuries after the *Principia*. Consequently, what the evidence from orbits showed about the accuracy within which Newton's theory holds remained the same, save for select exceptions, after the switch to Einstein's

theory. For example, Brown's 1903 conclusion about the accuracy of the exponent of r remained in place, for the calculated general relativistic effect on the precession of the lunar apogee is only 0.06 arc-seconds per century. The observed rate of this precession, once perturbing effects of other bodies are removed, can still serve as a preferred measure of the exponent. Only as advances like lunar laser-ranging yield extremely precise observations of the Moon's position will general relativistic corrections have to be made to the measured value.

This example, and the general conclusion that it illustrates, actually understate the consequences that the switch to Einstein's theory had on the empirical accuracy of Newton's theory. In some places the evidence for the specified level of accuracy of Newtonian claims became stronger after the switch. For example, the success of Einstein's theory not only with Mercury's perihelion, but also with the perihelia of the other three inner planets, has strengthened claims about the level of accuracy with which Newtonian theory gives the perturbing effects the other planets have on the perihelia. This is an example where the loop continued to be closed, yielding further reinforcing evidence in support of elements in the calculation introduced at an earlier stage.

The switch to Einstein's theory had more substantial consequences for what the evidence showed about the projectibility of Newtonian theory. The evidence following the switch has sharply circumscribed the domain over which the comparatively specific Newtonian principles presupposed in the orbital calculations apply to variants of the actual motions. The Newtonian principles remain projectible, but only over a domain defined ultimately by the level of precision of the observations with which the calculations are being compared. As argued in Section 10.2, even when the domain was taken to be universal, these principles were always, strictly speaking, being projected in their *quam proxime* form. The post-Einstein evidence has given numerical content to the "*quam proxime*" qualifier. In the process it has tied it to standards of observational accuracy when projecting the principles across the solar system, and it has placed tight constraints on projecting them outside this system.

There is another, perhaps better, way of thinking about what the switch to Einstein's theory did to the evidence for the projectibility of the Newtonian principles. The evidence never showed any of those principles to be projectible except in a *quam proxime* form. The evidence after the switch revealed that the Newtonian principles used in the calculations amount to surrogates for corresponding *quam proxime* principles of Einsteinian theory, principles that are not (yet) subject to the same constraints on projection. The Einsteinian principles could be substituted for the Newtonian in the calculations without altering any second-order phenomena at the level of observational precision with which these phenomena were specified. Viewed thus as surrogates for principles that have wider application, the Newtonian principles remain projectible because they approximate the Einsteinian principles over a restricted domain, where the restriction depends on a historically contingent level of precision of observation, and the Einsteinian principles drop this dependency.

Regardless of which of the these two ways one chooses to view their projectibility, the evidence has not ceased to show that the principles derived from Newton's law of gravity that were used in the calculations apply to a notable range of "possible" variants of the actual situations. Nor, correspondingly, has it ceased to show that the parameters entering into those principles delineate at least some of the kinds of changes that can make a difference in the orbital motions in our solar system. Indeed, the post-Einstein evidence has strengthened this claim about the parameters. For those same parameters—or at any rate parameters for which they are proxies—enter as well into the corresponding Einsteinian principles, though in some cases maybe requiring adjustments in the measurement processes through which they are assigned values. The only obvious departure is that the Einsteinian principles, as we saw earlier, contain at least one parameter, the speed of light, that does not enter into the Newtonian.

The switch from Newton's to Einstein's theory had the least effect on the evidence for the physical sources of Newtonian second-order phenomena. The counterfactual conditionals—*if the physical source had been otherwise, the second-order phenomena would not be what it is*—remain in place. This helps us see why the Newtonian physical sources remained in place after Einstein. The counterfactual conditionals in question describe situations only to a certain level of accuracy. At this level of accuracy the counterfactuals are equally supported by the Newtonian principles underlying them and their Einsteinian counterparts. The one impact the switch has had was to add such further possible physical sources of second-order phenomena as the curvature of space and the gravitational dilation of time. But in adding these, it has strengthened the evidence that various Newtonian physical sources have made the differences claimed for them. To provide an example, the post-Einsteinian success with the perihelia of the four inner planets has strengthened the evidence for the specific Newtonian gravitational effects the other planets have on those perihelia.

If anything, then, rather than undercutting the evidence for the conclusions reached before the switch to Einstein's theory, the switch has strengthened the evidence for them, at least once suitable qualifications are added about limits and conditions of precision. The strength of the evidence, however, varies from conclusion to conclusion. Each conclusion requires its own assessment of the strength of the evidence or, what amounts to the same thing, the stringency with which it has been tested over the course of the history of gravitational research on orbits. Such assessments involve just the sort of analyses one finds in review articles. The diagrams shown in Figures 10.6 and 10.8 depicting the structure of the evidence clarify what needs to be taken into consideration in any such analysis. In particular, more has to be taken into consideration than just the level of agreement with observation. The strength of the evidence for each claim depends most of all on the extent to which it has been presupposed in establishing further details that make a difference and the degree of robustness of these details.

10.4.3 Knowledge Achieved

The physical sources of Newtonian second-order phenomena turned out to be more resilient than Newton's theory. One might therefore naturally suggest that those physical sources have all along had more of a claim to being knowledge than the theory that enabled them to be identified. Let me spell this suggestion out more fully before assessing it.

What has the strongest claim to being knowledge coming out of gravity research are the configurational details that this research revealed as making specific detected differences in orbital motions in our solar system. The detected differences here include not only those that emerged historically, for which physical sources were then found, but also the discrepancies with clear signatures that emerge from counterfactually varying the parameters in the calculations in the manner discussed at the end of Section 10.4.1. The stronger the evidence is supporting any such detail and the differences it makes, the stronger is the claim they have to being knowledge. The claim to knowledge therefore turns on the number of times a detail and the differences it makes have been presupposed in successfully closing the loop, achieving still tighter agreement between observation and idealized calculations. So long as the details and the differences they make remain intact as the loop is closed, the claim to knowledge is not affected by the sort of asymptotically conservative theory-change that occurred with the transition from Newton's to Einstein's theory of gravity.

