THE DEBATE BETWEEN WHEWELL AND MILL ON THE NATURE OF SCIENTIFIC INDUCTION

Malcolm Forster

1 WHY THE DEBATE IS NOT MERELY TERMINOLOGICAL

The very best examples of scientific induction were known in the time of William Whewell (1994–1866) and John Stuart Mill (1806–1873). It is puzzling, therefore, that there was such a deep disagreement between them about the nature of induction. It is perhaps astounding that the dispute is unresolved to this very day!

What disagreement could there be about Newton's discovery of universal gravitation? Prior to Newton, it was well known that gravity acts on objects near the Earth's surface, and Copernicus even speculated that the planets have a spherical shape because they have their own gravity. But Newton was the first to understand that it's the Earth's gravity that keeps the Moon in orbit around the Earth, and that the Sun's gravity keeps the Earth and the Moon in orbit around the Sun. At the root of this discovery was Newton's explication of the kinematical concept of acceleration. To understand that the Moon (just like the fabled apple) is pulled by the Earth, one has to understand that the Moon is accelerating towards the Earth even if it is moving uniformly on the circular orbit around the Earth. Acceleration must not be defined as the time rate of change of speed, but as the time rate of change of velocity, where velocity has direction as well as magnitude. Thus, the Moon is accelerating towards the Earth because its velocity is changing its direction. Galileo, on the other hand, worked with a *circular* law of inertia, according to which uniform circular motion around the Earth was a "natural" motion that required no force.

Further explication of the new conception of acceleration led Newton to discover that if the line from a point O to a body B sweeps out equal areas (Kepler's second law), then B is accelerating towards O. If, in addition, the body follows an elliptical path with O at one focus (Kepler's first law), then the acceleration towards O is inversely proportional to the square of the distance of B from O. In the case of the planets moving around the sun, if we assume that the constant of proportionality is the mass of the sun, then Kepler's third law follows as well. Thus, Newton's new conception of acceleration causes Kepler's three laws to "jump together" in a way that tests the conceptions that Kepler had previously employed, involving ellipses,

areas swept out by the line *OB*, the mean length of that line, and its period of revolution around the sun.

For Whewell, the addition of the conceptions in each of these inductions is the defining characteristic of induction. Whewell introduced a new term for the process of binding the 'facts' by a new conception. He called it the *colligation of facts*, and used this phrase interchangeably with the word 'induction'. Mill reacted negatively to this 'improper' use of the term. Mill agreed that new conceptions are often applied to the 'facts' during an induction, but he insisted that they are not part of the induction, and certainly not a defining characteristic. For Mill, induction consisted in extrapolating or interpolating a regularity from the known instances to the unknown instances, as is classically the case in examples of simple enumerative induction such as: All observed swans are white; therefore all swans are white. Whewell agreed that interpolation and extrapolation does, in general, result from a colligation of facts, but should not be the property that defines induction.

It is tempting at this point to dismiss the debate as merely terminological. Whewell has an unusual conception of what induction is, but once it is taken on board, it is possible to translate between the two vocabularies. I agree that there is a large terminological component in the debate, but I insist that it is not *merely* terminological. Behind the difference in terminology is a very deep disagreement about the objectivity of human knowledge. Mill and Whewell both want to defend the objectivity of human knowledge. But they have quite distinctive views on how it comes about, and Whewell's idea is interesting and new.

Mill is entrenched in the rather extreme empiricist view that human knowledge is objective because it is built on an objective foundation of empirically given statements from which higher claims are inferred using the objective canons of inductive reasoning. Human knowledge maintains its objectivity (to the extent that it succeeds) by minimizing the influences of subjective elements at every stage of the process.

For Whewell, subjective and objective elements are inseparable parts of human knowledge at any level in the hierarchy of knowledge, from the concept-ladenness of perceptual knowledge at the bottom, to the concept-ladenness of the highest forms of scientific knowledge at the top. The counter-proposal is that empirical success at the higher levels of knowledge, captured in terms of what he called the *consilience of inductions*, can help to secure the lower levels as a kind of bootstrapping effect. For example, Kepler's colligations of facts are concept-laden in a way that makes them subjective at first, but once Newton used the new conception of force and acceleration to show how the facts, *described in terms of Kepler's colligations*, lead successfully to a higher level colligation of facts, then the subjective elements involved are successfully "objectified". Knowledge is like a building in which the addition of higher floors helps strengthen the lower levels.

Whewell harbored a deep distain for Mill's purely empiricist philosophy, which he saw as constantly downplaying the importance of the subjective component of knowledge, or as trying to reduce it to purely empirical elements at every stage.

In contrast, the conceptual components of knowledge are, for Whewell, the very instruments that ultimately explain how human knowledge is possible. They produce the colligations that may be confirmed by the consiliences of colligations, which serves to objectify the subject elements, making knowledge possible. The introduction of new conceptions in the colligation of facts is therefore a defining characteristic of induction.

There is a major problem in trying to understand the Whewell-Mill debate from what the authors wrote. Whewell was primarily a historian of science, but Mill did not have a good knowledge of the history of science. Whewell allowed Mill to center the debate on particular examples of induction such as Kepler's inference that Mars moves on an ellipse. They got so tied up in that example, that the larger philosophical differences got lost in the discussion. It's possible that Whewell's hierarchical view of knowledge led him to believe that the bigger picture is played out in smaller examples on a smaller scale. Unfortunately, Whewell did not recall the details of the Kepler example in sufficient detail bring out those features of it. In section 2, I attempt to remedy that problem by describing the Kepler example in a way that challenges Mill's picture of it. Section 3 turns to Whewell's bigger picture by discussing his tests of hypotheses, while section 4 argues that Whewell-Mill debate helps us understand why sophisticated methods of induction have not been programmed to run automatically on a computer. Finally, section 5 asks whether the Whewell-Mill debate may help us identify fundamental limitations in the scope of Bayesian and Likelihoodist theories of evidence and confirmation.

2   THE KEPLER EXAMPLE AND THE COLLIGATION OF FACTS

The colligation of facts was Whewell's name for scientific induction. Its defining characteristic is the introduction of a new conception not previously applied to the data at hand, which unites and connects the data. In curve fitting, the idea is easy to visualize. According to Whewell, "the Colligation of ascertained Facts into general Propositions" consists of (1) the *Selection of the Idea*, (2) the *Construction of the Conception*, and (3) the *Determination of the Magnitudes*. In curve fitting, these three steps correspond to (1) the determination of the *Independent Variable*, (2) the *Formula*, and (3) the *Coefficients*. Once the variables are chosen (Step 1), one chooses a particular functional relationship (Step 2; choose the Formula, Conception, family of curves) characterized in terms of some adjustable parameters (which Whewell calls coefficients), and then one fits the curves to the data in order to estimate the values of the parameters (Step 3; determining the magnitude of the coefficients).

Consider the simplest possible example. Suppose we hang an object on a beam balance in order to infer its mass from the distance at which a unit mass must be slid along the beam to counterbalance the object in question. If the units are chosen appropriately, and the device is built well, then the mass value can be read straight from the distance at which the unit weight balances the beam. The dependent variable chosen in step 1 of the colligation of facts is x (there is no

independent variable), and the family of "curves" or the formula chosen in step 2 is $x = m$, where x is the distance of the unit mass from the fulcrum and m is an adjustable parameter, which represents the mass of the object in question. Whewell's third step in the colligation of facts refers to the determination of the mass values by inferring them from the x values using the formula. The conception being introduced is the formula $(\forall o)(x(o) = m(o))$, where 'o' ranges over a set of objects. The formula is something added to or imposed upon the facts by mind of the investigator; it is not contained in, or read from, those facts. Of course, the *magnitude* of the mass is read from the facts; indeed, this is the third step in Whewell's colligation of facts. But that does not mean that the formula itself is determined by the facts.

