Whewell's Theory of Hypothesis Testing and a Relational View of Evidence[*]

Malcolm R. Forster and Ann B. Wolfe
Department of Philosophy
University of Wisconsin, Madison
5185 Helen C. White Hall
600 North Park Street
Madison, WI, 53706 U. S. A.
mforster@facstaff.wisc.edu

November 20, 2000

ABSTRACT: The debate between William Whewell and John Stuart Mill is not only hard in the sense that both sides are difficult to understand, but the issue itself is unresolved. Whewell's idea of predictive tests is similar to the method of cross validation in statistics and machine learning, except that Whewell applies it in a hierarchical way at multiple levels. Or at least, that is how Whewell argues that hypothesis testing works in science. In contrast, the received view of theory testing is that the confirmation of rival hypotheses is measured by their degree of fit with the *total* evidence, provided that the rival hypotheses are equally simple. However, there is a growing realization that predictive tests are stronger in many ways. What this suggests is that the history of science could be used as a source of examples against which theories of learning may be tested. The purpose of this paper is to explain and highlight some of the features of Whewell's theory of hypothesis testing that are relevant to the continuing controversies about confirmation.

1.   The Central Issues

After John Stuart Mill read William Whewell's *The Philosophy of the Inductive Sciences*, he decided to expand his own treatise on deductive logic to include a book on inductive logic. Mill extracted everything that made sense to him in Whewell, who was most famously a historian of science at that time. The end product, published as Book III of Mill's *System of Logic*, was the logical wheat of Whewell's philosophy of science winnowed away from the non-logical chaff. Or, at least, that is the perception of those who have failed to understand the parts of Whewell that Mill left out. Eventually Whewell responded to Mill in writing ('*Mr. Mill's Logic*', 1849, reprinted in Butts, ed., 1989), and Mill documented his continuing disagreement with Whewell in long footnotes added to subsequent editions of *System of Logic*. Mill's transmutation of Whewell's philosophy of science was more influential in the long term, and is closer to contemporary philosophy of science. The purpose of this essay is to make sense of Whewell's ideas in a way that makes them relevant to contemporary debates about the testing and confirmation of scientific theories.

Logicians view language as a fixed background framework in which rival theories are tested by their deductive consequences. While the language of science can be enriched and expanded according to Mill, the way in which this happens is highly restricted. Mill defends an empiricist viewpoint by which conceptual innovations are read from the newly discovered observational facts. Whewell took a completely opposing view—that rival scientific theories introduce their own conceptual innovations in a conjectural way, not determined by the observational facts or phenomena.[1] The conceptions introduced by rival theories are therefore part of what needs to tested and evaluated in science. To this end, Whewell introduces several "tests of hypotheses", including his famous *consilience of inductions*, and the tendency towards simplicity and unity.

We intend to argue that Mill and more recent philosophers of science have systematically misunderstood Whewell's philosophy of science in ways that have obscured its relevance to contemporary issues in the philosophy of science. Perhaps the single most common misconception is that Whewell's philosophy is "merely" an account of the psychology of scientific discovery, with no interesting or novel contributions to the normative question of how to tell the difference between good and bad science. On the contrary, we argue that Whewell's writings are a treasure-trove of *valuable* insights about theory evaluation.

The difficulty in extracting these insights from Whewell's writings arises from the following dilemma. On the one hand, the most clearly normative part of Whewell's philosophy of science looks like an early form of hypothetico-deductivism, according to which scientific theories are evaluated by comparing their deductive consequences with the observational facts (Butts, 1989; Jeong, 1991). In its basic form, hypothetico-deductivism says that evidence E confirms hypothesis H if and only if E is true, and H entails E. If this were correct, then the essence of Whewell's normative philosophy of science would be more clearly stated by Mill (1849). On the other hand, it is easy to find passages in Whewell's writings that have nothing to do with hypothetico-deductivism. But in these passages it appears that Whewell is really concerned with the *psychology* of discovery rather than the evaluation of theories. So, either way, if one is concerned with the evaluation of theories, then there appears to be nothing of interest in Whewell's writings beyond what is adequately described by Mill.

In our view, the dilemma is a false dilemma. It is true that Whewell talks about the processes of scientific discovery, and that this topic has nothing directly to do with the evaluation of theories. However, it must be understood that Whewell (rightly or wrongly) thinks that there is a constant interplay between subjective (psychological) and objective (logical) aspects of science. So, he often discussed both of these aspects in a single breathe. This "fundamental antithesis" (section 7) makes Whewell difficult to read and hard to understand. But at the same time, it makes his writings intriguing and insightful.

Historically, the debate between Whewell and Mill begins with a dispute about the word "induction." Both Mill and Whewell agree to use this word to denote whatever process in science leads from particular facts to general propositions; or from observed phenomena to theories. Yet for Whewell, the introduction of new conceptions is an *essential* part of scientific induction, while Mill refuses to see it that way (section 2). In sections 3 to 5, we argue that Whewell's description of the process better fits the historical examples. To the extent that these examples are examples of good science, this supports the view that Whewell's account has greater normative and descriptive value.

---

[1] Facts do not necessarily refer to the raw data. In this way, Whewell has already adopted the viewpoint of Bogen and Woodward (1988), whereby theory evaluation proceeds in two steps: (1) phenomena are inferred from the raw data, and (2) theories are compared with the phenomena.

A major obstacle to convincing philosophers of science of the value of Whewell's writings is the existence of some misleading interpretations of his view. For example, Whewell argues for the historical thesis that well tested theories gain a familiarity in the minds of scientists from which they acquire the status of necessary truths. For Whewell, necessity is a subjective property of theories, according to which the theory becomes so familiar and natural to scientists who are brought up with the view that they are unable to conceive of its contrary. This psychological claim is comparable to Kuhn's (1970) view that scientists working within a paradigm never question its truth. Yet, Butts (1989) suggests that Whewell's normative epistemology may amount to nothing more than the slogan "We must believe p, therefore p is true." In other places, Whewell describes the objective tests that a theory will pass before reaching this state in the minds of scientists. These tests may be forgotten by the scientist who are brought up in the theory, but it does not follow that the psychological entrenchment of a theory is its justification.

Another important misunderstanding has led to the widely held view that the *value* of a Whewellian test derives from the *psychological surprise* of predictions, which depends on the historical order of the deduction and the observation. Historical order *is* useful evidence, but only when one does not know the full logical facts. The fact that predictions are made before they are known to be true is evidence that they actually follow from the theory. The relevance of historical order in this weak sense does not imply that Whewell treats confirmation as a subjective commodity (section 9). For when the truth of the prediction is known before the prediction is derived (as was the case of Einstein's prediction of the precession of the perihelion of Mercury, and many other examples), then we need to check the derivation carefully. But if we do, and we find that the derivation is correct, then the prediction is convincing evidence for the theory. Historical order therefore plays no *essential* role in the process.

It is important to understand that the novel elements in Whewell's story about theory testing trace back to the initial debate about the role of concepts in scientific induction. For Whewell, theoretical concepts have an essential function in science. First, they bind the data together, which are otherwise separate and unconnected. At this stage, concepts *begin* life as conjectural elements of the inductive process, but may end up being essential elements of objective knowledge if that they prove their worth in the stringent testing that Whewell describes.

The second function of concepts is to enable us to represent what is common to otherwise disparate and unconnected phenomena. Its success is measured by their ability to extract information from one phenomenon to predict aspects of another. For example, let K be Kepler's laws applied to the motion of the moon, and G be Galileo's law of projectile motion. Prior to Newton, there was no known relation between these laws. After Newton, both phenomena were seen as arising for the influence of the earth's gravity. The objective test of Newton's theory was the agreement of independent measurements of the earth's mass from quantitative features of these phenomena. This is the consilience of inductions that justifies Newton's postulation of gravity as the common cause of the phenomena. Without the introduction of the concept of gravity, we would lack a representation of the common cause, and we would fail to learn anything about the reality *behind* the phenomena.

In sum, we argue for an interpretation of Whewell that takes much of what he says at face value, especially when he insists that the introduction of new conceptions is an *essential* part of every induction, not merely an aid in the psychological process of scientific discovery, but also as a necessary element in the expansion of our objective knowledge of the world.

2. The Role of Concepts in Science

Mill (1874, p. 208) defined induction as the operation of "discovering and proving general propositions", and Whewell (in Butts, 1989, p. 266) agreed with Mill's definition. It seems rhetorical to ask whether Whewell would agree that to infer "All humans are mortal" from the premise that "John, Peter and Paul, etc., are mortal" is an example of induction. Yet the surprising answer is that for Whewell, inferring "All As are Bs" from the premise that "All *observed* As are Bs" is *not* a genuine example of *scientific* induction. How can they disagree on this example if they agree on the definition? As Whewell explains it (Butts, ed., 1989, p.140):

> ...it appears to be frequently imagined that the general proposition results from a mere juxta-position of the cases, or at most, from merely conjoining and extending them. But if we consider the process more closely...we shall perceive that this is an inadequate account of the matter. The particular facts are not merely brought together, but there is a New Element added to the combination by the very act of thought by which they are combined. There is a Conception of the mind introduced in the general proposition, which did not exist in any of the observed facts.