The trouble is, this proposal, as it stands, makes no sense. Conclusions that some detail makes some specific difference in motions in our solar system cannot be divorced from the theory that tells us it makes this difference. This is especially true in this case insofar as the research yielding such conclusions had no means of intervening to manipulate the details in order to see what happens. Any claim to knowledge that any of the details have must accordingly include a claim to knowledge, in some form or other, for the theoretical principles that dictate what differences those details are supposed to be making.

What has strong claim to knowledge coming out of the three centuries of gravity research on orbital motion in our solar system is accordingly better described as an interpenetration, so to speak, of theory and the details that make a difference and the differences they make. The details have provided both continuing evidence bearing on the theory and increasingly precise values for its many parameters. The theory has been indispensable for identifying which details make a difference and the differences they make, for it has provided the projectible principles underlying the counterfactual conditionals that have licensed the conclusions about what differences each detail does and does not make.

What then are we to say about Newton's theory of gravity and any claim to knowledge it has had, whether before or after Einstein's theory? To know that the details in question make the differences claimed for them does not require Newton's law of gravity or the principles derived from it underlying the relevant counterfactuals—or, for that matter, Einstein's field equations and the relevant principles derivable from them—to be the final word. They do not have to be exceptionless generalizations that hold universally. *They only have to hold to an appropriate level of approximation across a restricted domain that includes the actual situations of concern and enough of a range of variants of them to give the relevant counterfactuals their content.* The required level of approximation is dictated by the level of precision assumed in describing the differences the details are claimed to make. In terms of the other way of viewing projectibility offered above, the Newtonian—or the Einsteinian—principles have to be surrogates for principles, perhaps not yet recognized, that hold at least more exactly over a wider domain. Either way, the question whether (and when) we know that certain details make certain differences on the basis of Newtonian theory depends on whether (and when) we know that the relevant Newtonian principles meet these requirements.

I originally adopted the phrasing "details that make a difference" in an initial version of this essay in an effort to be non-committal with respect to the philosophical literature on causation. (I now suspect that I picked up the phrasing years ago from H. L. A. Hart and Tony Honoré, who were trying to be similarly neutral in their *Causation in the Law*.[113]) I still intend the phrase to be as neutral as possible. Nevertheless, what I mean by "a detail that makes a difference" is at least closely akin to, if not the same as, what James Woodward calls "a contributing cause" in his *Making Things Happen: A Theory of Causal Explanation*.[114] More importantly, the view of what is required of a generalization to support a claim that a detail makes a difference stated in the preceding paragraph is at least very close to the view Woodward maintains on the question throughout his book.[115] For a philosophic defense of my requirements I therefore suggest you turn to his book.

At the beginning of Section 10.4, I granted that knowledge involves permanence. The question whether, and when, we have known that the relevant Newtonian

[113] Hart and Honoré (1959, 1985). See in particular the summary statement: "Human action in simple cases, where we produce some desired effect by the manipulation of an object in our environment, is an interference in the natural course of events which *makes a difference* in the way these develop" (1959, p. 27; 1985, p. 29).

[114] See Woodward (2003, pp. 45–61, especially p. 59).

[115] See, in particular, ibid., ch. 6, pp. 239–314.

principles meet my requirements comes down to how strongly the evidence, at any time, has supported the permanence of the conclusion that those principles hold at least to the appropriate level of approximation for the relevant features of orbital motions in our solar system.

An obvious way to interpret "permanence" here is with respect to further theory-change in the future history of gravity research. Will future theories of gravity invariably retain the kind of conservative relationship to Newtonian theory that Einstein required of general relativity? Keep in mind that this calls for much more of future theories than just their achieving no less of an agreement between calculated and observed orbital motions in our solar system than has been achieved so far. Curve-fits can do this. The question is whether future theories will conserve all—or at least all but a minor few—of the Newtonian conclusions about which details make specific differences in the motions.

To do this, any future theory will have to include principles for which the relevant Newtonian principles are at least approximate, restrictedly projectible surrogates in the respects spelled out above. The evidence from the first two hundred years of gravity research gave Einstein overwhelming reasons to impose the constraint that general relativity so conserve Newtonian theory and certain conclusions about the physical world reached with it. The evidence that has emerged in support of both general relativity and Newtonian theory in the last century has given future researchers even stronger reasons for imposing such a constraint on their theorizing. So, I think we can safely say that every effort will be made in the future to conserve the Newtonian conclusions about which details make a difference to orbital motions in our solar system and what differences they make.

That, however, does not answer the question. For, future research may yield phenomena that cannot be handled by *any* theory unless this constraint is compromised or abandoned. The question is whether the evidence from the three centuries of gravity research has ever been sufficient to eliminate all possibility of future theories nullifying Newtonian conclusions about which details make a difference and what differences they make. I, for one, do not see how any empirical evidence could ever do that.

At the beginning of this essay the issue of alternative theories was posed not in terms of what existing evidence entails about future theory-change, but in terms of the extent to which that evidence can rule out other possible theories. Increased precision of data from astronomical observation, in significant part from statistical methods, and tighter agreement of gravity-based calculations with observation over the three centuries since Newton have eliminated a whole host of possible theories that could not have been eliminated at the time of the *Principia*. This process, moreover, is continuing, thanks to such technologies as radar-ranging and lunar laser-ranging on the observation side and computers that allow for the direct numerical integration of the simultaneous equations of motion for bodies in our solar system. Nevertheless, observation remains inexact, the calculations still have to make the idealizing assumption that no other forces need to be taken into account, and no way is ever available to eliminate all possibility of the existing data being somehow misleadingly parochial. Consider only the last of these three. So far none of the evidence in the history of gravity research has come from the microphysical realm, yet quantum theory has given clear reason to think that macroscopic laws cease to be adequate at microphysical scales.

The evidence for Newton's theory of gravity, as a matter of historical fact, did not eliminate the alternative theory of gravity of general relativity; nor is the evidence for that theory ever going to eliminate all logically possible alternatives to it. On the view developed in this essay, however, this Duhemian claim is not of much consequence so long as the alternative theories that have yet to be eliminated are ones that conserve the Newtonian (and Einsteinian) conclusions about which details in the world make a difference and what differences they make. The real question is whether the evidence has eliminated, or will ever eliminate, all theories that do not conserve those conclusions. Again, I do not see how it can.