The underdetermination implies that the subjective elements in the colligation of facts make the inductive conclusions uncertain and conjectural. In order to defend the objectivity of our knowledge, we have two choices. We can choose the Millean strategy of denying that there is ever any such underdetermination, or go for the Whewellian strategy of allowing that the consilience of inductions can later test the conjecture, and upgrade its confirmational status in light of this higher-level empirical success. To take our hindsight wisdom for granted, as Mill does, and to suppose that the initial induction had this status all along, is to commit the kind of error that non-historians often make.

Though our beam balance example is not a real piece of history, the Millean mistake in that example would be to take the agreement of spring balance measurements of mass and beam balance measurements of mass for granted, and to assume that the justification for postulating 'mass' already existed prior to the consilience.

Unfortunately, the debate centers around the Kepler example and neither author gives the details of this important example in sufficient detail for the purpose at hand. It is especially confusing because Mill held the very strange and rather complicated view that Kepler did not perform any induction at all, even in the very broad sense in which Mill uses the term.

Mill's strategy is to make a distinction between a description and an explanation, and to argue that the inductive conclusion in the Kepler example is merely a description of the data, and therefore, there was no induction performed by Kepler. For example, in his view, when the ancients hypothesized that the planets move by being embedded on crystalline spheres, they put forward an *explanation* of celestial motions. But when Ptolemy and Copernicus conceived of the motions in terms of the combinations of circles, they were merely putting forward a description. In Mill's words:

> When the Greeks abandoned the supposition that the planetary motions were produced by the revolution of material wheels, and fell back upon the idea of "mere geometrical spheres or circles," there was more in this change of opinion than the mere substitution of an ideal curve for a physical one. There was the abandonment of a theory, and the replacement of it by a mere description. No one would think of call-

> ing the doctrine of material wheels a mere description. That doctrine
> was an attempt to point out the force by which the planets were acted
> upon, and compelled to move in their orbits. But when, by a great
> step in philosophy, the materiality of the wheels was discarded, and
> the geometrical forms alone retained, the attempt to account for the
> motions was given up, and what was left of the theory was a mere
> description of the orbits. [Mill 1872, Book III, Chapter ii, section 4]

It's true that no one would think of calling the doctrine of material wheels a
mere description. But it is very strange that Mill should insist that it becomes a
mere description as soon as the materiality of the wheels is discarded. For these
"mere descriptions" entail predictions that are not part of the data, and anything
that goes beyond the data goes from the known to the unknown should therefore
count as an induction, according to Mill's own definition. Thus, even if Kepler's
conclusion were a mere description, in the sense that Mill has just described, it
should not disqualify Kepler's inference as counting as an induction.

In order to be as charitable as possible to Mill, let me begin with the example
that he presents as the clearest in his favor. It is about the circumnavigation of
an island:

> A navigator sailing in the midst of the ocean discovers land: he cannot
> at first, or by any one observation, determine whether it is a continent
> or an island; but he coasts along it, and after a few days finds himself
> to have sailed completely round it: he then pronounces it an island.
> Now there was no particular time or place of observation at which he
> could perceive that this land was entirely surrounded by water: he
> ascertained the fact by a succession of partial observations, and then
> selected a general expression which summed up in two or three words
> the whole of what he so observed. But is there anything of the nature
> of an induction in this process? Did he infer anything that had not
> been observed, from something else which had? Certainly not. He had
> observed the whole of what the proposition asserts. That the land in
> question is an island, is not an inference from the partial facts which
> the navigator saw in the course of his circumnavigation; it is the facts
> themselves; it is a summary of those facts; the description of a complex
> fact, to which those simpler ones are as the parts of a whole. [Mill,
> 1872, Book III, ch. ii, section 3]

Astonishingly, even in this example, Mill's case is very weak. For if we think
carefully about what is *observed* in this example, it is the *similarity* of the view
of the shoreline at the start and the end of the circumnavigation. The views are
not *exactly* the same because the distance from the shore is different, the tides
are different, and the times of day are different. It is not given *in the facts* that
the views are *of the same shoreline*. That is a *conclusion*. The hypothesis that
an island has been circumnavigated *explains* why the views look similar. That
the conclusion is inductive is made plain by the fact that it makes a *prediction*,

which may be false. For it predicts that if we continue sailing further in the same direction, then we will see an ordered sequence of previously seen views of the shoreline. It is puzzling that Mill does not see this; he clearly defines induction, in his terms, as any inference from the known to the unknown. Perhaps he sees the logical gap as small in this case. But it gets much larger in the Kepler example because it is not merely a circumnavigation that is inferred, but also the exact path (Kepler's first law) and rate of motion (Kepler's area law). The puzzle is resolved a little once we look more carefully at Mill's description of the Kepler example. He continues from the previous passage.

> Now there is, I conceive, no difference in kind between this simple operation [in the island example], and that by which Kepler ascertained the nature of the planetary orbits: and Kepler's operation, all at least that was characteristic in it, was not more an inductive act than that of our supposed navigator.

> The object of Kepler was to determine the real path described by each of the planets, or let us say the planet Mars (since it was of that body that he first established the two of his three laws which did not require a comparison of planets). To do this there was no other mode than that of direct observation: and all which observation could do was to ascertain a great number of the successive places of the planet; or rather, of its apparent places. That the planet occupied successively all these positions, or at all events, positions which produced the same impressions on the eye, and that it passed from one of these to another insensibly, and without any apparent breach of continuity; thus much the senses, with the aid of the proper instruments, could ascertain. What Kepler did more than this, was to find what sort of a curve these points would make, supposing them to be all joined together. He expressed the whole series of the observed places of Mars by what Dr. Whewell calls the general conception of an ellipse. This operation was far from being as easy as that of the navigator who expressed the series of his observations on successive points of the coast by the general conception of an island. But it is the very same sort of operation; and if the one is not an induction but a description, this must also be true of the other. [Mill, 1872, Book III, ch. ii, section 3]

Mill's first naiveté is his passing from "the successive *apparent* places of the planet" to "the successive places of the planet", as if there is no important gap between the 3-dimensional positions of Mars and the angular position of Mars relative to the fixed stars. Then, without any additional argument, Mill simply affirms the analogy: "...if the one is not an induction but a description, this must also be true of the other." Let's read more.

> The only real induction concerned in the case, consisted in inferring that because the observed places of Mars were correctly represented

by points in an imaginary ellipse, therefore Mars would continue to re-
volve in that same ellipse; and in concluding (before the gap had been
filled up by further observations) that the positions of the planet dur-
ing the time which intervened between two observations, must have
coincided with the intermediate points of the curve. For these were
facts which had not been directly observed. They were inferences from
the observations; facts inferred, as distinguished from facts seen. But
these inferences were so far from being a part of Kepler's philosophical
operation, that they had been drawn long before he was born. As-
tronomers had long known that the planets periodically returned to
the same places. [Mill, 1872, Book III, ch. ii, section 3]

So, finally, Mill states why Kepler did not perform an induction. The induction
was already performed by astronomers before him who had concluded that the
planets returned to the same places after a fixed period of time. Yes, astronomers
before Kepler did assume that planets repeated exactly the same paths. But
that inductive conclusion is very vague because it does not say what the path
was. Specifying the path adds a great deal of predictive content, and so Kepler's
inference *does* take us from what is known to what is unknown *even if we treat
the periodicity of the orbits as known.*

The only way out for Mill is to insist that the full specification of the path
(the particular ellipse) was a part of the *data*. Mill seems to be assuming that
continuous sections of Mars's orbit were observed at various time, and over time,
these sections covered the whole ellipse. This is factually incorrect, as we shall see.
But even if it were true, it still does *not* follow that Kepler conclusion is a mere
description of the data, unless the observations are *exact*. Any margin of error can
allow for a multitude of possible paths that can disagree in the accelerations that
are attributed to the planets at different times. The consequences that Kepler's
laws have concerning the (unobserved) instantaneous accelerations of the planets
will be crucial in Newton's higher level colligation of Kepler's three laws, according
to which all the planets are attracted to the sun inversely to the square of their
distances to the sun.