It is not that Whewell is insisting upon some weird psychological thesis about what goes on in one's mind when one infers "All As are Bs" from the premise that "All *observed* As are Bs". He's objecting to this is being the proper pattern of scientific induction at all.

It follows that Whewell differs in his understanding of the word 'general' in Mill's definition of induction. What is necessary for Whewell is that (Butts, ed., 1989, p. 47) "In each inductive process, there is some general idea introduced, which is given, not by the phenomena, but by the mind." A new conception is needed in order for the facts to be "bound together so as to give rise to those general propositions of which science consists" (Whewell, 1840, pp. 201-202). Further, 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" (Butts 1989, pp. 140-141). The "pearls" are the data and the "string" is a new conception that *connects* and *unifies* the data. "All humans are mortal" does not qualify as general in Whewell's sense of the word because it introduces no *new* conception ('human' and 'mortal' do not count because they are used to describe the facts on which the inference is based).[2]

An intuitive way of characterizing Whewell's idea is that a genuine induction is an inference to the best explanation (Harman 1965), in which the explanation must introduce some theoretical "cause" that unifies and explains the facts of experience. "All humans are mortal" does not *explain* why John, Peter, and Paul are mortal by unifying those instances under a new conception. On the other hand, Copernicus explained why the retrograde motion of the outer planets occurs only in opposition to the sun by the proposition that "All planets revolve around the sun". This generalization introduces a new conception (not used to describe the phenomena to be explained), which ties the apparent motions of the planets to the motion of sun in exactly the right way. However, there is an important caveat to add. Whewell himself does not rely on a vague notion of explanation, and he is specific about what he requires in addition to the deductive subsumption of phenomena under general laws. The point of this paper is to describe the other components of Whewell's account and to argue that they are important to the *evaluation* of theories.

---

[2] Some propositions of the form "All A's are B's" may count as general in the relevant sense. For example, "All planets are moved by the gravity of the sun" is the result of a genuine induction because the unifying conception of "moved by the gravity of the sun" binds together the separate facts about the planetary motions, without being inferred directly in those facts.

Whewell introduces the special term "colligation" to refer to the conceptualization of facts. Long before Kuhn (1970) made the theory-ladenness of observation a major issue in the philosophy of science, Whewell was very explicit about the concept-ladenness of all observational facts. As Whewell describes it, colligation occurs in the most rudimentary observations:

> When anyone has seen an oak-tree blown down by a strong gust of wind, he does not think of the occurrence any otherwise than as a *Fact* of which he is assured by his senses. Yet by what sense does he perceive the Force which he thus supposes the wind to exert? By what sense does he distinguish the Oak-tree from all other trees? It is clear upon reflexion, that in such a case, his own mind supplies the conception of extraneous impulse and pressure, by which he thus interprets the motions observed, and the distinction of different kinds of trees… The Idea of Force, and the idea of definite Resemblances and Differences, are thus combined with the impressions on our senses and form an indistinguished portion of that which we consider as the Fact. (Butts, ed., pp. 123-124)

Mill claims to endorse everything Whewell says about the colligation of facts, except for Whewell's insistence that the colligation of facts is an essential part of every induction. Mill even agrees that concepts are added by the minds of scientists. Unlike Whewell, however, Mill thinks that a colligation takes place *prior* to an induction, properly so-called, and should not be conflated with it. Therein lies the first disagreement between them. Some philosophers may be tempted to dismiss this particular disagreement as largely terminological: Whewell and Mill agree that colligation occurs in science. Whewell calls every colligation an induction, whereas Mill uses the term "induction" in the more traditional sense of Aristotle and Hume. Of course, there are terminological issues involved. But it is not *merely* terminological for two vitally important reasons. (A) It is a substantive question whether examples of scientific inference may be adequately reconstructed so that all conceptualization occurs independently of, and prior to, the generalization, as Mill insists. (B) If some important scientific examples cannot be reconstructed according to Mill's description of induction, then conceptualization may be playing an essential role in the evaluation of theories, rather than merely in the psychology of discovery. We argue for both of these claims.

3. Examples of Scientific Inference

Who is right about the scientific examples when Whewell and Mill argued over the precise nature of Kepler's discovery of the elliptical motion of Mars? Both agree on the basic facts of the case, and how Kepler began with observations of the position of Mars relative to the sun at various times.[3] Naturally they differed on how these observations should be represented prior to Kepler's induction, properly so called.

Whewell and Mill agree that the conclusion of the inference is that "All points on Mars's orbit lie on ellipse b", where b specifies a particular ellipse. So, in this example, the predicate A is "is a position of Mars" and B is "lies on ellipse b." But Mill insists that the introduction of the ellipse belongs to the *descriptive* part of the process, which occurs prior to the induction itself. That is, he would say that the data are of the form "at time t_1 Mars lies on ellipse b; at time t_2 Mars lies on ellipse b; and so on." Notice that for Mill, the predicates that appear in the general

---

[3] For a description of the method Kepler used to obtain these data, see the appendix of Hanson (1973), or section 5 of Forster (1988).

proposition also appear in the description of the data, and so the inference proceeds from a statement of the form "All observed A's are B's" to a conclusion "All A's are B's."

On the other hand, Whewell insists that the conception of an ellipse is introduced as a part of the induction. Therefore, for Whewell, the data are not described using the concept of an ellipse, but in terms of more rudimentary ideas, such as 'position' and 'time'; "at time t_1 Mars is at position x_1, at time t_2 Mars is at position x_2, and so on." When Whewell (1840, pp. 201-202) says that the facts are "bound together…to give rise to those general propositions of which science consists", he is talking about facts that are not already described in terms of the binding concept. Only in that way is it true that the general proposition is not a "mere union of parts" or a "mere collection of particulars" (Butts, ed., 1989, p. 163.), or merely an extension of them.

Despite Mill's protests, Whewell (Butts, ed., p. 280) argues that "The fact that the elliptical motion was not merely the *sum* of the different observations, is plain from this, that other persons, and Kepler himself before this discovery, did not find it by adding together the observations." Like many of Whewell's assertions, this quote may be misread. It could mean that the ellipse is not the sum of different observations because it extends to new cases. Or it could mean that the ellipse is not the mere sum of observations because the observations do not contain the conception of an ellipse. Whewell means the latter. On Whewell's view, Tycho Brahe, whose data Kepler used, did not report instances of Kepler's inductive conclusion. Such instances are only *deduced* from the generalization after the induction is completed. Therefore, instances of the general proposition are not the evidence on which the induction was based.

The earlier 'mortality' example is an induction for Mill, but not for Whewell. Kepler's example is an induction for Whewell and apparently for Mill as well, though for quite different reasons.[4] Does this mean that Mill has a more liberal definition of induction? Or are there examples of induction for Whewell that Mill would exclude? Apparently, the answer is 'yes', for Mill says that "any process in which what seems the conclusion is no wider than the premises from which it was drawn, does not fall within the meaning of the term" (Mill, 1874, p.210). While such examples may be rare in practice, they are possible in principle. Here is a hypothetical example. Imagine that an economist wants to explain the inflation rate of the Soviet Union, and uses the known inflation rates for all the years from 1917 to 1990. For Whewell, this can count as a genuine inductive inference provided that the explanation introduces a new conception; maybe the concept of price control. But Mill would be forced to say that there is no induction because there are no new instances of the proposition, due to the disintegration of the Soviet Union in 1990. We think that intuition is on Whewell's side in this case, for the induced proposition does have implications about what the inflation rate *would* have been in 1991 had the Soviet Union survived. Moreover, the general proposition may be well tested by observations of other nations in relevantly similar circumstances.[5]

---

[4] Zac Ernst and Byung-Hoon Jeong have been quick to point out that Mill viewed Kepler's contribution to the induction as merely providing the right *description* of Mars' orbit. This is because he thought that (1) the periodicity of Mars's motion was already established by prior inductions, and that (2) the orbit being an ellipse is uniquely specified by those observations. That is, the inferential part was done prior to Kepler, and therefore Kepler did not perform an induction at all (Mill 1874, page 214). Mill is wrong on both counts, we think. However, even Mill could agree that "the Kepler example" is an example of induction if it includes the part that was not done by Kepler.

[5] This is one of disagreements that we have with Snyder (1997), who sees Whewellian induction as necessarily involving both the colligation of facts and a generalization to new instances. Whewell talks of generalization as being a consequence of the colligation of facts, but this is a generalization of the scope of a concept, not a generalization to new instances. If colligation of facts that does not cover new instances, it can still take part in higher level consiliences of inductions. And if it can take part in a consilience of *inductions*, then it is an induction.

4.   Empiricism about Concepts and Recent Hypothetico-Deductivism

Mill claims that the property of "lying on ellipse *b*" is part of the description of the data, and is therefore determined by, or read from, the data themselves. According to Mill (1874, p.216, Mill's emphasis) "Kepler did not *put* what he had conceived into the facts, but *saw* it in them . . ." However, his argument for this (Mill, 1874, p. 216) is entirely spurious: "A conception implies, and corresponds to, something conceived: and though the conception itself is not in the facts, but in our very mind, yet if it's to convey any knowledge relating to them, it must be a conception *of* something which really is in the facts…" Mill's idea is that the ellipse must be in the facts if the inferred proposition is true. Whewell does not deny Mill's premise (Butts, ed., p. 280): "Kepler found it in the facts, because it was there, no doubt, for one reason; but also, for another, because he had, in his mind, those relations of thought which enabled him to find it." So, the issue is not whether the orbit of Mars is really elliptical. The issue is whether Kepler, or any scientists, could *see* that it is elliptical *prior* to the induction. Whewell says no, and Mill has no convincing reply.