Even so, consider the burden of proof the evidence has put on any theory that repudiates the Newtonian conclusions about which physical factors make detectable differences in our solar system. The evidence consists of sequences of progressively less obtrusive Newtonian second-order phenomena in which later members of the sequence presuppose not only the validity of the earlier members, but also the Newtonian physical sources identified for them. Put another way, these are sequences of increasingly smaller discrepancies between calculation and observation that have clear signatures which would be masked, or at least have less clear signatures, were it not for the earlier members and the physical sources identified for them. This evidence includes not just the historical sequences, like the one that culminated in Spencer Jones's fluctuation in the rotation of the Earth. It includes as well any such sequence that can be generated by exploring modified values of the parameters in the calculations and comparing the results with observation.

Any theory that repudiates Newtonian conclusions about which physical details make a difference in our solar system will have to provide grounds for saying that the degree of agreement with observation achieved by Newtonian calculations over the last three hundred years has been a mere accident, a subtly masked form of curve-fitting. But this is not all it will have to do. It will have to provide grounds for saying that the nested sequences of Newtonian second-order phenomena are mere coincidence, as is each case within any sequence in which a discrepancy with a clear signature emerges only after physical sources for preceding discrepancies have been incorporated into the calculations. This burden of proof is ever so much larger than the burden on alternative theories that do conserve the Newtonian conclusions. For, according to them neither the overall agreement with observation achieved by Newtonian calculations nor the sequences of Newtonian second-order phenomena have in any way been mere coincidence.

I said at the outset that much of the struggle of twentieth century philosophy of science in its efforts to characterize the nature, scope, and limits of the knowledge science can achieve has come from having to consider such matters in the light of the overthrow of Newtonian theory. The switch to Einstein's theory definitely did undercut any claim Newton's theory had to being the final word. But it did not undercut the claim to knowledge of the Newtonian conclusions about which details of our solar system make detectable differences to the orbital motions within it. To the contrary, the switch to Einstein's theory strengthened the claim to knowledge of these conclusions. That is the knowledge claim, even if only provisional, that the evidence coming out of three centuries of gravity research, properly analyzed, really supports. If my analysis of the evidence is correct, then philosophers and historians of science have indeed over-reacted in their response to the Einsteinian revolution.

10.5 CONCLUDING REMARKS

Anyone inclined to grant my analysis of evidence and knowledge in gravity research might naturally ask whether it extends to other areas of research, or maybe even to science generally.[116] My knee-jerk reaction is to reject this question as, at best, wildly premature. It violates everything I want my work in the philosophy and history of science to stand for. So far as I can see, whatever claim any area of science has to lofty epistemic status derives from the extent to which it constantly revisits previously accepted claims, bringing continuing evidence to bear on them that ideally subjects them to ever more stringent tests. Because of the rightful preoccupation science has with discovering new things about the world, any such further testing often takes place in silence. To assess the force of the continuing evidence therefore calls for careful, critical analysis of the history of evidence in it. I see this essay as a step in that direction for gravity research, and I hope it encourages others to examine the history of evidence in other areas. I definitely do not know enough about the history of evidence in other areas to begin assessing whether, or how, my analysis of gravity research might extend to them.

On top of this, there are respects in which gravity research has surely been atypical. Save for Eötvös-type experiments and the variants of Cavendish's experiment that began in the latter half of the nineteenth century, experimentation has contributed virtually nothing to gravity research, and even the experiments that did contribute

[116] Audiences put this question to me each of the five times I presented forerunners of this essay as talks.

were of no consequence in orbital research.[117] Experiments, needless to say, generally offer a more powerful way of developing evidence than passive observation, especially evidence about which factors make a difference and the differences they make. Experiments, however, often involve steps that complicate assessments of the force of that evidence. In particular, many experiments, including a surprisingly large fraction of the ones most heralded in the history of science, managed to yield well-behaved data only after trial-and-error revealed a need to introduce special provisions or restrictions in the experimental set-up. Often the reasons why those provisions or restrictions were needed did not become clear until decades later.

To give just one historically prominent example, Galileo's inclined-plane experiments gave well-behaved results only by restricting the angle of inclination of the plane to just a few degrees. Huygens, a half-century after Galileo published, appears to be the first to appreciate that a mixture of rolling and falling confounds the results at higher angles of inclination.[118] To what extent did the restriction to small angles reduce the force Galileo's evidence that acceleration is uniform in vertical free-fall? This sort of question arises with any experiment that involves *ad hoc* provisions or restrictions. The limited role of experiments in gravity research has freed it from such questions.

Those engaged in orbital research would gladly have lived with complications like these if they could have intervened. Only with space flight has it become possible. The inability to intervene, however, has not been the main evidential problem in gravity research on orbits. The one source of evidence, even after space flight, has been the orbital motions themselves, and these are extraordinarily complicated, opening the way to multiple representations of them to whatever level of precision is then current. Newton saw the possibility of turning those very complexities into a continuing source of evidence, provided the world turned out to be simple in one crucial respect, namely that the motions are predominately gravitational phenomena.

117 In the *Principia* Newton describes the first Eötvös-type experiment, the goal of which in his case was to show that weight of a body varies as its mass; he always thought of mass as inertial and therefore would never have asked whether gravitational mass is equal to inertial mass. All the Cavendish-like experiments, which are discussed in Smith (*forthcoming*), can be viewed as primarily aiming to measure the universal constant of gravity G; but in the process they ended up providing the first clear empirical evidence that the gravitational attraction toward a body varies as its mass. As noted earlier, the masses of celestial bodies are inferred from the gravitational attraction toward them, and these masses then enter into orbital calculations only in the form GM, defining the gravitational attractions toward them. So, the numerator in Newton's law of gravity has been almost entirely irrelevant to orbital calculations. This is why the experiments verifying the terms in it contributed virtually nothing to orbital research.

118 See Huygens (1929, p. 437). Domenico Bertoloni Meli first called my attention to this note of Huygens, who died before he published it.