It's time to correct this series of mistakes (see also [Harper, 1989; 1993; 2002;
Harper *et al.*, 1994]).[1]

Mill's first mistake was to ignore the difference between angular positions and
3-dimensional positions; this is a huge mistake. The correct story is complicated
because it's not so easy to fill this logical gap. To do it, Kepler first needed to
determine earth's orbit around the sun in relation to a particular point on the
orbit of Mars. The measured period of the Martian orbit was 687 days, which
is a little under two years. Tycho Brahe's observations from earth at E, and at
E_1 687 days later, Kepler obtained the angle SE_1M directly, and obtained ESE_1
from well known tabulations of the (angular) motion of the sun across the fixed
stars. (See Fig. 1.) Mill is right that Kepler simply assumed that the orbits were

---

[1]I follow Hanson's [1970, pp. 277–282] account.

periodic, even though it could never have been justified as exactly true (because it is not).
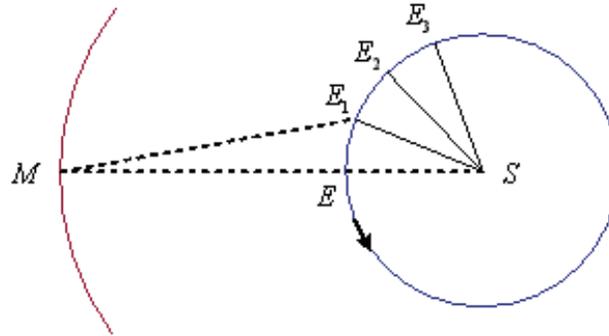

Figure 1. The first step in Kepler's determination of Mars orbit was the calculation of the earth's orbital motion. S denotes the sun, and M refers to Mars.

As a check, Kepler might also have compared the two apparent positions of Mars relative to the fixed stars to obtain the third angle in the triangle, SME_1 (given that Mars returns to the same position M after 687 days). This is an important check given that the periodicity assumption is not entirely secure.

The shape of the triangle SE_1M is thereby given, and this determines the distance SE_1 as a ratio of the (unknown) distance SM. Similar calculations for triangles SE_2M, etc, obtained when Mars had returned to the point M again, then give the distances SE_2, etc, as a ratio of SM also. By then fitting a smooth elliptic orbit to these discrete data points, Kepler determined the motion of the Earth around the sun.

Only now is Kepler able to return to the main problem of measuring the distance of Mars from the sun at different stages of its orbit. Consider another observation of Mars at M' in opposition with the earth at E_0' 687 days later at E_1'. (See Fig. 2.) Again, the shape of the triangle $SE_1'M'$ is determined from the knowledge of its angles, and this gives the distance SM' as a ratio of SE_1'. But the distances SE_1' are known (as a ratio of SM) from the previous colligation of the facts concerning the orbit of earth. Therefore, SM', SM'', etc, are determined as ratios of SM. Kepler then fitted another elliptic curve to obtain the orbit of Mars around the sun as a continuous function of time, which he described in his first law (elliptic path) and second law (equal areas swept out in equal times). Here Kepler is adding a new conception by applying his elliptical formula to the inferred data. Although

these inductions are suggestive, and he may have eliminated many competing hypotheses, Kepler himself did not succeed in *fully* justifying his results. That was left to Newton.
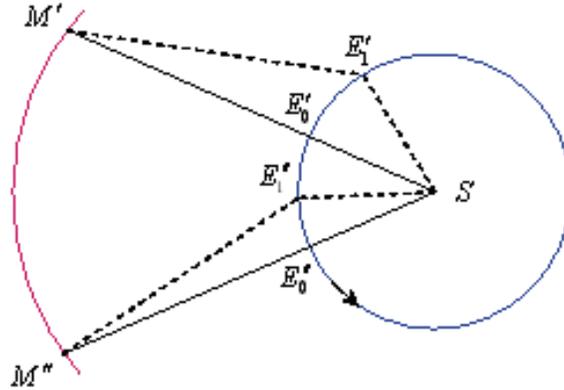

Figure 2. The second step in Kepler's calculation of the Martian orbit.

Against Mill, it is now clear that Kepler's data was only a discrete sampling of points on Mars' orbit. Moreover, each was inferred from measurements of the angles of a triangle and distance ratios that were inferred from another colligation of facts. They were hardly the incorrigible "givens" that empiricists like Mill assume to be the bedrock of inductive inferences. The 3-dimensional positions attributed to Mars were determined in a heavily theory-laden way. However natural it might seem to *assume*, in hindsight, that the planets live in a 3-dimensional space, such attributions are not part of any theory-neutral observation language [Kuhn, 1970]. But, for Whewell, this does not signal the end of the objectivity of science. Higher-level consiliences discovered by Newton will eventually ground the validity of these lower-level conceptions.

The same point applies to Kepler's ellipse. Yes, the ellipse hypothesis might have produced the best fit with the data out of the nineteen hypotheses that Kepler tried, but that does not mean that was completely secure at that time. It was later confirmed by the intimate connection between the inverse square law and Kepler's first and second laws discovered by Newton when he proved that any planet moving such that the line from sun sweeps out equal areas in equal time is accelerating towards the sun, and further, that if the path is an ellipse, the sun-seeking acceleration is inversely proportional to the square of the distance. Furthermore, Kepler's third law is icing on the cake because it also follows from

the inverse square law that the ratios R^3/T^2 are independent measurements of the Sun's mass, adding to the consilience of inductions.

Colligation, for Mill, is a part of the *invention* process, whereas induction (properly so-called) is relevant to questions of *justification*. Whewell's characterization of induction, Mill objects, belongs to (what we call) the 'context of discovery'. Accordingly, Mill [1872, Book III, ch. ii, section 5] charges that "Dr Whewell calls nothing induction where there is not a new mental conception introduced and everything induction where there is." "But," he continues, "this is to confuse two very different things, Invention and Proof." "The introduction of a new conception belongs to Invention: and invention may be required in any operation, but it is the essence of none." Abstracting a general proposition from known facts without concluding anything about unknown instances, Mill goes on to say, is *merely* a "colligation of facts" and bears no resemblance to induction at all.

In sum, Mill thinks that the colligation of facts are mere descriptions that have nothing to do with the justification of scientific hypotheses.

Contrary to what Mill thinks, colligations are not mere descriptions. They do add something unknown to the facts; any general proposition (in Whewell's sense) *can be tested further*, either by untried instances, *or by the consilience of inductions*. It does, therefore, go beyond the data. Yes, mental acts are essential to invention and discovery. But they are *also* essential to justification. Conceptions are essential to the *justification* of the hypothesis that results from a colligation of facts *in spite of* the fact that conceptions are mental, and therefore subjective. Conceptions are essential because there can be no consilience of inductions without them. For, the consilience of inductions often consists of the agreement of magnitudes (Step 3 in the colligation of facts) determined in separate inductions, which derive from the new conception imposed upon the facts in those inductions. Mill has *no good reason* to accuse Whewell of confusing invention and proof. At its core, the dispute is really about the nature of evidence and *justification* — about how hypotheses are tested and confirmed.