In fact, there is a very simple reason why Kepler could not have *seen* that Mars's orbit was elliptical, for the observed points of Mars's orbit did not lie exactly on any ellipse. Tycho Brahe's data showed discrepancies between the points on Mars's orbit and points on an ellipse. The question for Kepler was whether the discrepancies could be explained away as arising from observational errors or the action of some other law. There was no way that Kepler could *see* whether this was so from these data themselves. Whewell does not question Mill's assumption that Mars's orbit is exactly elliptical. perhaps because Newton showed how the discrepancies arise from the action of another law; namely, the gravitational pull of other celestial bodies. This view may sound a little strange to the modern ear, but it conforms to the way that Newton talked of Kepler's laws (see Forster 1988, section 5, for further discussion). While Kepler was unaware, Newton and Whewell were *fully* aware that Mars's trajectory is not exactly elliptical even within the limits of the observational precision obtained by Tycho Brahe, though we wonder whether Mill was aware of this fact.

Mill's empiricism about concepts reminds us of the logical positivists' more sophisticated view that theoretical terms should either be eliminated from science or reduced to observational statements. Both views were widely discredited during the twentieth century, which led to a revised version of hypothetico-deductivism in which the 'hypothetico' part, which includes the construction of theories and the discovery of new concepts, belongs entirely within the domain of the psychology (*e.g*, Popper, 1959). On this view, it doesn't matter where theories or concepts come from. They may come from dreams or psychic visions. The only thing of relevance to the philosophy of science is the *justification* and evaluation of theories, which depends *solely* on the observational statements that can be *deduced* from the theories.

For Whewell, mental acts are an essential part of scientific induction.   But is Whewell's thesis a claim about the psychology of discovery, or about the justification of theories? Mill (1874, p. 222) dismisses the latter possibility rather quickly by claiming that Whewell confuses invention and proof, where invention is a matter of psychology and proof is about justification: "The introduction of a new conception belongs to Invention: and invention may be required in any operation, but it is the essence of none." In truth, Whewell talks about invention (i.e. colligations) and justification (the tests of Hypotheses) separately, and he does not conflate them. Moreover, the concepts introduced by the colligations of facts do play an essential role in testing scientific hypotheses.

If this is correct, then anyone who believes that Whewell's views about testing theories are purely hypothetico-deductive is agreeing that Whewell was confused about the distinction

7

between invention and proof. At best, the only relevant issue to the philosophy of science about which Whewell was right and Mill was wrong was Whewell's objection to Mill's empiricism about the origin of concepts. We take issue with this minimization the debate. We are therefore disagreeing with Butts (1989, p. 17) when he labels Whewell as "one early form" of hypothetico-deductivism, with Jeong (1991), who endorses Butts's view, and with Ruse (1976, p. 231), who describes Whewell as combining "inductive discovery and deductive justification."[6] Our view is that Whewell's normative theory of science goes far beyond hypothetico-deductivism. In fact, we argue that the Whewell-Mill debate is an unresolved issue in contemporary philosophy of science.

Mill dismisses Whewell's ideas very quickly. Likewise, we suspect that most contemporary philosophers of science would balk at giving a justificatory role to mental concepts either because it introduces a subjective element into the confirmation of theories, or because talk of concepts is likely to be vague and unhelpful. For that reason, there is much work to be done in explicating and defending Whewell's ideas.

5. Newton's Apple and the Relational Nature of Evidence

The Whewell-Mill debate centered around Kepler's discovery of the elliptical orbit of Mars because that is the example that Mill chose to discuss. But, as we have explained, Kepler's is an example of a colligation of facts, which Whewell regards as incompletely tested at the time of its introduction.   Therefore, if our concern lies with the *confirmation* of Kepler's laws (as opposed to their discovery), then we cannot address the issue by merely examining Kepler's arguments in ever increasing detail. We must look at look at a broader historical context, which must extend at least to Newton's *Principia*. The verification of the Keplerian concept of elliptical orbits is tied up with Newton's introduction of another conception—gravity.

According to the myth, Newton thought of the idea of gravity when an apple fell on his head as he was gazing at the moon in the day-time sky. The fable is not true, but it does point to Newton's *reason* for introducing the concept of gravity, for his key idea was that the moon moving at uniform speed around the earth and the apple falling to the ground have something in common. In particular, both were *accelerating* toward the center of the earth. His solution required a re-conceptualization of acceleration as a change of *direction* and speed rather than merely a change of speed. Thus, the moon accelerating because its *direction* of motion constantly changes, while the apple accelerates because its speed constantly changes. Then the observed rates of acceleration yield measure earth's gravity influence, quantified in terms of the earth's mass, agree with each other (once Newton obtained a corrected value of the earth's radius, though that is another story). Whewell refers to this agreement as a consilience of inductions.

Note that the inductions introduced the concept of acceleration, not gravity, and the raw data involved only the positions of bodies at particular times. Instantaneous acceleration is not *seen* in the data because it is an *instantaneous* rate of change derived from a *continuous* trajectory. It is an unobserved quantity because we can only observe a finite number of points on a trajectory. For example, suppose that the true continuous motion were to contain imperceptibly small oscillations superimposed on the usual Newtonian motion. In such cases, our point observations would be the same, but the true acceleration would fluctuate by many orders of magnitude from

---

[6] Snyder, 1997, has recently argued against Butts, although she fails to make a clear distinction between the normative and psychological issues. Forster, 1988, also argues against the view that Whewell is a hypothetico-deductivist.

the constant value, *g*, inferred by Newton (the *average* acceleration would still be *g*, but *instantaneous* acceleration is the key concept introduced by Newton). So, acceleration is theoretical quantity inferred from the observations with the aid of Newton's theory. Acceleration is a concept superimposed on the facts, and is not contained in them.

In particular, Newton calculated the acceleration of the moon towards the earth using a Keplerian formula and showed that Galileo's law of projectile motion on earth is a truncated form of elliptical motion. The apple, in other words, if thrown sideways would fall on a parabolic trajectory, and this motion would continue as a Keplerian ellipse were it not for its impacting the surface of the earth.[7] He then showed that these formulae were derivable from the inverse square law of gravitation, thereby *unifying* two previous unconnected domains of data; terrestrial phenomena on the one hand and lunar motion on the other.

For Whewell, Kepler's colligation of the facts is confirmed and interpreted by the successful unification under Newton's theory of universal gravitation. In particular, the appropriateness of Newton's conception of instantaneous acceleration was confirmed by the fact that the same constant *g* appears in the description of disparate phenomena, such as the motion of pendula, the balancing of beams, and projectile motion. This is a further *consilience of inductions* and a further confirmation of Newton's inverse square law conception of gravity. In contrast, there is no consilience that justifies the introduction of small oscillatory motions, even though they are consistent with the observed facts (Forster, 1988). Thus, the successful generalization of Newton's concept of acceleration further justifies Kepler's induction.

What is being confirmed is not the literal truth of every aspect of Kepler's laws, for at the same time, Newton's theory lead to a numerical correction of the predictions provided by the laws. This is an important qualification, which is taken up in more detail in section 10.

By definition, a consilience of inductions leads to a unification of two colligations of facts because an induction *is* a colligation of facts for Whewell, and colligations involve the introduction of a new conception. Therefore, Whewell's story necessarily involves conceptions. But is Whewell right? Isn't the role played by concepts mysterious and vague? Can't we do without them? This is certainly Mill's view of the matter. For Mill concepts are merely a necessary part of the psychological habits of human beings, and they are not essential to the story about the justification of induction. Mill therefore replaces Whewell's consilience of inductions with what he refers to as the deductive subsumption under a simple set of higher-level laws. As Mill describes it, "The subsumption (as it has been called) of one law under another [is] the gathering up of several laws into one more general law which includes them all" (Mill, 1874, p. 335). To "include" the several laws in more general law means to *deduce* several laws from the more general one, as Mill makes clear. Thus, the subsumption under general laws is a *deductive* account of theory organization and structure.

Mill quickly adds that the laws of nature state the fewest and simplest assumptions, which being granted, the whole existing order of nature would result (Mill, 1874, p.230). This means that Mill is adding an element that goes one step beyond a purely deductive account. Naturally, Mill goes on to claim that the "most splendid example of this operation was when terrestrial gravity and the central force of the solar system were brought together under the general law of gravitation" (1874, p.335).

---

[7] In the Scholium after Proposition X, Problem V, Book I, Newton puts it this way: "If the ellipse, by having its center moved to an infinite distance, degenerates into a parabola, the body will move in this parabola; and the force, now tending to a center infinitely remote, will become constant. This is Galileo's theorem."

There are two questions to consider concerning Mill's proposal. (1) Can the role of concepts be eliminated in some way yet to be specified? (2) Can the role of concepts be eliminated in the way that Mill proposes? We doubt that the answer to the first question is affirmative, and we answer the second question strongly in the negative.