There is, however, no reason to think that the main evidential problem in other areas of research lies in the complexity of the sources of evidence, and hence no reason to think that Newton's solution to the problem of evidence in gravity research applies elsewhere. In optics, for example, the problem is not the complexity of sources of evidence, but the complexity of light itself. Evidence has had to come from forcing light into extraordinary situations that diminish its complexity. This is a very different form of evidence from that of gravity research, where the phenomenon of gravity itself seems comparatively simple. Perhaps different areas of research face their own distinctive evidence problems, and the historical solutions to these problems have been largely peculiar to those areas. If so, then trying to generalize from the case of gravity research will likely just confuse matters. Better we should identify the fundamental obstacles to developing evidence in each area of research and ask how they have been surmounted historically before we attempt any general account.

In spite of these reasons not to generalize from the case of gravity research, I want to end with a thought for future consideration about science generally. Perhaps philosophy of science has created unnecessary difficulties for itself with the view that the knowledge science is ultimately pursuing is that of final theory. Maybe we would be better off with the view that each area of science is most of all pursuing knowledge of which factors make a difference within the domain it covers and what differences they make. One need not put undue weight on the practical applications of research in adopting this view. The pursuit of the details that make a difference in orbital motions in our solar system offered no practical gains to speak of before the space age. The real reason to take this view seriously is it better describes science as practiced over the last 400 years.

Consider, for example, a field like chemistry which has been constituted, at least until very recently, far less by theory than by laboratory practices. Such a large fraction of the information chemists have that has real claim to being knowledge consists of nuances in laboratory practice that were discovered through trial-and-error and have been handed down to the next generation through apprenticeship of one form or another, not through reading textbooks and papers. These nuances are very much details that make a difference. Philosophers of science have tended to treat chemistry as a different kind of science—that is, different from those branches of physics that have been dominated by theory—in large part because the sort of knowledge chemistry seems classically to have pursued does not fit the archetype supplied by such areas of physics as Newtonian gravitational mechanics. Chemistry does not seem nearly so different, however, if the archetype of knowledge being pursued centers on which details make a difference and the differences they make.

The same point can be made about those areas of research in engineering and medicine that are customarily referred to, pejoratively, as "empirical." As an engineer

I have worked in one such area for decades, cyclic-stress induced initiation and propagation of fatigue cracks in metals. At least in the foreseeable future, if not forever, there is no hope of theoretical physics shedding any quantitative light on such cracks. Instead, since the phenomenon first emerged with the advent of railroads in the nineteenth century, we have always relied on controlled laboratory testing of each different alloy to identify which factors make a difference and what differences they make.[119] Philosophers of science have almost no interest in a field like this, yet it has made impressive progress over its 150 years in identifying such factors. Taking this to be the knowledge that is being pursued in *all* areas of scientific research provides a basis for applying the word "science" across all of them.

In putting this thought forward, I by no means intend to diminish the importance of theory in science. There are strong *prima facie* reasons to think that areas of research that have managed to develop theories of the right sort have historically been more effective in establishing which details make a difference and what differences they make. One need look no further than orbital astronomy before and after Newton to see that illustrated. If we take the primary goal of science to be discovery of such details, the philosophic questions concerning theory of greatest interest become: What sorts of theory enable a field of research to become more effective in establishing details that make a difference and differences they make and, as in this essay, how does theory of this sort do it? I propose that the best way to answer such questions is through studies of the history of evidence in different areas of research, including ones not dominated by theory as well as ones that are.

ACKNOWLEDGMENTS

The first two sections of this essay began as a talk at a conference on "Assessing Evidence in Physics" put together by William Harper at the University of Western Ontario in May 2005. Revised versions of that talk were subsequently given at the University of Pittsburgh in November 2005, UCLA in October 2006, and as the Howard Stein

[119] Lest the example of fatigue fracture seem too recondite, the remarks made about it apply no less to the first of Galileo's "two new sciences," the strength of materials. A century after Galileo the key parameter, *stress*, had been identified by Cauchy and his contemporaries in France. But controversy remains to this day about how to combine biaxial tensile stresses with shear stresses to define an "effective stress" parameter governing instantaneous fracture. (See, for example, Dowling [1999], ch. 7.) As with fatigue, the ultimate tensile strength of each separate alloy (always as a function of temperature) has to be determined in laboratory experiments. The anomaly theoretical physics has for some time been trying to resolve is why these ultimate tensile strengths are typically an order of magnitude below the chemical bond strengths in the crystals forming the alloy.

Lecture in the Philosophy of Science at the University of Chicago a month later. The final version of that talk was given, under the title of this essay, as the first lecture in the series "Turning Data into Evidence: Three Lectures on the Role of Theory in Science" that I gave at Stanford University under the auspices of the Patrick Suppes Center for the Interdisciplinary Study of Science and Technology at Stanford University in February and March of 2007; those lectures remain available online at http://www.stanford.edu/dept/cisst/events0506.html. Too many people presented me with probing questions and remarks for me to try to list all of them here, but I should single out Michel Janssen, Eric Poisson, James McGuire, Sheldon Smith, Howard Stein, Michael Friedman, and several graduate students at Stanford.

Section 10.4 and the redirection of the talk to focus on the question of knowledge achieved with Newtonian gravity emerged when I began turning my Stanford lecture series into essays. The talk was not the one I gave at the conference on "Newton as Philosopher" organized by Andrew Janiak and Eric Schliesser in Leiden in 2007, but Eric urged me to publish it rather than an essay version of that talk in the book arising from the conference. Several people have provided helpful comments on earlier versions of the essay, most notably Michael Friedman, Chris Smeenk, and Kenneth Wilson, in addition to Janiak and Schliesser. I also benefited enormously from an encouraging discussion of the essay at a meeting of the Bay Area Philosophy of Science Discussion Group at the University of San Francisco in April 2009, where Bas van Fraassen objected to it far less than I expected him to. Most of all, however, I must acknowledge Eric Schliesser for urging me to stay with the questions addressed in this essay for two decades, and Curtis Wilson, for while I did not burden him with drafts of it, his writings on Newton and celestial mechanics were what forced those questions on me in the first place.