3   WHEWELL'S TESTS OF HYPOTHESES

Whewell distinguishes four tests of scientific hypotheses (although the last one is more like a sign than a test). By 'instances' he is referring to empirical data that can be fitted to the hypothesis in question:

1. The Prediction of Tried Instances.

2. The Prediction of Untried Instances;

3. The Consilience of Inductions; and

4. The Convergence of a Theory towards Simplicity and Unity.

Keep in mind that Whewell uses the term 'colligation of facts' interchangeably with 'induction'. A consilience of inductions occurs when two, or more, colligations of

facts are successfully unified in some way. Newton's theory of gravity applied the same form of equation to celestial and terrestrial motions (the inverse square law), and in the case of the moon and the apple, both colligations of facts made use of the same adjustable parameter (the earth's mass). Consequently, the moon's motion and an apple's motion provided independent measurements of the earth's mass, and the agreement of these independent measurements was an important test of Newton's hypothesis. This test is more than a prediction of tried or untried instances. It leads to a prediction of facts *of a different kind* (facts about celestial bodies from facts about terrestrial bodies, and vice versa).

The consilience of inductions leads to a convergence towards simplicity and unity because unified theories forge connections between disparate phenomena, and these connections may be tested empirically, usually by the agreement of independent measurements. So, a theory *can be* unified in response to a successful consilience of inductions. Simplicity and unity are necessary conditions for the consilience of inductions, but not sufficient. A theory like 'everything is the same as everything else' is highly unified, but not consilient. As Einstein once described it, science should be simple, but not too simple.

In the *Novum Organon Renovatum*, Whewell [1989, 151] speaks of the consilience of inductions in the following terms:

> We have here spoken of the prediction of facts *of the same kind* as those from which our rule was collected [tests (1) and (2)]. But the evidence in favour of our induction is of a much higher and more forcible character when it enables us to *explain* and determine cases of a *kind different* from those which were contemplated in the formation of our hypothesis. The instances in which this has occurred, indeed, impress us with a conviction that the truth of our hypothesis is certain. No accident could give rise to such an extraordinary coincidence. No false supposition could, after being adjusted to one class of phenomena, exactly represent a different class, where the agreement was unforeseen and uncontemplated. That rules springing from remote and unconnected quarters should thus leap to the same point, can only arise from that being the point where truth resides.
>
> Accordingly the cases in which inductions from classes of facts altogether different have thus *jumped together*, belong only to the best established theories which the history of science contains. And as I shall have occasion to refer to this peculiar feature of their evidence, I will take the liberty of describing it by a particular phrase; and will term it the *Consilience of Inductions*. [Whewell, 1989, 153]

"Real discoveries are . . . mixed with baseless assumptions" (Whewell, 1989, 145), which is why Whewell considers the consilience of inductions to provide additional guidance in finding the "point where the truth resides."

Whewell has been soundly criticized over the years for his claim that the consilience of inductions "impress us with a conviction that the truth of our hypothesis

is certain" and that "no false supposition could, after being adjusted to one class of phenomena, exactly represent a different class, where the agreement was unforeseen and uncontemplated." Given the explication of the notion of truth that we use today, according to which a hypothesis is false if any small part of it is false, Whewell's claims cannot be defended. But if they are suitably qualified, they cannot be so easily dismissed. It *is true* that such cases "belong only to the best established theories which the history of science contains."

In place of the consilience of inductions, Mill talks about the deductive subsumption of lower level empirical laws under more fundamental laws, which is a well-known part of hypothetico-deductivism. Whewell's account of consilience gets around the common objection that deductive subsumption is too easy to satisfy. For instance, hypothetico-deductivism tries to maintain that Galileo's theory of terrestrial motion, call it G, and Kepler's theory of celestial motion, K, are subsumed under Newton's theory N because N deductively entails G and K. The problem is that G and K are also subsumed under the mere conjunction of $(G\&K)$, so deductive subsumption by itself cannot fully capture the advantage that N is more unified or consilient. Many respond to the problem by saying that unification and simplicity must be added to confirmational equation as non-empirical virtues. But this is to short-change empiricism, because N does make empirical predictions that $(G\&K)$ does not. Namely, N predicts the agreement of independent measurements of the earth's mass from celestial and terrestrial phenomena. That is why Whewell's theory is better than Mill's theory.

Many of these ideas about confirmation have been raised in the literature before [Forster, 1988]. Earman [1978] uses the idea that unified hypotheses have greater empirical content to make sense of Ramsey's argument for realism. Friedman [1981; 1983] uses a similar idea to makes sense of arguments for the reality of spacetime. Glymour [1980] discusses ideas about theory and evidence that have a distinctly Whewellian flavor. Norton [2000a; 2000b] emphasizes the overdetermination of parameters, Harper and Myrvold [2002], Harper [2002; 2007] emphasize the importance of the agreements of independent measurements, and provide excellent detailed examples. These authors appreciate the nuances involved in real examples of scientific discovery, yet there is still a failure to see two things very clearly: (1) The depth of difficulties for *standard* theories of confirmation, such as Bayesianism, and (therefore) a failure to appreciate (2) the relevance of Whewell's ideas to contemporary debates about theory and evidence. To defend the objectivity of knowledge, we need to understand how conceptions introduced in our best explanations are "objectified" by the agreement of independent measurements in a hierarchy of successive generalizations. None of this is going to "fall out" of standard formal theories of epistemology.

4   DISPUTES ABOUT INDUCTION THAT HAVE IGNORED THESE LESSONS

Hempel [1945] made an important distinction between the direct and indirect confirmation of hypotheses. Direct confirmation is the familiar process by which a generalization is confirmed by observed instances of it, while indirect confirmation arises from its place in a larger network of hypotheses. For example, the law of free fall on the moon is directly confirmed by the experiments done on the moon by the Apollo astronauts, but was indirectly confirmed long before that by being deduced from Newton's theory of gravitation, which has its own support. Whewell's discussion of what he termed successive generalizations and the consilience of inductions can be seen as an account of indirect confirmation.

Whewell's idea is this: The aim of any inductive inference is to extract information from the data that can then be used in higher level inductions. For example, Copernicus's theory can be used to infer 3-dimensional positions of the planets relative to the sun from 2-dimensional positions relative to the fixed stars. The 3-dimensional positions were then used by Newton to provide instances of the inverse square law of gravitation, which enable us to make predictions about one planet based on observations of other planets. It was only this higher-level empirical success that finally confirmed Copernicus's conjecture that the earth moved with the sun at the center. Only then can we fully trust the inferences about 3-dimensional positions inferred from Copernicus's theory on which Newton's inductions we based. Whewell explains why this circle is not vicious.

Mill's mistake is to reduce Whewell's innovative idea of the consilience of inductions solely as the deductive subsumption of lower-level generalizations under higher-level laws. The problem with Mill's idea is that it seems to involve a kind of circular reasoning: A is confirmed because A entails B and B is confirmed; but wait, B is now better confirmed because A is confirmed and A entails B. Mill fails to notice that higher-level generalizations have a direct kind of empirical confirmation in terms of the agreement of independent measurements of theoretically postulated quantities. In the case of Newton's theory of planetary motions, it was the agreement of independent measurements of the earth's mass obtained by observing the moon's motion and terrestrial projectiles, and the agreement of independent measurements of the sun's mass, and of Jupiter's mass, and so on. The consilience of inductions thereby relies on *aspects* of the data that play no role in the confirmation of lower-level generalizations. This is why indirect confirmation, on the Whewellian view, avoids the Millian circle.