It's not that Whewell denies that there is a deductive subsumption of Kepler's and Galileo's laws under Newton's laws. More is going on, and the extra part is relevant to the testing of hypotheses. Recall the earlier point of disagreement about whether induction begins from the instances of the induced laws deduced from them. For Whewell, it does not begin with deduced instances if these are described in terms of the conception introduced during the induction. In other words, the colligation of facts (*i. e.*, induction, for Whewell) begins with uncolligated facts (that is, uncolligated by the *new* conception). Therefore the initial inductions must introduce new concepts, and the consilience of inductions leads to the "jumping together" of those concepts.

Mill recognizes that simplicity is an essential part of the story, but he vaguely alludes to a notion of simplicity that makes no reference to the generalization of concepts. The trouble with Mill's alternative is that he does not tell us what counts as "fewest" or what counts as simple. Aside from a few examples, he gives no definition or analysis. More importantly, the idea does not stand up to critical scrutiny.[8] To see why, let G be Galileo's theory of terrestrial motion, and K be Kepler's theory of celestial motion. Presumably, for Mill, K and G would not be subsumed under the conjunction $K\&G$ because there are two assumptions here, rather than one. The problem is that N could also be expressed as two assumptions, or as three, or as however many assumptions one is ingenious enough to devise. On the other hand, $K\&G$ could be viewed as one statement. Why should it count as two? How does Mill answer these questions without reverting to Whewell's idea about the unification of concepts? Therefore, there is nothing concrete in Mill that goes beyond saying that N deductively entails G and K. Unfortunately for Mill, the mere deducibility of K and G from N is not sufficient.

Whewell's account of consilience succeeds where Mill fails, for Whewell *can* explain why the mere conjunction $K\&G$ does not unify the phenomena. He can point out that Newton's conception of instantaneous acceleration uncovers an *agreement* between G and K, which confirms that Newton's inference, and leads to an interpretation of G and K as arising from the action of gravity. The mere conjunction of $K\&G$ does not introduce Newton's unifying conception of acceleration and produces no consilience.

In the pages immediately following Whewell's description of his tests of hypotheses (section 6), Whewell devotes considerable space to the need for checking the deductive relationship from theory to fact. It is this chapter VI "Of the Logic of Induction" in the *Novum Organon*, Whewell (Butts, ed., 1989, p. 175.) looks like a hypothetico-deductivist: "The special facts which are the basis of the inductive inference, are the conclusion of the train of deduction," "Deduction is a necessary part of Induction," "Deduction justifies by calculation what Induction had happily guessed," and so on.

Yet, at the same time, Whewell repeatedly emphasizes that:

> when we say that the more general proposition *includes* the several more particular ones, we must recollect what has before been said, that these particulars form the general truth, not by being merely enumerated and added together, but by being seen *in a new light*. No mere verbal recitation of the particulars can decide whether a general proposition is true... On the contrary, the Inductive truth is never the mere *sum* of the facts. (Butts, ed., 1989, pp. 169-170)

---

[8] The same debate has been played out in the contemporary literature, when Friedman (1974) proposed something along Mill's line, which Kitcher (1976) rebutted.

That inductive truth is not the mere sum of the facts does not mean that that truth of a general proposition is *more* than the sum of its deductive consequences. For it is a standard theorem of logic that any statement is equivalent to the conjunction of all its deductive consequences. However, if the deductive consequences are restricted to facts that formed the basis of the inductive inference, the general proposition is more than the sum of those deductive consequences.

For example, let E_K denote the data from which Kepler's laws, K, was inferred. K differs from E_K by its addition of the ellipse. Whewell (Butts, ed., 1989, p. 301) actually agrees with Mill's statement that: "if Newton had been obliged to verify the theory of gravitation not by deducing from it Kepler's laws, but by deducing all the observed planetary positions which had served Kepler to establish those laws, then Newtonian theory would probably never have emerged from the state of an hypothesis." Whewell (Butts, ed., p. 300) even agrees with Mill that the purpose is "to compare the results of the deduction, not with one individual instance after another, but with general propositions expressive of the points of agreement which have been found among many instances." However, according to Mill's empiricist view of concepts (section 4), the "points of agreement" found among many instances are determined by E_K.

Mill therefore makes no distinction between the *empirical* content of K and the *empirical* content of E_K, so the confirmation of Newton's laws would be served equally well by the deduction of either of them from Newton's Laws. At first sight, this seems to be a reasonable statement. After all, confirmation means *empirical* confirmation, and K and E_K share the same *empirical* content, even if they do not have the same theoretical content. Yet this is exactly the claim that we think Whewell should deny. Perhaps the best way to explain the point is to divide empirical content into two parts; a non-relational part, and a relational part. Thus, the empirical content of E_K, appears to consist of the non-relational positions of the planets at particular times, together with the relational fact that these positions collectively fit to an ellipse. However, the second claim is not quite correct, for the raw data also fit many other curves. The relational facts contained in E_K are therefore disjunctive in nature and logically weak.

The situation for K is different because the ellipse relation has *been* superimposed upon the facts. Moreover, K smoothes over the data, and therefore sacrifices information about the observed positions in favor of particular relations amongst the data. While this content may be conjectural and unconfirmed, it is empirical. There is therefore a huge difference in the empirical content of K and the empirical content of E_K.

An immediate retort, on behalf of Mill, is to say that there is a confusion here between the total empirical content, which includes unconfirmed instances, and the *observed* empirical content. E_K is the observed data, and its content exhausts the empirical content of K *that has any role in the justification of Newton's Laws*. K is a convenient surrogate for E_K, which works because a deduction of K from Newton's Laws is sufficient to show that Newton's Laws provide a good fit with the data in E_K given that K provides a good fit with E_K.

To see the fallacy in this argument, let N be the version of Newton's Laws that applies to an idealized earth-moon system,[9] such that the mass of the earth, M, and the mass of the moon, m, are adjustable parameters. Then from N we may deduce that the earth and moon follow the elliptical paths around their common center of gravity at a focus of each ellipse such that the line from the line from the center of gravity sweeps out equal areas in equal times. From this, it also follows that the moon travels in an elliptical path with the earth at one focus, and the line from

---

[9] In other words, let us suppose that the earth and the moon are spherically symmetric bodies, and the gravitational influence of the sun, and other bodies is neglected.

the earth to the moon sweeps out equal areas in equal times. This last consequence is K, Kepler's Laws applied to the moon. K is a disjunctive set of many possible elliptical motions of the moon around the earth. So, K does not imply the particular motion followed by the moon, and therefore does not entail the observed data E_K. That is, there is no deductive relationship between K and E_K.

Nevertheless, this seems like an easy detail to fix. All we need to do is to use the known data to fix the adjustable parameters M and m (or at least their ratio) so that N provides the best fit to E_K. If we plug these values into N, then we get a particular Newtonian hypothesis, which we might call $N(E_K)$, since the hypothesis picked out depends on E_K. Now, from $N(E_K)$ we can deduce $K(E_K)$, which is a particular elliptical motion, from which we can predict E_K, at least approximately.

A potential problem is that the process is circular: We have used E_K in order to predict E_K. Of course, in this case the process is not viciously circular, and there is real confirmation going on. But here is the crucial question. What distinguishes this case from one that *is* viciously circular?[10] This may sound like a trite question, but it is crucial to understanding the role of relational evidence.

A sufficient answer is to point out that N, or K, have been fitted to data, and they have had their predictions tested against *fresh* data. For example, the ellipse that Kepler fitted to Tycho Brahe's data led to predictions that have tested against data collected after Tycho Brahe. In other words, Kepler's Laws have been fitted to one set of data, E_1, and the ellipse singled out as best fitting E_1, has been tested against an entirely new set of data E_2. These laws, in other words, have *an established history of successful projection*. Because E_1 and E_2 are *different* sets of data, there is no circularity involved. This record of successful projection provides strong evidence that the conception of the ellipse has succeeding in capturing a relation between E_1 and E_2, which is a relational fact about the evidence E_K.

Many readers will agree that successful projection in the past is a *sufficient* condition for E_K to provide a non-circular test of K. But is it *necessary* to look at it this way? Can't we just point to the fact that K fits the total evidence E_K well. To say that K fits E_K well is to say that K *accommodates* the data E_K. So, the question is whether accommodation and prediction have an equivalent evidential status. It has always been the bug-bear of hypothetico-deductivism that it has failed to draw a logical distinction between them. Rather, the hypothetico-deductivists have seen the distinction between prediction and accommodation as relying on the historical order in which the data is collected and applied. We wish to argue that there is logical distinction between them, though not one that is captured in Mill's terms.