References

Adams, J. C. (1847). "An Explanation of the Observed Irregularities in the Motion of Uranus, on the Hypothesis of Disturbances Caused by a More Distant Planet; With a Determination of the Mass, Orbit, and Position of the Disturbing Body." *Memoirs of the Royal Astronomical Society* 16: 427–459.

——. (1853). "On the Secular Variation of the Moon's Mean Motion." *Philosophical Transactions of the Royal Society* 143: 397–406.

——. (1860). "Reply to Various Objections Against the Theory of the Secular Acceleration of the Moon's Mean Motion (with Postscript)." *Monthly Notices of the Royal Astronomical Society* 20: 225–240, 279–280.

——. (1880). "Investigation of the Secular Acceleration of the Moon's Mean Motion, Caused by the Secular Change in the Eccentricity of the Earth's Orbit." *Monthly Notices of the Royal Astronomical Society* 40: 211–223.

Adams, M. M. (1987). *William Ockham*. Notre Dame: University of Notre Dame Press.

Airy, S. G. B. (1834). *Gravitation*. London: Charles Knight.

Audoin, C., and B. Guinot. (2001). *The Measurement of Time: Time, Frequency, and the Atomic Clock.* Translated by S. Lyle. Cambridge, UK: Cambridge University Press.

Ball, R. (1893). *An Essay on Newton's "Principia."* London: Macmillan.

Baum, R., and W. Sheehan. (1997). *In Search of Planet Vulcan: The Ghost in Newton's Clockwork Universe.* New York: Plenum Trade.

Blanchet, L. (2001). "On the two-body problem in general relativity." *Comptes rendus de l'Académie des sciences* 2, Series 4: 1343.

Bouvard, M. A. (1821). *Tables astronomiques publiées par le Bureau des longitudes de France, contenant les tables de Jupiter, de Saturn et d'Uranus, construites d'après la théorie de la mécanique céleste.* Paris: Bachelier et Huzard.

Bowie, W. (1917). *U.S. Coast and Geodetic Survey, Special Publication* 40. Washington, DC: United States Government Printing Office.

Brown, E. W. (1903). "On the Verification of the Newtonian Law." *Monthly Notices of the Royal Astronomical Society* 64: 396–397.

——. (1919). *Tables of the Motion of the Moon.* 3 vols. New Haven, CT: Yale University Press.

——. (1926). "The Evidence for Changes in the Rate of Rotation of the Earth and Their Geophysical Consequences, with a Summary and Discussion of the Deviations of the Moon and Sun from Their Gravitational Orbits." *Transactions of the Astronomical Observatory of Yale University* 3: 207–235.

Brumberg, V. A. (1991). *Essential Relativistic Celestial Dynamics.* Bristol, UK: Adam Hilger.

Buchwald, J. Z., and G. E. Smith. (2002). "Incommensurability and Discontinuity of Evidence." *Perspectives on Science* 9 (4): 463–498.

Caroti, S. (2004). "Some Remarks on Buridan's Discussion on Intension and Remission." *Vivarium* 42: 58–85.

Ciufolini, I., and J. A. Wheeler. (1995). *Gravitation and Inertia.* Princeton, NJ: Princeton University Press.

Clemence, G. M. (1947). "The Relativity Effect in Planetary Motions." *Reviews of Modern Physics* 19: 361–364.

——. (1949). "First-order Theory of Mars." *Astronomical Papers Prepared for the Use of the American Ephemeris and Nautical Almanac* XI: 224–500.

——. (1961). "Theory of Mars—Completion." *Astronomical Papers Prepared for the Use of the American Ephemeris and Nautical Almanac* XVI: 265–333.

Cohen, I. B. (1999). "A Guide to Newton's Principia." In I. B. Cohen, A. Whitman, and J. Budenz, trans. *The Principia: Mathematical Principles of Natural Philosophy* by Isaac Newton. Berkeley: University of California Press, 1–370.

Cole, G. H. A. (2006). *Wandering Stars: About Planets and Exo-Planets, an Introductory Notebook.* London: Imperial College Press.

Darwin, G. H. (1907). *Scientific Papers.* Cambridge, UK: Cambridge University Press.

de Sitter, W. (1927). "On the Secular Accelerations and the Fluctuations of the Longitudes of the Moon, the Sun, Mercury, and Venus." *Bulletin of the Astronomical Institutes of the Netherlands* 4: 21–38, correction p. 70.

Dowling, N. E. (1999). *Mechanical Behavior of Materials: Engineering Methods for Deformation, Fracture, and Fatigue.* Upper Saddle River, NJ: Prentice Hall.

Duhem, P. (1982). *The Aim and Structure of Physical Theory.* Princeton, NJ: Princeton University Press.

Duncombe, R. L. (1956). "Relativity Effects for the Three Inner Planets." *The Astronomical Journal* 61: 174–175.

——. (1958). "Motion of Venus 1750–1949." In *Astronomical Papers Prepared for the Use of the American Ephemeris and Nautical Almanac*, Vol. 16, Part I. Washington, DC: United States Government Printing Office, 3–258.

Dunthorne, R. (1749). "A Letter from the Rev. Mr. Richard Dunthorne to the Reverend Mr. Richard Mason F. R. S. and Keeper of the Wood-Wardian Museum at Cambridge, Concerning the Acceleration of the Moon." *Philosophical Transactions of the Royal Society* 46: 162–172.

Einstein, A. (1995). *The Collected Papers of Albert Einstein.* Vol. 4, *The Swiss Years: 1912–1914.* Edited by M. J. Klein, A. J. Cox, J. Renn, and R. Schulmann. Princeton, NJ: Princeton University Press.

——. (1996a). *The Collected Papers of Albert Einstein.* Vol. 6, *The Berlin Years: Writings, 1914–1917.* Edited by A. J. Cox, M. J. Klein, and R. Schulmann. Princeton, NJ: Princeton University Press.

——. (1996b). *The Collected Papers of Albert Einstein.* Vol. 6, *The Berlin Years: Writings, 1914–1917 (English Translation Supplement).* Translated by A. Engel. Princeton, NJ: Princeton University Press.