Whewell's writings were responsible, in part, for the existence of Book III *On Induction* in Mill's *System of Logic*, in which many footnotes and sections are devoted to the important task of separating Mill's views from Whewell's. In 1849, Whewell published a reply called "Of Induction, with Especial reference to Mr. Mill's *System of Logic*". Near the beginning of his commentary, Whewell [1989, p. 267] main complaint is that Mill "has extended the use of the term *Induction* not only to cases in which general induction is consciously applied to particular

instances; but to cases in which the particular instance is dealt with by means of experience in the rude sense in which *experience* is asserted of brutes; and in which, of course, we can in no way imagine that the law is possessed or understood as a general proposition. Mill has thus "overlooked the broad and essential difference between speculative knowledge and practical action; and has introduced cases which are quire foreign to the idea of science, alongside with cases from which we may hope to obtain some views of the nature of science and the processes by which it must be formed." In a footnote to chapter i, Book III, Mill [1872] replies: "I disclaim, as strongly as Dr. Whewell can do, the application of such terms as induction, inference, or reasoning, to operations performed by mere instinct, that is from an animal impulse, without the exertion of any intelligence." But the essence of Whewell's complaint is that simple enumerative induction, and Mill's other methods of induction, are no more complicated than animal impulses even when it is consciously employed; at least, not different in a way that accounts for the difference in intelligence.

If the complaint is about the established use of the word "induction", then I tend to think that Whewell is the one swimming against the tide. But it would be a mistake to think that this is merely a linguistic debate about the use of the word 'induction'; for as Whewell notes, there is always a proposition that accompanies every definition, and the proposition in this case is something like: Simple enumerative induction (such as inferring that all humans are mortal from John, Paul,... are mortal) adequately represents the habit of mind that brings about the highest forms of human knowledge. This is an assumption that should be questioned in light of what we know today.

Whewell expands upon his worries by characterizing most generalizations of the form "All humans are mortal" as a mere juxtapositions of particular cases [Whewell, 1989, 163]. Whewell agrees that induction is the operation of discovering and proving general propositions, but he appears to have a different understanding of the term "general". For Whewell (1989, 47) it is necessary that "In each inductive process, there is some general idea introduced, which is given, not by the phenomena, but by the mind." The inductive conclusion is, therefore, composed of facts *and* conceptions "bound together so as to give rise to those general propositions of which science consists". "All humans are mortal" is not general in the appropriate sense because there has been no conception *added* to the fact that John, Paul,... are mortal.[2] Whewell insists that in every genuine induction, "The facts are known but they are insulated and unconnected . . . The pearls are there but they will not hang together until some one provides the string" [Whewell, 1989, 140-141]. The "pearls" are the data points and the "string" is a new conception that *connects* and *unifies* the data. The "pearls" in "All *A*s are *B*s" are unstrung because "All *A*s are *B*s", though general in the sense that it is

---

[2]But it would be incorrect to say that Whewell thinks that no generalization of the for All *A*s are *B*s can introduce a new conception. For example, it could be that "All metals conduct electricity" qualifies as an induction conclusion because the term 'metal' may represent a new conception not contained in the facts. I owe this point to Dan Schneider.

universally quantified, does not connect or unify the facts; it does not *colligate* the facts. For Whewell, this process of uniting the facts under a general conception, which he calls the *colligation of facts*, is an essential step in the formation of human knowledge. Mill would gladly transfer Whewell's description of the colligation of facts to his own pages, but fails to see that it has the kind of importance that Whewell attaches to it.

There are two worries that everyone should have about simple enumerative induction:

(1) It is *not* a habit of mind that we have in a great many cases; in fact, it is the subject of well known philosophical jokes. A philosopher jumps from the Empire State Building and is heard to say as he falls past the 99th floor "99 floors and I'm not dead!" As a different example, imagine a study of radioactive decay in which all the samples observed are radioactive, yet the very law of radioactive decay discovered from these observations leads us to deny that any finite sample will be radioactive *for all times*.

(2) When such a habit of mind is desirable, it is very easy to implement. Simple associative learning is not what marks the difference between human intelligence and animal intelligence. I say 'salt' and you think 'pepper'. Pavlov's dogs are the most famous case of a kind of associative learning in animals known as *classical conditioning.* In more recent times, the same learning ability has been demonstrated in animals as primitive as sea slugs (*Aplysia californica*). It's not just that "brutes" do it, sea slugs do it! A strong 1-sec electric shock to the mantle of the slug (called the *unconditioned stimulus* UCS) elicits a *prolonged* withdrawal of its siphon. The UCS in Pavlov's dogs is the smell of meat, which elicits salivation. The aim of the experiments is to demonstrate an ability to learn to predict the UCS from a *conditioned stimulus* (CS). In Pavlov's dogs, the CS was the sound of a bell. When presented immediately prior the presentation of food on several occasions, the bell would eventually trigger the salivation response *by itself without the smell of meat*, thereby indicating that the dogs had learned to predict the presence of meat from the sound of the bell. In the case of the sea slugs, one CS was a short tactile stimulation of the siphon, which elicited a *short* withdrawal of the siphon. When the CS was presented a short 0.5 sec before the UCS, and this was repeated 15 times, the CS would produce a siphon withdrawal that is more 4 times as long as what would have resulted without the learned association between the CS and the UCS. Just as Pavlov's dogs appear to learn to "predict" the presence of food from the sound of a bell, the sea slugs appear to anticipate a large electrical shock from a short tactile stimulation of the siphon.[3] Sea slugs have about 20,000 nerve cells in its central nerve system arranged in nine ganglia [Macphail, 1993, p. 32] compared to the approximately 10^{12} neurons in a human being, some of which may have several thousand synaptic contacts [Nauta and Feirtag, 1986]. What is the function of these extra neurons? To learn a billion more associations of the same

---

[3]No such association is learned when the CS is presented after the UCS. See [Macphail, 1993, pp. 103-5], for a more complete description of the experiment, or the original source; Carew, Hawkins, and Kandel 1983.

kind? If so, how are these learned associations organized or associated together?

The most influential part of the *System of Logic* is Mill's four methods of induction [Mill, 1972, Book III, Chapter VIII, IX]; but these are also the butt of many jokes. A philosopher goes to a bar on Monday and drinks whiskey and soda water all night. The next day he drinks vodka and soda. The following night, gin and soda, and then the night after that, bourbon and soda. Finally, on Friday, he comes into the bar and complains that he's been too inebriated for the past week to get much work done, so tonight he's going to drink whiskey *without the soda*. The philosopher has used Mill's the method of agreement to observe that the only common thread in the four times he's been inebriated is that he's been drinking soda water. Therefore, soda water causes inebriation. So much the worse for simple inductive rules mindlessly applied.

Of Mill's four methods, Whewell [1989, p. 286] writes: "Upon these methods, the obvious thing to remark is, that they take for granted the very thing which is the most difficult to discover, the reduction of the phenomena to formulae such as are here presented to us. When we have any set of complex facts offered to us; for instance. . . the facts of the planetary paths, of falling bodies, of refracted rays, of cosmical motions, of chemical analysis; and when, in any of these cases, we would discover the law of nature which governs them, or if any one chooses so to term it, the feature in which all the cases agree, where are we to look for our A, B, C, and a, b, c? Nature does not present to us the cases in this form. . ."

Whewell's point is very simple. In order to discover a connections between two disparate phenomena, we need to be able to extract the *relevant* information from each domain, that is, introduce quantities that will prove to be connected, yet we don't know that until after we collect the right kind of data and see whether the quantities fit together in higher-level regularities. This kind of catch-22 makes discovery extremely difficult, though not impossible for human beings. But for present-day machines, computer systems, and primitive organisms, it has not been possible.

A failure to see the depth of the problem is the root cause of the overly optimistic forecasts in the 1960s about how the AI systems would match human intelligence within 20 years. Even the apparent exceptions to this, such as the Deep Blue chess-playing program, prove the rule. In 1996, Deep Blue became the first computer system to defeat a reigning world champion (Garry Kasparov) in a match under standard chess tournament time controls. But it did it by brute force computing power, rather than the pattern-recognition techniques of the human chess masters, which enable them to play 40 opponents at once. (See [Dreyfus, 1992] for an in-depth analysis.)