In fact, we wish to argue that the prediction test is a stronger than the accommodation test. Our argument for this claim is straightforward: There are examples in which accommodation would be successful, but the prediction test would fail. Here is one such example: Suppose our hypothetical law says that a quantity y is a function of x, where the form of the law connecting them is a polynomial with 100 terms. Suppose the data consists of 100 observed values of pairs (x, y). To relate this to the Kepler example, think of x as 'time' and y as the position of Mars. Then it is a mathematical theorem that one of the family of 100-degree polynomials will fit the data perfectly and therefore accommodate the data perfectly. Now suppose we break the data into two subsets, with 50 data points in each. Call the first data set the calibration data and the

---

[10] Glymour (1980b) put forward a bootstrapping account of confirmation as a way of confronting the circularity problem, which it does by simply ruling out the case in which the circularity is vicious (the non-trivialization clause). We note that Whewell's theory is different and has not received the attention it deserves.

second set the test set (Browne 2000). The prediction test requires that we fit the polynomial to calibration data and test its predictions on the test data. Will the fitted polynomial successfully pass this test? No, it will not. In fact, there will be infinite subset of polynomials that fit the calibration data perfectly. One of these will fit will also fit the test set perfectly, but this merely an accommodation of the test data, and not a successful prediction of it. (Notice that the argument works no matter how the calibration and test data are chosen. It has nothing to do with the historical order in which the data are collected. We return to that point in a later section.)

A common intuition about this example is that we should use simplicity to rule out laws with high numbers of parameters, so that accommodation *is* a meaningful test of hypotheses. In fact, there are sophisticated criteria that correct the degree of fit of the total data to take account of the accommodation effect (called overfitting), in order to judge the predictive accuracy of hypotheses. One such method is described in Forster and Sober (1994). Perhaps the same idea could be applied to a hypothetico-deductivist approach?

However, there is some reason to believe that correction for overfitting still provides a weaker test than looking at the past predictive success. The argument derives from the fact that overfitting disappears as data sets become larger, so the correction for overfitting disappears in that limit. Therefore, the criteria that take the number of adjustable parameters into account reduce to measures of accommodation in the large sample limit. The problem is that more complex hypotheses accommodate the data better even in this large same limit, yet they do more poorly at predicting data *of a different kind*. One might say that the solution is to put a greater weight on simplicity, but nobody knows exactly how to do this. So, in the absence of alternatives, the idea of testing hypotheses in terms of their *past* projections is a sensible alternative. Computer simulations performed by Busemeyer and Wang (2000) and Forster (2000) verify the merit of this idea.[11]

The point is that the difference between Mill and Whewell on testing hypotheses is a live issue in theoretical statistics, and the merit of Whewell's ideas go far beyond the psychology of discovery. There is plenty of evidence against the claim Mill's subsumption under general laws is an adequate surrogate for Whewell's tests of hypotheses.

Kepler's Laws are therefore subjected to a strong Whewellian test when they are fitting to data pertaining to Mars's orbit at one time and tested against new observations of Mars's motion. However, Whewell's tests of hypotheses does not stop there. He extends the same idea to consilience between different colligations of facts. That is, Newton's equations N succeed in relating fact about the moon's motion, E_K, to facts about the motion of projectiles on Earth, E_G, pertaining to Newton's apple and other terrestrial projectile. The conjunction of Kepler's and Galileo's laws conjunction does not past the same test because the conjunction does allow for any prediction from E_K to E_G, or *vice versa*. It is crucial to understand that it is not predictions simpliciter that are at issue. Galileo's Laws are already sufficient for the purpose of predicting the motion of projectiles. The question is whether any predictions can be made *solely* from the moon's motion.

---

[11] It is important that the choice of calibration data and test data has a kind of asymmetry built into it, as does Whewell's consilience of inductions, which tests prediction from one induction to another. A well known way of making cross validation 'direction-free' is as follows: If there are n data, perform n cross-validation tests by calibrating on all the data except a single data point, testing on that data point, and averaging the results of the n tests. This leave-one-out cross-validation is asymptotically equivalent to AIC (Stone, 1977), which one of the well known criteria that corrects for overfitting by taking the number of adjustable parameters into account (Forster and Sober 1994).

Such a connection is imposed by Newton's new concept of acceleration, together with the inverse square law of gravitation. The gravity of the Earth, measured by its mass, captures what is common to the acceleration of the moon and the acceleration of the apple. This relational content that is not contained in the mere conjunction of *K&G*. The confirmation of this relationship is the evidence for moving beyond the laws of phenomena, described in terms of *K* and *G*, to the laws governing the causes of the motions. The reason does not depend on some vague notion that Newton *explains* the phenomena more deeply than the combination of Kepler's and Galieo's laws. Rather, the test is whether Newton postulation of a common cause provides a cross prediction that is confirmed by the a numerical agreement of two independent measurements of the earth's mass.[12]

The complaint that Whewell's concept of colligation and consilience is vague has a two-part reply. First, to the extent that it is vague, it is nevertheless better than no theory at all. Second, Whewell has far more to say about the colligation of facts and the consilience of inductions than we have mentioned here. The following sections present some of those details, but ultimately any judgment of the vagueness of Whewell's account should be made only after one is fully conversant with Whewell's writings.

Unfortunately, most, if not all, philosophers of science who have read Whewell have failed to see that the differences between Whewell and Mill are highly relevant to the normative evaluation of theories. Even less have they seen that Whewell's writings contain important insights into the issues facing scientists and statisticians today. We hope that the briefly sketched arguments of this section will help philosophers and historians of science to read Whewell as describing an interesting normative theory of science, and not merely a psychology of discovery.

6.   The Consilience of Inductions and other Tests of Hypotheses

In this section, we will explain the role that colligation of facts plays in Whewell's story about justification. Be warned, however, that Whewell's words are easily misunderstood when he talks about evidential and psychological considerations in the same breath. Consequently, we admit that it is hard to separate the normative and psychological aspects of his theory. Yet the effort will reward us with a novel view of empirical evidence, which emphasizes the importance of relational features amongst the data.

Whewell distinguishes four tests of hypotheses.

Test 1) The initial step in verification is rigorously reapplying the hypothesis to the observed phenomena and establishing that the induced proposition accurately represents all of the facts of observation.

"We are never to rest in our labours or acquiesce in our results, till we have found some view of the subject that is consistent with *all* the observed facts…" (in Butts ed. 1989, p.151). In contemporary terms, this is no more than the weak requirement that a hypothesis *accommodates* the known data. It is a mistake to interpret the word "consistent" in the logical sense, because that would require that the hypothesis must fit the fact *perfectly*. It is clear that this is not what Whewell means when he talks about curve fitting methods, such as the method of least squares (Butts, ed., pp. 233-234). As such, Whewell is aware that this test is somewhat vague and fairly easy to satisfy. He therefore proceeds to his second test:

---

[12] This is the basis of the reply given in Forster (1988) to Cartwright's (1983) allegation that the existence of component causes is not empirically grounded in the observational facts.

Test 2)   "Hypotheses," says Whewell, "[O]ught to do more (than account for what has been observed): . . . our hypotheses ought to *foretell* phenomena which have not yet been observed; at least all phenomena of the same kind as those which the hypothesis was invented to explain . . ..That it does this with certainty and correctness, is one mode in which the hypothesis is to be verified as right and useful" (in Butts, ed. 1989, p.151).

A hypothesis is subjected to test 2 when it is tested again against data not used in the construction or calibration of the hypothesis, provided that the test data is of the same kind. So this includes the case in which Kepler's Laws fitting to Mars's orbit at one time is tested against new observations of Mars's position.

Whewell also acknowledges that the history of science contains many examples of false hypotheses that made some correct predictions (e.g. the theory of phlogiston); he does not, therefore, think that a limited predictive ability is sufficient to convince us of the truth of a hypothesis. It is when a hypothesis passes a more stringent test, the *consilience of inductions*, that, according to Whewell, the evidence leads us to the conviction that the theory is true:

Test 3) The *consiliences of inductions* offer the most convincing evidence for a hypothesis' validity in either of two circumstances:

(a) When a hypothesis "enables us to explain and determine cases of a *kind different* from those which we contemplated in the formation of our hypothesis" (Butts, ed., 1989, p. 153).

(b) "[W]hen the explanation of two or more different kinds of phenomena (as the revolutions of the planets, the fall of a stone, and the precession of the equinoxes,) lead us to *the same* cause, such a coincidence give a reality to the cause. We have, in fact, in such a case, a Consilience of Inductions." (Butts, ed., 1989, p. 296).

The literature on the consilience of inductions either focuses on (a) or describes (b) in terms of inference to the best explanation without emphasizing that in *both* cases, there is a prediction test involved. Predictions are explicitly mentioned in (a), so our claim is clear in that case. However, in (b), Whewell is referring to a particular kind of explanation, which has led "to *the same* cause." In the case of the quantitative sciences, the successful explanation of the two kinds of phenomena leads to an agreement of the independent measurements of postulated common cause; namely, the gravitational mass of the earth, in our example. Inference to the best explanation has been criticized as being too vague (e. g., van Fraassen 1980), so this constraint on the kind of explanation required is philosophically important.

In *Laws of Phenomena and of Causes* (Butts, ed., 1989, pp. 177-185), Whewell distinguishes between lower-level statements (Laws of Phenomena) and higher-level theories (Laws of Causes) that tell us *why* phenomena occur. A law of causes describes the causes of these phenomena. The law of causes is *confirmed* to the extent that it explains disparate phenomena in terms of a common cause:

That the cause explains *both* classes, gives it a very different claim upon our attention and assent than from that which it would have if it explained one class only. The very circumstance that the two explanations coincide, is a most weighty presumption in their favour. It is the testimony of two witnesses in behalf of the hypothesis; and in proportion as these two witnesses are separate and independent, the conviction produced by their agreement is more and more complete. (Butts, ed., 1989, p.330).