Euler, L. (1769). "Recherches sur irrégularities du mouvement de Jupiter et de Saturne. Piéce qui remporté le Prix propose par 'Académie Royale des Sciences, pour l'année 1752." *Recueil des pièces qui ont remporté les prix de l'Académie des Sciences.* VII. Paris: Panckoucke. Available in The Euler Archive, Document 384, http://www.math.dartmouth.edu/~euler/.

——. (1777). "Réponse à la question proposée par l'académie royale des sciences de Paris, pour l'année 1772." The Euler Archive, Document 486, http://www.math.dartmouth.edu/~euler/. Reprinted in Leonhard Euler, Opera Omnia, Series 2, Vol. 24, pp. 167–190.

Forbes, E. G. (1971). *The Euler-Mayer Correspondence (1751–1755).* New York: American Elsevier.

Garavaglia, T. (1987). "The Runge-Lenz Vector and Einstein Perihelion Precession." *American Journal of Physics* 55: 164–165.

Glymour, C. (1980). *Theory and Evidence.* Princeton, NJ: Princeton University Press.

Hall, A. (1894). "A Suggestion in the Theory of Mercury." *The Astronomical Journal* 14: 49–51.

Hansen, P. A. (1838). *Fundamenta nova investigationis orbitae verae quam luna perlustrat.* Gotha, Germany: Cardum Glaeser.

——. (1857). *Tables de la lune construites d'après le principe newtonien de la gravitation universelle.* London: G. E. Eyre et G. Spottiswoode.

Harper, W. (2002). "Howard Stein on Isaac Newton: Beyond Hypotheses?" In D. B. Malament, ed. *Reading Natural Philosophy: Essays in the History and Philosophy of Science and Mathematics.* Chicago: Open Court, 71–112.

Hart, H. L. A., and T. Honoré. (1959). *Causation in the Law.* 1st ed. Oxford: Oxford University Press.

——. (1985). *Causation and the Law.* 2nd ed. Oxford: Oxford University Press.

Heiskanen, W. (1921). "Über den Einfluss der Gezeiten auf die säkuläre Acceleration des Mondes." *Annales Acadamiae Scientiarum Fennicae* 18: 1–84.

Hempel, C. G. (1965). "Aspects of Scientific Explanation." In *Aspects of Scientific Explanation and Other Essays in the Philosophy of Science.* New York: The Free Press, 331–496.

——. (1988). "Provisos: A Problem Concerning the Inferential Function of Scientific Theories." In A. Grünbaum and W. G. Salmon, eds. *The Limitations of Deductivism.* Berkeley: University of California Press, 19–36.

Hertz, H. G. (1953). "The Mass of Saturn and the Motion of Jupiter 1884–1948." *Astronomical Papers Prepared for the Use of the American Ephemeris and Nautical Almanac* XV: 169–215.

Hill, G. W. (1905). "On the Part of the Motion of the Lunar Perigee Which Is a Function of the Mean Motions of the Sun and Moon." In *Collected Mathematical Works of George William Hill*. Vol. 1. Washington, DC: Carnegie Institution of Washington, 243–270.

Huygens, C. (1690). "Discours de La Cause de La Pesanteur." In *Traité de la Lumière avec Discours de La Cause de La Pesanteur*. Leiden, The Netherlands: Pierre Vander, 125–180. A facsimile edition was published in Brussels by Culture et Civilisation, 1967.

——. (1929). *Oeuvres complètes de Christiaan Huygens*. Vol. 16. The Hague: Martinus Nijhoff.

——. (1944). *Oeuvres complètes de Christiaan Huygens*. Vol. 21. The Hague: Martinus Nijhoff.

Janssen, M. (2002). "COI Stories: Explanation and Evidence in the History of Science." *Perspectives on Science* 10: 457–522.

——. (2007). "What Did Einstein Know and When Did He Know It? A Besso Memo Dated August 1913." In J. Renn, ed. *The Genesis of General Relativity. Vol. 2, Einstein's Zurich Notebook, Commentary and Essays*. Dordrecht, The Netherlands: Springer, 785–837.

Janssen, M., and J. Renn. (2007). "Untying the Knot: How Einstein Found His Way Back to Field Equations Discarded in the Zurich Notebook." In J. Renn, ed. *The Genesis of General Relativity. Vol. 2, Einstein's Zurich Notebook, Commentary and Essays*. Dordrecht, The Netherlands: Springer, 839–925.

Jeffreys, H. (1916). "The Secular Perturbations of the Four Inner Planets." *Monthly Notices of the Royal Astronomical Society* 77: 112–118.

——. (1918). "The Secular Perturbations of the Inner Planets." *Philosophical Magazine* 36: 203–205.

——. (1920). "Tidal Frictions in the Shallow Seas." *Philosophical Transactions of the Royal Society, Series A* 221: 239–264.

——. (1924). *The Earth: Its Origin, History and Physical Constitution*. 1st ed. Cambridge, UK: Cambridge University Press.

——. (1937). "On the Figures of the Earth and Moon." *Monthly Notices of the Royal Astronomical Society* 97: 3–15.

——. (1941). "On the Figures of the Earth and Moon." *Monthly Notices of the Royal Astronomical Society* 101: 34–36.

——. (1948). "On the Figures of the Earth and Moon." *Monthly Notices of the Geographical Society* 5: 219–247.

——. (1976). *The Earth: Its Origin, History and Physical Constitution. 6th ed.* Cambridge, UK: Cambridge University Press.

Jones, H. S. (1926). "The Rotation of the Earth." *Monthly Notices of the Royal Astronomical Society* 87: 4–31.

——. (1932). "Discussion of Observations of Occultations of Stars by the Moon, 1672–1908, Being a Revision of Newcomb's 'Researches on the Motion of the Moon, Part II.'" *Annals of the Cape Observatory* 13, Part 3, 3–70.

——. (1939). "The Rotation of the Earth, and the Secular Acceleration of the Sun, Moon, and Planets." *Monthly Notices of the Royal Astronomical Society* 99: 541–558.

Kant, I. (1900). "Untersuchung der Frage, ob die Erde in ihrer Umdrehung um die Achse… einige Veränderung seit den ersten Zeiten ihres Ursprungs erlitten habe." In Immanuel Kant: *Gesammelte Schriften*. Vol. 1. Berlin: Akademie, 183–191.