In 1987, researchers based at Carnegie Mellon University (CMU) published a book called *Scientific Discovery: Computational Explorations of the Creative Process* by Langley, Simon, Bradshaw, and Zytkow. Again, the basic Whewellian criticism was raised about the computer programs such as Bacon, an AI system that rediscovered numeric laws such as Kepler's third law, which equates the period of revolution of a planet around Sun to the 3/2 power of the mean radius. It's one

thing to ask how to relate one variable to another *when the variables are already given*, but quite another to discover Kepler's laws from raw data about the angular positions of the planets at various times. Even knowing that 'position relative to the fixed stars' and 'time' can be functionally related is a major step forward. Nothing like this has been replicated by any computer system. That's not to say that it's impossible (indeed Langley and Bridewell (in press) speak in terms that remind me of Whewell). After all, our brains are computers and a network of these computers did solve the problem. But we must recognize that the requisite "explication of the conceptions", to use Whewell's term, is difficult.

The most recent instance of this kind of disagreement surrounds the work by another group at CMU headed by Spirtes, Glymour and Scheines [1993], who have developed algorithms for discovering causal models or Bayes nets. Humphreys and Freedman [1996] published a critique, while Spirtes, Glymour and Scheines [1997] and Korb and Wallace [1997] published a reply. Again, this research in computer-automated algorithms of scientific discovery is an extremely valuable. The question is whether it could be improved by an implementation of Whewellian ideas (see [Forster, 2006]).

In 1981, Hinton and Anderson edited an important volume on *Parallel Models of Associative Memory*, which was followed up by the very famous work on parallel distributed processing edited by Rumelhart and McClelland in 1986, which gave birth to a thriving industry on connectionist networks, otherwise known as artificial neural networks. The breakthrough was made possible by the mathematical discovery about how to implement a learning algorithm in neural networks that propagates backwards in the network to adjust connection weights so as to reduce the error in the output [Rumelhart *et al.*, 1986]. Yet again, the lesson turned out to be the same: An all-purpose neural network is able to approximate any function *in principle*; but in practice too much flexibility creates difficulties. Top-down constraints need to be imposed on the network before data-driven search methods can match any of the cognitive abilities of human beings. My only point is that, in each of these episodes, it has taken quite some time to rediscover some of the points that were raised 150 years ago in the Whewell-Mill debate.

5   IMPLICATIONS FOR PROBABILISTIC THEORIES OF EVIDENCE AND CONFIRMATION

Allow me to predict a new example of the same thing. At the present time, there seems to me to be an overestimation of what the methods of statistical inference can achieve. In philosophy of science, major figures in the field endorse the view that Bayesian or Likelihoodist approaches to statistical reasoning can be extended to cover scientific reasoning more generally. In [Forster, 2007], I have argued that standard statistical methods of model selection, such as AIC [Akaike, 1973] and BIC [Schwartz, 1978], are fundamentally limited in their ability to replicate the methods of scientific discovery. (Note that connectionist networks are also implementing a standard statistical learning rule known as the method

of least squares.) In [Forster, 2006], I put forward a positive suggestion about how Whewellian ideas about the consilience of inductions enrich the relationship between theory and evidence, which could improve the rate of learning and the amount that can be learned.

Continuing on the same theme, philosophers of science, such as Hesse [1968; 1971], Achinstein [1990; 1992; 1994], and more recently Myrvold [2003], have tried to capture the confirmational value of consilience and unification in terms of standard probabilistic theories of confirmation, but with limited success. The reason for their limited success is illustrated by the following schematic example. Suppose we have a set of three objects $\{a, b, c\}$ that can be hung on a mass measuring device, either individually or in pairs, $a*b$, $a*c$, and $b*c$, where $a*b$ denotes the object consisting of a conjoined with b, and so on. Suppose that the *Data* consists of six measurements of the distances at which the counterweight need to be hung from the center of a beam balance in order to balance the object being measured. Let's denote this observed distance as $x(o)$, where o is the name of the object being measured. In order to talk about the consilience of inductions, we need two, or more, separate inductions; so let's divide the data into two parts, and consider inductions performed on each part.

$$\text{Data1} = \{x(a) = 1,\ x(b) = 2,\ x(c) = 3\},$$

and

$$\text{Data2} = \{x(a*b) = 3,\ x(a*c) = 4,\ x(b*c) = 5\}.$$

The core hypothesis under consideration is the assertion that for all objects o, $x(o) = m(o)$, where $m(o)$ denotes a theoretically postulated property of object o called *mass*.

$$M: \qquad\qquad\qquad (\forall o)(x(o) = m(o)).$$

The quantity x can be repeatedly measured, but no assumption is made that its value will be the same on different occasions. That depends on what the world is like. On the other hand, the hypothesis M *asserts* that masses are constant over time. The postulated constancy of m, combined with the equation, *predicts* that repeated measurements on the same object will be the same. It's easy to equate some new quantity m with the outcome of measurement x, but it's not so easy to defend the new quantity as representing something real underlying the observable phenomena.

If we apply the conception that $x(o) = m(o)$ to the two data sets, we notice that the hypothesis *accommodates* the data in each case, and there is no test of the hypothesis in the precise sense that the hypothesis would not have been refuted had the data been "generated by" a contrary hypothesis [Mayo, 1996]. The predictive content is not tested by single measurements of each mass. Yet, we

can arrive at an inductive conclusion from the data according to standard rules. In the case of Data1, we arrive at the hypothesis

$$h_1: \qquad M\&\{m(a) = 1, m(b) = 2, m(c) = 3\}.$$

Note that $h_1 \Rightarrow$ Data1, where '$\Rightarrow$' means 'logically entails'. I have no problem with the claim that the data Data1 confirms the hypothesis h_1, although it does so by pointing to the particular predictive hypothesis out of all those compatible with M, rather than confirming M itself.

Now let's consider the inductive conclusion arrived at on the basis of Data2:

$$h_2 \qquad M\&\{m(a * b) = 3, m(a * c) = 4, m(a * c) = 5\}.$$

Again, $h_2 \Rightarrow$ Data2, and the data confirms the inductive hypothesis. On my understanding of Whewell and Mill, they would agree on this.

To explain the difference between Whewell and Mill, let's consider a stronger inductive conclusion that includes the standard Newtonian conception that *the mass of a composite object such as a*b is the sum of the masses of the parts*. We shall call this the *law of the composition of masses* (LCM), and write it more formally as:

$$\text{LCM} \qquad (\forall o_1)(\forall o_2)(m(o_1 * o_2) = m(o_1) + m(o_2)).$$

Let's denote the stronger inductive conclusions drawn from the data sets by $H_1 = h_1$ & LCM and $H_2 = h_2$ & LCM, respectively. Again, the data confirms the respective hypotheses, but only by picking out the mass values that correctly apply to the objects. There is no confirmation of the general propositions in the inductive hypotheses by Data1 or Data2.

But all this changes when we consider the bigger picture; for H_1 and H_2 entail more than the data from which they were inductively inferred, they predict the other data set as well. That is, $H_1 \Rightarrow$ Data2, and $H_2 \Rightarrow$ Data1. This is an illustration of the idea behind Whewell's *consilience of inductions*... "That rules springing from remote and unconnected quarters should thus leap to the same point, can only arise from that being the point where truth resides" [Whewell, 1989, p. 153]. The hypotheses h_1 and h_2 enjoy no such relationship with the data.

Another way of seeing the same thing is to note that the two data sets, Data1 and Data2, provide independent measurements of the theoretically postulated masses, $m(a)$, $m(b)$, and $m(c)$, and the independent measurements agree.[4] From Data1, we obtain values of $m(a)$, $m(b)$, and $m(c)$, and from Data2, we obtain values of $m(a) + m(b) = 3$, $m(a) + m(c) = 4$, and $m(b) + m(c) = 5$. Since there are three equations in three unknowns, these equations yield an independent set

---

[4] "Independent" just means that the measurements are calculated from non-overlapping sets of data.

of values for the three masses, which agree with the first set. Therefore H is confirmed by agreement of independent measurements of its postulated quantities, while $h = h_1 \& h_2$ is not.