There are two important features of this passage. One is Whewell's insistence that the two explanations should coincide. We believe that the coincidence is an empirical test, when either (a) the theory determines, or predicts, cases of a kind different or (b) there is a coincidence

between the values of independently measured quantities. Both of these descriptions amount to the same thing in terms of the logic of the situation. To determine cases of a different kind requires us to *assume* the agreement of the unifying quantity, such as gravity, and make predictions about the second phenomenon. Or else, we can accommodate the theory separately to the two kinds of data, obtain independent measurements of the postulated common cause, and see that they agree. A theory will pass the test in the first way if and only if it passes the test in the second way. Each story may suggest a different historical sequence of events, but this order of events need have no bearing on the strength of the test. We return to this issue in section 10.

The second point is that Whewell talks of the *weight* of presumption and completeness of the agreement. It is a common misconception of Whewell that he thinks that the consilience of inductions is conclusive proof of the truth of the hypotheses. The misimpression arises from passages like this:

> The instances in which [the consilience of inductions] 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 unseen 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 the truth resides. (Butts, ed., 1989, p. 153)

The thesis is that the consilience of inductions *impress us with a conviction* that the truth of our hypothesis is certain. Given his explicit mention of degrees of conviction in the previous passage, it is charitable to assume that Whewell means the subsequent sentences to be an elaboration of the same psychological thesis. We return to this point in our discussion of the evidential nature of necessity in section 8.

Whewell's picture of common cause explanation is not the same as the modern principle of common cause made famous by Reichenbach (1956). Anyone familiar with the modern conception might think an induction leads to the law of phenomena connecting observable quantities X and Y, such that X is correlated with Y. For example, the ancients knew that the phases of the moon (X) are correlated with the height of high tides (Y). But this is a phenomenological law par excellence, for it would be mystical to think that patterns of light on the moon cause the tides. So, to explain the correlation, we must discover the causal law by which the same third observable quantity W is the common cause of X and Y. For example, W might be the relative position of the earth and the sun, which causes the phases of the moon, and accounts for the movements of the tides at the same time. That is a typical example of how causal explanation is understood in Mill's four methods of experimental inquiry (Mill, 1874, Ch. VIII, Book III). On Mill's account of causal explanation, one doesn't need new concepts to describe the effects to be explained.

However, this not an example of Whewellian common cause explanation. First of all, the fact that the sun causes the illumination of the moon was a causal law that was well known before Newton. So, in Newton's time there is only one law of phenomenon in this example; namely, the correlation between W and Y. This law of phenomenon was explained by Newton in terms of a law of causes; namely, by a composition of the effects on the tides of the moon's and sun's gravitational pull (see Forster 1988 for discussion of Whewell's treatment of this example). A Whewellian kind of common cause explanation enters this example only when Newton derives other observable of effects of the moon's and the sun's gravity, for example on the orbit of the earth and the moon system around the sun.

Van Fraassen (1982) has argued correctly that Reichenbach's principle is refuted by some quantum mechanical correlations (Bell 1964). A noticeable feature of these quantum mechanical correlations is that quantum theory provides no separate explanations for the occurrence of each variable. In Bohm's version of the example, each correlated variables are the outcomes of the measurement of electron spin, in which there are two possibilities—up and down. If the measurements are simultaneous (space-like separated), then quantum theory gives the probability of ½ for each outcome conditional on the entire state of the universe prior to the measurement. So quantum theory provides no separate colligations of these facts, and therefore no consilience of *inductions* in the example, and therefore no common cause is *needed* to explain the correlation.

It would be nice if a more complete physics were able to provide a separate colligation of the facts, as Einstein believed, but wishful thinking cannot undo the proof that it cannot work (without violating locality or some other plausible constraints). The local hidden variable interpretation of quantum mechanics *attempts* to introduce common causes to explain the correlations, but makes false predictions that are refuted by experiment. It fails because its predictions about data of a different kind turn out to be false.

In sum, Whewell's theory says the right thing in both cases. (1) Not every correlation is a consilience of inductions, and there no consilience of inductions in Bell's example, and therefore no common cause is needed. (2) The local hidden-variable hypothesis fails the test of consilience, and therefore should be rejected.

In Newton's argument for universal gravitation, the common cause was represented by the gravitational mass of the earth. In this case, "the same cause" refers to the same *particular* cause. However, it is also clear that in other cases, the common effects such as "the revolutions of the planets, the fall of a stone, and the precession of the equinoxes," do not all refer to the same *particular* cause because the fall of the stone and the precession of the equinoxes are effects of the earth's gravitational influence and the *sun's* gravitational pull, respectively. In that case, the consilience of inductions cannot refer to the agreement in the measurements of the same quantity. Instead, it must refer to something like the common *form of the equations* governing those motions, or the fact that they separately exhibit the same *kind* of cause. A clearer example would be to count the observation of a twin-star system undergoing Keplerian motion as a kind of consilience with other applications of Kepler's laws. We shall distinguish the two kinds of consilience as quantitative consilience (the numerical agreement of independently measured quantities) and formal consilience (agreement in the form of explanations).

This is the kind of consilience is found Darwin's argument for natural selection. Perhaps that is why Darwin was so pleased with Whewell's methodology. No single example in the *Origin of Species* provided convincing evidence of the existence of evolution, and it was also difficult to meet the deductive requirements of Mill's hypothetico-deductive scheme. But it was possible to convince his readers that the pieces of evidence fit together such that it would be implausible to dismiss the overall pattern of the evidence as a mere coincidence.[13]

Nevertheless, the two cases have much in common. Each has a kind of persuasiveness that rests on coincidence argument—that such an agreement would be improbable on the basis of

---

[13] Whewell did not accept or endorse Darwin's theory. That is another story, however, which involves discussion of Whewell's personal relationship with Darwin and Whewell's own theological commitments. The interested reader should see Ruse (1975) and Ruse (1976).

chance alone. In both cases, a kind of likelihood argument[14] applies, in which one argues that the unified theory makes the agreement more probable than it would be on the basis of chance alone. This is not evidence for this particular theory as opposed to all other theories, since there are many extensions and modifications of the theory that may exhibit the same consilience. Rather it is evidence that the concepts introduced have latched onto *something* real, and the phenomena really do have *something* in common. More specific explanations of the coincidence may have to tested by a more general consilience with as-yet-undiscovered phenomena.[15]

The difference is that the agreement between the independent measurements is more easily thought of as an *empirical* fact, whereas the agreement in the form of different laws is not. However, from a Whewellian point of view, there is not so much difference between them. The agreement of independent measurements depends on the theoretical causes introduced by the theory, as is the case of gravitational mass. Likewise, the agreement in the form of laws requires that the two laws actually fit their respective phenomena, so formal consilience has an empirical component. Both kinds of consilience have theoretical and experiential elements, exactly as Whewell describes in his fundamental antithesis of philosophy (which we discuss in the next section).

The idea that formal consilience is (in part) *empirical* evidence in support of a theory is something that is denied by standard deductive and probabilistic accounts of evidence, which assume that evidence must be described independently of the theory under test. The agreement of equation forms cannot be described independently of the theory under test. If this is right, then it is a surprising endorsement of Kuhn's (1970, p. 85) use of the metaphor of a duck-rabbit gestalt to claim that different theories place "the same data as before in a new system of relations with one another by giving them a different framework." Whewell and Kuhn have a similar view about concepts imposing relations amongst the data, except that Whewell is a realist about the existence of such relations in the world. For Kuhn, the concept-ladeness of observation counts against the objectivity of science, while for Whewell it is a necessary step on the path towards truth.

Finally, we arrive at Whewell's fourth test of hypotheses. If a theory redescribes the phenomena as effects of a common cause, then the theory represents the phenomena in simpler and more unified way than if the causes were separate. For subsequent generations of scientists immersed in the theory, the failure of consilience is then psychologically impossible to conceive because it is a necessary consequence of the unified theory (Glymour 1980a) . Whewell describes this tendency towards unity and simplicity as a further test of hypotheses:

> Test 4) "The last two sections of this chapter direct our attention to *two* circumstances, which tend to prove, in a manner which we may term irresistible, the truth of the theories which they characterize:—the *Consilience of Inductions* from different and separate classes of facts;—and the progressive *Simplification of the Theory* as it is extended to new cases. These two Characters are, in fact, hardly different; they are exemplified by the same cases. ... The Consiliences of our Inductions give rise to a constant Convergence of our Theory towards Simplicity and Unity." (Butts 1989, p. 159).

---

[14] "Likelihood" is a technical term that refers to the probability of evidence given a theory rather than the probability of a theory given the evidence. Likelihood arguments do not require that theories have probabilities, which is a distinctly Bayesian idea.

[15] For instance, Belot (1998) has argues that Aharanov-Bohm effect (which is a quantum effect) teaches us that the vector potential in classical electromagnetism should be interpreted as real (contrary to traditional interpretations).