Koyré, A. (1968). *Newtonian Studies*. Chicago: University of Chicago Press.

Kriesel, G. (1949). "Some Remarks on Integral Equations with Kernels: L(ξ1-x1,...,ξn-xn;α)." *Proceedings of the Royal Society of London*, Series A 197 (1049): 160–183.

Kuhn, T. (1970). *The Structure of Scientific Revolutions. 2nd ed.* Chicago: University of Chicago Press.

Lambeck, K. (1980). *The Earth's Variable Rotation: Geophysical Causes and Consequences*. Cambridge, UK: Cambridge University Press.

Laplace, P. -S. (1786–1788). "Sur l'équation séculaire de la lune." *Mémoires de l'Académie Royale de Sciences de Paris*, 235–264.

——. (1966). *Celestial Mechanics*. Translated by N. Bowditch. New York: Chelsea Publishing Company.

Leverrier, U. J. J. (1859). "Théorie du mouvement de Mercure." *Annales de l'Observatoire de Paris* 5: 1–196.

Lykawka, P. S., and T. Mukai. (2008). "An Outer Planet beyond Pluto and the Origin of the Trans-Neptunian Belt Architecture." *The Astronomical Journal* 135: 1161–1200.

Mayer, J. T. (1753). "Novae tabulae motuum solis et lunae." *Commentarii Societatis Regiae Scientarium Gottingensis* 2: 383–430.

Misner, C. W., K. S. Thorne, and J. A. Wheeler. (1973). *Gravitation*. New York: Freeman.

Morando, B. (1995). "The Golden Age of Celestial Mechanics." In R. Taton and C. Wilson, eds. *Planetary Astronomy from the Renaissance to the Rise of Astrophysics, Part B: The Eighteenth and Nineteenth Centuries*. Cambridge, UK: Cambridge University Press, 211–239.

Munk, W. H., and G. J. F. MacDonald. (1960). *The Rotation of the Earth: A Geophysical Discussion*. Cambridge, UK: Cambridge University Press.

Murray, C. D., and S. F. Dermott. (1999). *Solar System Dynamics*. Cambridge, UK: Cambridge University Press.

Newcomb, S. (1876). *Investigation of Corrections to Hansen's Tables of the Moon; with Tables for the Application*, Forming Part III of Papers Published by the Commission on the Transit of Venus. Washington, DC: United States Government Printing Office.

——. (1878). "Researches on the Motion of the Moon: Part I. Reduction and Discussion of Observations of the Moon before 1750." In *Washington Observations of 1875*. Appendix II. Washington, DC: United States Government Printing Office.

——. (1882). "Discussion and Results of Observations on Transits of Mercury, from 1677 to 1881." *Astronomical Papers Prepared for the Use of the American Ephemeris and Nautical Almanac*. Vol. 1. Washington, DC: United States Government Printing Office, 367–487.

——. (1895). *The Elements of the Four Inner Planets and the Fundamental Constants of Astronomy, Supplement to the American Ephemeris and Nautical Almanac for 1897*. Washington, DC: United States Government Printing Office.

——. (1898). *Tables of the Four Inner Planets. Vol. 6 of Astronomical Papers Prepared for the Use of the American Ephemeris and Nautical Almanac*. Washington, DC: Bureau of Equipment, US Navy Department.

——. (1903). "On the Desirableness of a Re-investigation of the Problems Growing Out of the Mean Motion of the Moon." *Monthly Notices of the Royal Astronomical Society* 63: 316–324.

——. (1907). *Investigation of Inequalities in the Motion of the Moon Produced by the Action of the Planets*. Assisted by Frank E. Ross. Carnegie Institution of Washington Publication No. 72. Washington, DC: Carnegie Institution of Washington.

——. (1909). "Fluctuations in the Moon's Mean Motion." *Monthly Notices of the Royal Astronomical Society* 69: 164–169.

——. (1912). "Researches on the Motion of the Moon: Part II. The Mean Motion of the Moon and Other Astronomical Elements Derived from Observations of Eclipses and Occultations Extending from the Period of the Babylonians Until A.D. 1908." *Astronomical Papers Prepared for the Use of the American Ephemeris and Nautical Almanac* 9, Part I: 5–249.

Newton, I. (1959–1977). *The Correspondence of Isaac Newton*. Edited by H. W. Turnbull, J. F. Scott, A. R. Hall, and L. Tilling. 7 vols. Cambridge, UK: Cambridge University Press.

——. (1702). "Theoria Lunae." In D. Gregory. *Astronomiae Physicae & Geometricae Elementa*. Oxford. (English Edition, London, 1715.) The English version published as a pamphlet in 1702 and reprinted in facsimile in *Isaac Newton's Theory of the Moon's Motion (1702)*. Introduction by I. B. Cohen. Folkstone, UK: Dawson, 1975.

——. (1962). *Unpublished Scientific Papers of Isaac Newton*. Edited by A. R. Hall and M. B. Hall. Cambridge, UK: Cambridge University Press.

——. (1999). *The Principia: Mathematical Principles of Natural Philosophy*. Translated by I. B. Cohen, A. Whitman, and J. Budenz. Berkeley: University of California Press.

Norton, J. (1984). "How Einstein Found His Field Equations: 1912–1915." *Historical Studies in the Physical Sciences* 14: 253–316.

Ohanian, H. C., and R. Ruffini. (1994). *Gravitation and Spacetime. 2nd ed.* New York: Norton.

Rindler, W. (1977). *Essential Relativity: Special, General, and Cosmological*. New York: Spring-Verlag.

Roseveare, N. T. (1982). *Mercury's Perihelion: From le Verrier to Einstein*. Oxford: Oxford University Press.

Ross, F. E. (1917). "New Elements of Mars." *Astronomical Papers Prepared for the Use of the American Ephemeris and Nautical Almanac* 9, Part II, 255–274.

Schliesser, E., and G. E. Smith. (*forthcoming*). "Huygens's 1688 Report to the Directors of the Dutch East India Company on the Measurement of Longitude at Sea and the Evidence It Offered Against Universal Gravity." *Archive for History of Exact Sciences*.