The intuition just described is far more forceful if we were to embellish the example by including a set of mass measurements on a larger set of objects; say 25 objects. Then Data1 consists of 25 measurements of the 25 objects, whereas Data2 consists of 300 measurements of all possible pairings of the 25 objects, which provides 12 more independent measurements of each mass. That fact that 13 independent measurements of mass agree for each of 25 different objects is *very strong* evidence for the hypothesis H.

Unfortunately, we cannot obtain this conclusion (that H is better supported by the *Data* than h) from the standard theories of confirmation used in contemporary philosophy of science or in statistics, such as Bayesianism and Likelihoodism.[5] These views are committed to a likelihood theory of evidence that says that degree to which a total evidence, the *Data* in our example, supports a hypothesis, such as H or h, is *fully exhausted by* likelihoods $P(Data|H)$ and $P(Data|h)$. But, $H \Rightarrow Data$, and $h \Rightarrow Data$, and, therefore, $P(Data|H) = 1 = P(Data|h)$. The relationship between theory and evidence is therefore the same for each of the hypotheses according to these (well respected) accounts of the nature of evidence.

I suspect that the Bayesians and Likelihoodists will respond to this example along the following lines. Instead of considering the hypotheses as I have defined them, which include the "determination of the magnitudes" (as Whewell would put it), we should consider just the generalizations M and (M & LCM). Then we can argue that (M & LCM) gives the *Data* a greater probability (i.e., the hypothesis has a greater likelihood). They may argue that $P(Data|\,M\&\mathrm{LCM}) > P(Data|\,M)$.[6] The idea behind this claim is very simple, but first you need to understand that (by the axioms of probability) the likelihood of a family of hypotheses is equal to a weighted average of the likelihoods of the hypotheses in the family. (M & LCM) is a family of hypotheses in which one member, namely H, has likelihood 1, while all the others have likelihood 0 because they get at least one mass value wrong (out of the masses that have been measured). The same applies to M; it contains one hypothesis with likelihood 1 and the rest with likelihood 0. (Having likelihood 0 usually means that the hypothesis is refuted by the data.) Thus, (M & LCM) has a greater likelihood because its likelihood is calculated by

---

[5]The one exception that I know of is Mayo [1996]. Her take on this example would be that H is *severely* tested by the *Data* because the probability is high that H would be refuted if H were false. But h is not severely tested by the *Data* because it would not be refuted if h were false. The uneasiness I have with this approach is the reference to counterfactual data. Other things being equal, I prefer a theory of confirmation that focuses only on the actual data.

[6]**Proof**: $P(Data|\,M\&\mathrm{LCM}) = P(\mathrm{Data1}|M\&\mathrm{LCM})P(\mathrm{Data2}|M\&\mathrm{LCM}\ \&\ \mathrm{Data1})$.

But $P(\mathrm{Data2}|M\&\mathrm{LCM}\ \&\ \mathrm{Data1}) = 1$, so $P(Data|\,M\&\mathrm{LCM}) = P(\mathrm{Data1}|M\&\mathrm{LCM})$.

But now it is clear that the hypotheses "say the same thing" about Data1, so $P(\mathrm{Data1}|M\&\mathrm{LCM}) = P(\mathrm{Data1}|M)$, and it is obvious that $P(\mathrm{Data1}|M) > P(Data|\,M)$. Thus, the result follows.

averaging over a larger set of other hypotheses, all of which have zero likelihood. In other words, the likelihood of M is smaller because its maximum likelihood is washed out by averaging over a greater number of hypotheses.

The first problem with this reply is that it *changes the subject*. We began by talking about the confirmation of H and h, and ended about talking about something else. But let's consider the confirmation of (M & LCM) and M. The problem is that under any sensible way of averaging likelihoods, it turns out to be zero, zilch, nil. This is because there is only one point hypothesis that has non-zero likelihood, so any weighting that averages (integrates) over an infinite number (a continuum) of point hypotheses will yield an average likelihood of zero (Forster and Sober 1994). So, the claim that $P(Data|M\&\text{LCM}) > P(Data|M)$ is incorrect. It should have been $P(Data|M\&\text{LCM}) \geq P(Data|M)$. And under the rather general conditions I have stated, $P(Data|M\&\text{LCM}) = P(Data|M)$.

The core part of the Bayesian argument, the part that was right, derives from the inequality

$$P(\text{Data2}|M\&\text{LCM \& Data1}) = 1 > P(\text{Data2}|M \text{ \& Data1}) = 0.$$

But this inequality is just what lies at the heart of Whewell's consilience of inductions! Once we see that the inequality is what's crucial, then we can express what should be said about the *original* example in the language of probability, *without changing the subject*. For note that the hypothesis (M & LCM) & Data1 is logically equivalent to H_1, as we previously defined it, and M & Data1 is logically equivalent to h_1. So the inequality is just

$$P(\text{Data2}|H_1) = 1 > P(\text{Data2}|h_1) = 0,$$

to which we could add

$$P(\text{Data1}|H_2) = 1 > P(\text{Data1}|h_2) = 0.$$

In other words, the part of the likelihood analysis that makes sense *rests on Whewellian principles*. Why try to wrap it up in a Bayesian package with trappings that are false at worst, and irrelevant at best? I suggest that it is philosophically more fruitful to understand the relationship between theory and evidence in Whewellian terms right from the beginning.

To repeat, as Whewell points out, nature does not present inductive problems in a form that lends itself to any simple methods of induction. In the mass measurement example, we began with two sets of data, with two phenomena, each of which is colligated by the formula $x(o) = m(o)$, but we can discover no deeper connection between them until we explicate the concept of mass by introducing the law of composition of masses (LCM). **Question**: How do we explain why these thirteen independent measurements agree? **Answer**: By concluding that they are measurements *of the same quantity*, the effects of a common cause. Arguing that we should explain many effects in terms of a common cause is the easy part of the discovery. The harder part is to arrive at the problem in this form.

The same is true of the Kepler example.

ACKNOWLEDGEMENTS

BIBLIOGRAPHY

[Achinstein, 1990] P. Achinstein. Hypotheses, Probability, and Waves. *British Journal for the Philosophy of Science* **41**: 73-102, 1990.

[Achinstein, 1992] P. Achinstein. Inference to the Best Explanation: Or, Who Won the Mill-Whewell Debate? *Studies in the History and Philosophy of Science* **23**: 349-364, 1992.

[Achinstein, 1994] P. Achinstein. Explanation v. Prediction: Which Carries More Weight? In David Hull and Richard M. Burian (eds.), *PSA* 1994, vol. 2, East Lansing, MI, Philosophy of Science Association, 156-164, 1994.

[Akaike, 1973] H. Akaike. Information Theory and an Extension of the Maximum Likelihood Principle. B. N. Petrov and F. Csaki (eds.), *2nd International Symposium on Information Theory:* 26781. Budapest: Akademiai Kiado, 1973.

[Carew *et al.*, 1983] T. J. Carew, R. D. Hawkins, and E. R. Kandel. Differential classical conditioning of a defensive withdrawal reflex in *Aplysia* californica. *Science* **219**: 397-400, 1983.

[Dreyfus, 1992] H. L. Dreyfus. *What Computers Still Can't Do: A Critique of Artificial Reason* MIT Press: Cambridge, Mass, 1992.

[Earman, 1978] J. Earman. Fairy Tales vs. an Ongoing Story: Ramsey's Neglected Argument for Scientific Realism. *Philosophical Studies 33:* 195-202, 1978.