Many Whewell scholars, such as Blake *et al* (1960, p. 212) and Butts (1973, p. 125), believe that the consilience of inductions and the simplicity or unity of theories amount to the same thing for Whewell. We admit that they are intimately connected—after all they are "exemplified by the same cases." Nevertheless, like inertial mass and gravitational mass, they are conceptually different. To treat them as identical would not make sense of Whewell's claim that one "gives rise" to the other. Perhaps the issue is whether the convergence towards simplicity and unity is an *independent* test. Perhaps it is just another way of recognizing the presence of a consilience of inductions, just as we argued that descriptions (a) and (b) of test 3 were equivalent from a logical point of view. If that is the issue, then we would tend to agree that test 4 is really a different way of seeing that the theory has passed the consilience of inductions test.

An important conclusion of this section is that Whewell's consilience of inductions requires that "inductions" be understood as colligations of facts, and not as "inductions" in Mill's sense. The agreement of independently estimated parameter values, or an agreement between form of laws, is undefined prior to the introduction of the form of the law. The form of the law (e.g., Kepler's ellipse) is *added* to the facts, *and it is not contained in them*. For Whewell, the formula itself is the new conception superimposed on the data in the case of curve fitting examples (Butts, ed., 1989, pp. 213). Therefore, in curve-fitting, there is no consilience of inductions without inductions *in Whewell's sense*. Hence, any reason to value the consilience of curve-fitting inductions is a reason to believe that Whewell is right and Mill is wrong about the nature of induction in science.

7. The Fundamental Antithesis between Subject and Object

According to Whewell, conceptions are added by the mind, and are therefore "subjective" components of theories. A central thesis of Whewell's epistemology is that all objective knowledge must necessarily arise through the interplay of "subjective" and "objective" elements. "We can have no knowledge," says Whewell, "except that we have both impressions on our senses from the world without, and thoughts from our minds within" (Butts, ed., 1989, p.57).

Whewell's point is not to divide our beliefs into two mutually exclusive categories; subjective beliefs on the one hand, and objective knowledge on the other. For Whewell, the objective and subjective elements are inseparable parts of all beliefs about the world. Mill and Whewell share the same optimism about the possibility of objective knowledge. The difference lies in their understanding of foundation of knowledge and the processes the lead us towards the truth. For an empiricist like Mill, science is founded on observational beliefs that are of the most objective and incorrigible kind. If the processes of induction are to produce objective conclusions, then they do so by insulating the process from the subjective elements. For Whewell the reasoning process *introduces* subjective elements at every stage of the inductive process. So, it is an incidental fact for Whewell that the observations themselves are necessarily concept-laden:

> [I]f we take the terms Reasoning and Observation; at first sight they appear to be very distinct. Our observation of the world without us, our reasonings in our own minds, appear to be clearly separated and opposed. But yet we shall find that we cannot apply these terms absolutely and exclusively. I see a book lying a few feet from me: is this a matter of observation? At first, perhaps, we might be inclined to say that it clearly is so. But yet, all of us, who have paid attention to the process of vision, and to the mode in which we are enabled to judge of the distance of objects, and to judge them to be distant objects at all, know that this judgment involves inferences drawn from various sensations; -- from the impressions on our two eyes; --

from our muscular sensations; and the like. These inferences are of the nature of reasoning. . . .
All observation involves inferences, and inference is reasoning. (Butts, ed., 1989, p.61)

 Whewell's vision of science is fundamentally different from Mill's. His idea is that the testing of hypotheses plays the role eliminating or correcting the subjective elements (the conceptions) that fail his tests of hypotheses. For Whewell, science is an active process of seeing the truth through the subjective constructs of the mind, whereas for Mill it is a passive process of shielding science from the intrusion of subjective elements at every stage of the process.

Whewell's fundamental antithesis between "object" and "subject" manifests itself in many different ways. We have seen how observation is the result of sensation and inference. The same idea manifests itself as an antithesis between fact and theory. Facts are on the objective side of the dichotomy, while theories are the products of our mind, on the subjective side. However, there is not absolute distinction between fact and theory for Whewell. For the claim that counts as a theory in one context may play the role of a fact in another:

> Is it a Fact or a Theory that the planet Mars revolves in an Ellipse about the Sun? To Kepler, employed in endeavoring to combine the separate observations by the conception of an Ellipse, it is a Theory; to Newton, engaged in inferring the law of force from a knowledge of elliptical motion, it is a Fact. There are, as we have already seen, no special attributes of Theory and Fact which distinguish them from one another. Facts are phenomena apprehended by the aid of conceptions and mental acts, as Theories also are. (Butts ed., 1989, p.176)

This sets the stage for Whewell's special view of how it is possible for our objective knowledge to accumulate in spite of the unavoidable presence of subjective elements. In Whewell's words:

> In the progress of science, both the elements of our knowledge are constantly expanded and augmented. By the exercise of observation and experiment, we have a perpetual accumulation of facts, the materials of knowledge, the objective element. By thought and discussion, we have a perpetual development of man's ideas going on: theories are framed, the materials of knowledge are shaped into form; the subjective element is evolved; and by the necessary coincidence of the objective and subjective elements, the matter and the form, the theory and the facts, each of these processes furthers and corrects the other: each element moulds and unfolds the other. (Butts, ed., 1989, p.75).

So, in spite of the fact that our representation of the world is subjective in the beginning, in the middle and at the end, our subjective representation of the world can evolve towards the truth. For Whewell there is no contradiction is saying that a theory is the invention of mind and a true fact at the same time. Whewell certainly made both claims about Newton's theory of motion. Not all theories are true, although most theories in the history of science have a measure of truth. By the process of theory testing described in the previous section, there is a process of elimination, reinterpretation and correction that leads to objective progress in science. We shall say more about the process of correction in section 10.

8.   The Evidential Nature of Necessity

The opposition between subject and object has many manifestations for Whewell, including "Necessary and Experiential Truth, Ideas and Senses, Thoughts and Things, and Theory and Fact" (Butts, ed., 1989, p. 57). Notice that necessity falls on the *subjective* side of the dichotomy in this quote. According to Whewell's definition (Butts, ed., 1989, p. 55) "necessary truths are those *of which we cannot distinctly conceive the contrary*." The reference to "what we cannot be distinctly conceive" has to be read literally. For Whewell, necessity is a *psychological* property of theories. This is the opposite of what we are taught in contemporary philosophy, where

'necessity' is an objective commodity, as in "logical necessity' or "physical necessity," and experience is on the psychological side of the divide.

Unfortunately, Whewell muddies the waters by talking about necessary *truth* in his definition, for the concept of 'truth' falls on the 'object' side of the antithesis. He sometimes slips from one to the other, perhaps because he thought, as a matter of *contingent* fact, that laws that achieved the status of being necessary were always true.[16] It does not mean that laws are justified as true *because* their contrary is inconceivable. Mill's misunderstanding of Whewell leads to the following bewilderment on exactly this point:

> Now I can not but wonder that so much stress should be laid on the circumstance of inconceivableness, when there is such ample experience to show, that our capacity of conceiving a thing has very little to do with the possibility of the thing in itself; but is in truth very much an affair of accident, and depends on the past history and habits of our own minds. There is no more generally acknowledged fact in human nature, than the extreme difficulty at first felt in conceiving any thing as possible, which is in contradiction to long established and familiar experience; or even to old familiar habits of thought. And this difficulty is a necessary result of the fundamental laws of the human mind. (Mill, 1874, pp.177-178)

Mill is correct to say that necessity, in Whewell's psychological sense, is a poor criterion on which to justify the truth of theories. On the other hand, if *empirical* theories attain this psychological status if and only if they are *empirically* well established, then there is no reason why necessity cannot be an *indicator* of truth. Because this is also a source of controversy in contemporary Whewellian scholarship, we shall examine it more fully.[17]

Butts (1989, p. 23) is someone who follows Mill in complaining that Whewell "seems to have confused psychological conviction with objective empirical truth, and with inferential validity." As Butts points out, it would be like saying: "I believe *p*, therefore *p* is true," or more accurately, "I *must* believe *p*, therefore *p* is true." But is this what Whewell is saying? Nowhere does Whewell say that if a belief reaches the status of necessity, then this must be the *sole* source of its justification. In fact, Whewell explicitly denies this when he says (Butts, ed., 1989, p. 58, Whewell's emphasis) "*the terms which denote the fundamental antithesis of philosophy cannot be applied absolutely and exclusively in any case.*" That is, there are no scientific laws that are necessary in an absolute and exclusive sense. That is why Whewell's cannot believe that scientific laws are true because "I *must* believe *p*, therefore *p* is true."

Nevertheless, Whewell has an early paper (1834), published ten years before his fundamental antithesis of philosophy, which appears to provide some evidence for Butts and Mill's interpretation of Whewell. In that paper (reprinted in Butts, ed., 1989, pp.79-100), Whewell argues that Newton's laws are necessarily true and yet their contrary *is* conceivable. Whewell then observes that (*ibid.*, p.80) "[It is an apparent contradiction] that a law should be necessarily true and yet the contrary of it conceivable." Yet, clearly, his view is that it is *not* contradictory, because he then states that he will endeavor to explain it by pointing out the true grounds for the laws of motion. Doesn't this imply that the word "necessarily" in the phrase "necessarily true" is different from "necessary" in his psychological sense? Isn't this enough to show that Butts's slogan "I *must* believe *p*, therefore *p* is true" is a false characterization of Whewell's view?