Seidelmann, P. K., ed. (1992). *Explanatory Supplement to the Astronomical Almanac*. Sausalito, CA: University Science Books.

Smeenk, C., and G. E. Smith. (in preparation). "Newton on Constrained Motion: A Commentary on Book 1, Section 10 of Newton's *Principia*."

Smith, G. E. (*forthcoming*). "Pending Tests to the Contrary: The Question of Mass in Newton's Law of Gravity."

——. (2001a). "J. J. Thomson and the Electron: 1897–1899." In Jed Z. Buchwald and A. Warwick, eds. *Histories of the Electron: The Birth of Microphysics*. Cambridge, MA: MIT Press, 21–76.

——. (2001b). "The Newtonian Style in Book II of the *Principia*." In Jed Z. Buchwald and I. B. Cohen, eds. *Isaac Newton's Natural Philosophy*. Cambridge, MA: MIT Press, 249–312.

——. (2002). "From the Phenomenon of the Ellipse to an Inverse-Square Force: Why Not?" In D. Malament, ed. *Reading Natural Philosophy: Essays in the History and Philosophy of Science and Mathematics to Honor Howard Stein on His 70th Birthday*. La Salle, IL: Open Court, 31–70.

——. (2007). "Turning Data into Evidence: Three Lectures on the Role of Theory in Science." Delivered at Stanford University, Feb–Mar 2007. Available at: http://www.stanford.edu/dept/cisst/events0506.html.

Stachel, J. (2002). "Einstein's Search for General Covariance, 1912–1915." In *Einstein from "B" to "Z"*. Boston: Birkhäuser, 301–337.

Standish, E. M., Jr. (1993). "Planet X: No Dynamical Evidence in the Optical Observations." *The Astronomical Journal* 105: 2000–2006.

Stephenson, F. R. (1997). *Historical Eclipses and Earth's Rotation*. Cambridge, UK: Cambridge University Press.

Sussman, G. J., and J. Wisdom. (1992). "Chaotic Evolution of the Solar System." *Science* 257: 56–62.

Taylor, G. I. (1920). "Tidal Friction in the Irish Sea." *Philosophical Transactions of the Royal Society, Series A* 220: 1–33.

Tisserand, F. (1889). *Traité de la Mécanique Céleste*. Paris: Gauthier-Villars.

Truesdell, C. (1968). "Whence the Law of Moment and Momentum." In *Essays in the History of Mechanics*. New York: Springer-Verlag, 239–271.

van Helden, A. (1989). "The Telescope and Cosmic Dimensions." In R. Taton and C. Wilson, eds. *Planetary Astronomy from the Renaissance to the Rise of Astrophysics, Part A: Tycho Brahe to Newton*. Cambridge, UK: Cambridge University Press, 106–118.

Waff, C. B. (1995). "Clairaut and the Motion of the Lunar Apse: The Inverse-Square Law Undergoes a Test." In R. Taton and C. Wilson, eds. *Planetary Astronomy from the Renaissance to the Rise of Astrophysics, Part B: The Eighteenth and Nineteenth Centuries*. Cambridge, UK: Cambridge University Press.

Weinberg, S. (1972). *Gravitation and Cosmology: Principles and Applications of the General Theory of Relativity*. New York: Wiley.

——. (1993). *Dreams of a Final Theory: The Scientist's Search for the Ultimate Laws of Nature*. New York: Vintage Books.

——. (2001). "The Non-Revolution of Thomas Kuhn." In *Facing Up: Science and Its Cultural Adversaries*. Cambridge, MA: Harvard University Press, 187–206. Originally published as "The Revolution That Didn't Happen" in *The New York Review of Books*, October 8, 1998, 45: 15.

Whittaker, E. T. (1904). *A Treatise on the Analytical Dynamics of Particles and Rigid Bodies*. 4th ed. 1947, Cambridge, UK: Cambridge University Press.

Will, C. M. (1993). *Theory and Experiment in Gravitational Physics. Rev. ed.* Cambridge, UK: Cambridge University Press.

Wilson, C. (1980). "Perturbations and Solar Tables from Lacaille to Delambre: The Rapprochement of Observation and Theory." *Archive for History of Exact Sciences* 22: 53–304.

——. (1985). "The Great Inequality of Jupiter and Saturn: From Kepler to Laplace." *Archive for History of Exact Sciences* 33: 15–290.

——. (1989). "Predictive Astronomy in the Century after Kepler." In R. Taton and C. Wilson, eds. *Planetary Astronomy from the Renaissance to the Rise of Astrophysics, Part A: Tycho Brahe to Newton*. Cambridge, UK: Cambridge University Press, 161–207.

——. (2002). "Newton and Celestial Mechanics." In I. B. Cohen and G. E. Smith, eds. *The Cambridge Companion to Newton*. Cambridge, UK: Cambridge University Press, 202–226.

Woodward, J. (2003). *Making Things Happen: A Theory of Causal Explanation*. Oxford: Oxford University Press.

Newton and Empiricism

Edited by
Zvi Biener
and
Eric Schliesser

OXFORD
UNIVERSITY PRESS

OXFORD
UNIVERSITY PRESS

Oxford University Press is a department of the University of Oxford.
It furthers the University's objective of excellence in research, scholarship,
and education by publishing worldwide.

Oxford New York
Auckland Cape Town Dar es Salaam Hong Kong Karachi
Kuala Lumpur Madrid Melbourne Mexico City Nairobi
New Delhi Shanghai Taipei Toronto

With offices in
Argentina Austria Brazil Chile Czech Republic France Greece
Guatemala Hungary Italy Japan Poland Portugal Singapore
South Korea Switzerland Thailand Turkey Ukraine Vietnam

Published in the United States of America by
Oxford University Press
198 Madison Avenue, New York, NY 10016

Cataloging in Publication data on file with the Library of Congress

ISBN 978-0-19-933709-5

1 3 5 7 9 8 6 4 2
Printed in the United States of America
on acid-free paper

CONTENTS

LIST OF ILLUSTRATIONS

Figures

Tables