[Forster, 1988] M. R. Forster. 'Unification, Explanation, and the Composition of Causes in Newtonian Mechanics. *Studies in the History and Philosophy of Science* **19**: 55–101, 1988.

[Forster, 2006] M. R. Forster. Counterexamples to a Likelihood Theory of Evidence, *Mind and Machines*, **16**: 319-338, 2006.

[Forster, 2007] M. R. Forster. A Philosopher's Guide to Empirical Success, *Philosophy of Science*, Vol. 74, No. 5: 588-600, 2007.

[Forster and Sober, 1994] M. R. Forster and E. Sober. How to Tell when Simpler, More Unified, or Less *Ad Hoc* Theories will Provide More Accurate Predictions. *British Journal for the Philosophy of Science* **45:** 1–35, 1994.

[Friedman, 1981] M. Friedman. Theoretical Explanation. In *Time, Reduction and Reality*. Edited by R. A. Healey. Cambridge: Cambridge University Press. Pages 1–16, 1981.

[Friedman, 1983] M. Friedman. *Foundations of SpaceTime Theories*. Princeton, NJ: Princeton University Press, 1983.

[Glymour, 1980] C. Glymour. Explanations, Tests, Unity and Necessity. *Noûs* **14:** 31–50, 1980.

[Hanson, 1973] N. R. Hanson. *Constellations and Conjectures,* W. C. Humphreys, Jr. (ed.) D. Reidel: DordrechtHolland, 1973.

[Harper, 1989] W. L. Harper. Consilience and Natural Kind Reasoning. In J. R. Brown and J. Mittelstrass (eds.) *An Intimate Relation:* 115152. Dordrecht: Kluwer Academic Publishers, 1989.

[Harper, 1993] W. L. Harper. Reasoning from Phenomena: Newton's Argument for Universal Gravitation and the Practice of Science. In Paul Theerman and Seeff, Adele F. (eds.) *Action and Reaction*, Newmark: University of Delaware Press, 144–182, 1993.

[Harper, 2002] W. L. Harper. Howard Stein on Isaac Newton: Beyond Hypotheses. In David B. Malament (ed.) *Reading Natural Philosophy*: *Essays in the History and Philosophy of Science and Mathematics*. Chicago and La Salle, Illinois: Open Court. 71–112, 2002.

[Harper, 2007] W. L. Harper. 'Newton's Method and Mercury's Perihelion before and after Einstein. *Philosophy of Science* **74**: 932-942, 2007.

[Harper *et al.*, 1994] W. L. Harper, B. H. Bennett and S. Valluri. "Unification and Support: Harmonic Law Ratios Measure the Mass of the Sun." In D. Prawitz and D. Westerståhl (eds.) *Logic and Philosophy of Science in Uppsala:* 131-146. Dordrecht: Kluwer Academic Publishers, 1994.

[Hinton and Anderson, 1981] G. E. Hinton and J. A. Anderson, eds. *Parallel Models of Associative Memory*. Hillsdale, NJ: Lawrence Erlbaum Associates, 1981.

[Hempel, 1945] C. G. Hempel. Studies in the Logic of Confirmation. *Mind, vol.* **54**, 1945. Reprinted in Hempel [1965].

[Hempel, 1965] C. G. Hempel. *Aspects of Scientific Explanation and Other Essays in the Philosophy of Science.* New York: The Free Press, 1965.

[Hesse, 1968] M. Hesse. Consilience of Inductions. In I. Lakatos (ed.), *Inductive Logic.* North-Holland, Amsterdam, 1968.

[Hesse, 1971] M. Hesse. Whewell's consilience of inductions and predictions, *The Monist* , 55, 520-524, 1971.

[Humphreys and Freedman, 1996] P. Humphreys and D. Freedman. The Grand Leap. *British Journal for the Philosophy of Science* **47**: 113-123, 1996.

[Korb and Wallace, 1997] K. B. Korb and C. S. Wallace. In Search of the Philosopher's Stone: Remarks on Humphreys and Freedman's Critique of Causal Discovery. *British Journal for the Philosophy of Science* **48**: 543-553, 1997.

[Kuhn, 1970] T. Kuhn. *The Structure of Scientific Revolutions,* Second Edition. Chicago: University of Chicago Press, 1970.

[Kuhn, 1970a] T. Kuhn. *The Structure of Scientific Revolutions,* Second Edition. Chicago: University of Chicago Press, 1970.

[Langley *et al.*, 1987] P. H. Langley, H. A. Simon, G. L. Bradshaw, and J. M. Zytkow. *Scientific Discovery: Computational Explorations of the Creative Process.* MIT Press, Cambridge, Mass, 1987.

[Langley and Bridewell, in press] P. H. Langley and W. Bridewell. Processes and constraints in explanatory scientific discovery. *Proceedings of the Thirtieth Annual Meeting of the Cognitive Science Society.* Washington, D.C., in press.

[Macphail, 1993] E. M. Macphail. *The Neuroscience of Animal Intelligence: From the Seahare to the Seahorse.* Columbia University Press, New York, 1993.

[Mayo, 1996] D. G. Mayo. *Error and the Growth of Experimental Knowledge.* Chicago and London, The University of Chicago Press, 1996.

[Mill, 1872] J. S. Mill. *A System of Logic, Ratiocinative and Inductive: Being a Connected View of the Principles of Evidence and the Methods of Scientific Investigation*, 1872. Eighth Edition (Toronto: University of Toronto Press, 1974).

[Myrvold, 2003] W. Myrvold. A Bayesian Account of the Virtue of Unification, *Philosophy of Science* **70**: 399-423, 2003.

[Myrvold and Harper, 2002] W. Myrvold and W. L. Harper. Model Selection, Simplicity, and Scientific Inference, *Philosophy of Science* **69**: S135-S149, 2002.

[Nauta and Feirtag, 1986] W. J. H. Nauta and M. Feirtag. *Fundamental Neuroanatomy.* New York: W. H. Freeman, 1986.

[Norton, 2000a] J. D. Norton. The Determination of Theory by Evidence: The Case for Quantum Discontinuity, 1900-1915, *Synthese* **97**: 1-31, 2000.

[Norton, 2000b] J. D. Norton. How We Know about Electrons. In Robert Nola and Howard Sankey (eds.) *After Popper, Kuhn and Feyerabend*, Kluwer Academic Press, 67-97, 2000.

[Rumelhart *et al.*, 1986] D. E. Rumelhart, J. McClelland, *et al. Parallel Distributed Processing, Volumes 1 and 2.* MIT Press, Cambridge, Mass, 1986.

[Rummelhart *et al.*, 1986a] D. E. Rumelhart, G. Hinton, and R. J. Williams. *Nature* **323:** 533536, 1986.

[Schwarz, 1978] G. Schwarz. Estimating the Dimension of a Model. *Annals of Statistics* **6:** 4615, 1978.

[Spirtes *et al.*, 1993] P. Spirtes, C. Glymour and R. Scheines. *Causation, Prediction and* Search. New York: Springer-Verlag, 1993.

[Whewell, 1840] W. Whewell. *The Philosophy of the Inductive Sciences* (1967 edition). London: Frank Cass & Co. Ltd. 1840.

[Whewell, 1847] W. Whewell. *Philosophy of the Inductive Sciences* , 2 vols. (London, John W. Parker), 1847.

[Whewell, 1858] W. Whewell. *Novum Organon Renovatum*, Part II of the 3rd the third edition of *The Philosophy of the Inductive Sciences*, London, Cass, (1858), 1967.

[Whewell, 1989] W. Whewell. *William Whewell: Theory of Scientific Method.* Edited by Robert Butts. Hackett Publishing Company, Indianapolis/Cambridge, 1989.