---

[16] We think that Whewell is wrong here, but we also think his mistake does not undermine his epistemology.

[17] Harper (1989) and Metcalfe (1991) accuse Forster (1988) of ignoring Whewell's views about necessity, and his psychologism. This is a response to those criticisms.

Not quite, because Whewell goes on to argue that Newton's Laws have an experiential component (which accounts for the fact that we can conceive the contrary) and a necessary part, which he states in terms of three *axioms* that he thinks are necessarily true (in his psychological sense). If these axioms play a role in the justification for the extension of Newton's laws to untried instances, as Whewell maintains, then the necessity of the three axioms is indirectly playing a part in the justification of Newton's Laws. So, Whewell's explanation of the apparent contradiction may be that Newton's Laws are necessarily true in the oblique sense of being true and containing components whose contraries cannot be conceived.

For example (*ibid.*, p.81), "*Axiom I: Every change is produced by a cause*" expresses "a universal and constant conviction of the human mind" according to Whewell. Applied to motion (*ibid.*, p. 83), Axiom I asserts that (*ibid.*, p. 83) "when no force acts, the properties of motion will be constant." He then concludes that:

> …so far as the laws are announced in this form, they will be of absolute and universal truth, and independent of any particular experiment or observation whatsoever.
>   But though these laws of motion are necessarily and infallibly true, they are, in the form in which we have stated them, entirely useless and inapplicable. It is impossible to deduce from them definite and positive conclusions, without some additional knowledge or assumption. (*ibid.*, pp.83-84)

That is to say, the three necessarily true axioms do not by themselves yield Newton's three laws of motion without the addition of empirical claims. It becomes very clear that this is what Whewell is doing when he explicitly states (*ibid.*, p.97) the necessary and empirical parts of Newton's first law as "Velocity does not change without a cause" and "The time for which a body has already been in motion is not a cause of change of velocity," respectively. Why is "Velocity does not change without a cause" necessary for Whewell? Because velocity is a property of motion and Axiom I says that no properties of motion can change without a cause.

Our purpose is not to defend Whewell's 1834 analysis of Newton's laws. In fact, we agree with Butts, and others, that Whewell's analysis is incoherent. For example, consider Axiom I, in the form "Every change is produced by a cause." This is plainly consistent with the Aristotelian idea that a force is needed to produce any change of *position*. Yet the Aristotelian version of Axiom 1 is actually denied by Newton's first law, which allows that a body can move with a constant nonzero velocity *without* a cause. That is, changes in position may be *uncaused* on Newton's theory. But now it is apparent that the contrary of Axiom I is not only conceivable, but that the contrary of it is actually implied by Newton's first law of motion![18]

Our conclusion is that *even if* Whewell's analysis were successful, there still would be no grounds for concluding that Whewell's epistemology is *exhausted* by the slogan "I *must* believe *p*, therefore *p* is true."[19] Whewell never says that "The contrary of Newton's laws are inconceivable, therefore Newton's laws are true." He can only be read as implying that the contrary of Axiom I is inconceivable, therefore Axiom I is true. Yet at the same time Whewell is very clear that Axiom I has the same status as an arithmetical truth like "2+2=4". Thus, whatever one thinks of Whewell's views on necessity in this early paper, it is very clear that, in the final

---

[18] Note that Whewell's analysis is not contradicted by the fact that history has since shown that Newton's laws are false. Whewell would have the option of blaming the empirical part of those laws.

[19] For example, Buchdahl (1971, p. 345) says that "Whewell's necessitarianism was meant to provide an alternative to Mill's own theory of validation." However, if this implies that this is all that Whewell has to offer as an alternative, then the statement is very misleading.

analysis, they cannot provide anything close to a complete justification of *scientific* laws, for scientific laws always have an empirical component as required by his fundamental antithesis of philosophy. Therefore, one should concentrate on what Whewell says about testing hypotheses in his later writings in the *Novum Organon Renovatum* (see section 6).

9.   Logical versus Historical Theories of Confirmation

We have argued that Mill is wrong to claim that Whewell confuses invention and proof. Yet we certainly admit that Whewell is not always careful in distinguishing between the psychology of belief and the normative implications of theory testing. He often talks in psychological terms, about the conviction of scientists, or their inability to conceive contraries. Nevertheless, he also talks of truth, which clearly lies on the objective side of his fundamental antithesis. In most cases, we can infer his normative views from what he says about the psychological convictions of scientists. However, there are two cases in which this is not so straightforward. One case concerns the role of necessity in Whewell's theory, which we have just discussed. The second case concerns whether evidence is stronger if it is discovered after, rather than before, the invention of a hypothesis.

Mill's subsumption under general laws is a *logical* account of theory evaluation which gives no place to temporal considerations of this kind. How else can Whewell's account be different from Mill's theory if not by endorsing the relevance of historical considerations? It is natural to suspect that Whewell's is a historicist account of theory evaluation. We plan to argue that while Whewell agrees that scientists may be *psychologically* influenced by such considerations, and that the historical order of events may provide us with useful information, Whewell's underlying theory of hypothesis testing lies squarely on the logical side of the divide.

Mill thinks that Whewell's views are historicist, and he explicitly challenges him on these grounds.[20] Here is how Whewell recounts Mill's charge of historicism:

> These expressions of Mr. Mill have reference to a way in which hypotheses may be corroborated, in estimating the value of which, it appears that he and I differ. "It seems to be thought," he says (ii, 23), "that an hypothesis of the sort in question is entitled to a more favourable reception, if, besides accounting for the facts previously known, it has lead to the anticipation and prediction of others which experience afterwards verified." And he adds, "Such predictions and their fulfillment are indeed well calculated to strike the ignorant vulgar;" but it is strange, he says, that any considerable stress should be laid upon such a coincidence by scientific thinkers. (Butts, ed., 1989, pp. 293-4)

Unfortunately, Whewell addresses the psychological issue in his response (Butts, ed., 1989, pp.294):

> It was not the ignorant vulgar alone, who were struck by the return of Halley's comet, as an evidence of Newton's theory. Nor was it the ignorant vulgar, who were struck by those facts which did so much strike men of science, as curiously felicitous proofs of the undulatory theory of light…predicted by the theory and verified by experiment.

Mill is surely right that, in the case of Halley's comet, the mere return of the comet should not have impressed anyone. It doesn't take a rocket scientist to predict that a comet observed in 1531, 1607, and 1682, will return in 1759. Assuming that the period of motion is exactly

---

[20] See Musgrave (1974) for an excellent introduction to the distinction between logical and historicist accounts of confirmation. Other good discussions are found in Achinstein (1994), Mayo (1993, ch. 8) and Snyder (1994). Each of these authors credits Whewell with the historicist view.

constant, then simple calculations shows that it would have passed the perihelion (the closest point to the sun) in the middle of 1759. However, the point of the example is that Halley's comet did not return in the middle of 1759, but nearer the beginning of 1759. Moreover, that fact was predicted *in advance* by Clairaut's computations of the perturbations due to the attractions of Jupiter and Saturn (Moulton, 1970, p. 431). Whewell, the historian of science, was well aware of these details.

What if Clairaut has not completed his calculations until after the comet passed its perihelion on March 13, 1759? For those who were capable of understanding and checking Clairaut's calculations, the postdiction need not have been any less impressive. For those "ignorant vulgar", who were not capable of checking whether Clairaut's computations were fudged, the result would have been less convincing. That is, historical facts about the temporal order of a prediction and its verification, may be relevant to evaluating the strength of confirmation when the full logical facts are *not* known. However, this is perfectly consistent with a logical view of confirmation, which is only committed to the irrelevance of temporal considerations *if the full logical details of the calculation are fully taken into account*.

Our view is that Whewell's view of the psychological impact of *pre*diction (as oppose to postdiction) is consistent with a logical theory of confirmation.[21] For instance, a sentence just below the ones already quoted directly supports our viewpoint: "If we can predict new facts which we have not seen, *as well as explain those we have seen*, it must be because our explanation is not a mere formula of observed facts, but a truth of a deeper kind." (Butts, ed., 1989, p. 294, our emphasis.) A historicist would insist that if the prediction is already deduced, then only future facts can lead to confirmation, yet Whewell insists that the explanation of facts already seen also contributes to the confirmation.

Since we appear to be the only ones to have interpreted Whewell in this way, let us consider another supporting passage in which Whewell addresses the question explicitly. The passage introduces Whewell's second test of hypotheses:

> ...our hypotheses ought to *foretell* phenomena which have not yet been observed; at least all phenomena of the same kind as those which the hypothesis was invented to explain. For our assent to the hypothesis implies that it is held to be true of all particular instances. *That these cases belong to past or to future times, that they have or have not already occurred, makes no difference in the applicability of the rule to them. Because the rule prevails, it includes all cases; and will determine them all, if we can only calculate its real consequences. Hence it will predict the results of new combinations, as well as explain the appearances which have occurred in old ones*. (Butts, ed., p.151, our emphasis)

If one were to read only the first sentence of this quote, one might think of that temporal order is essential for Whewell's test. But the italicized sentences make it clear it makes no difference to the test whether the instances belong to the past, present, or future. The only difference temporal order makes is whether one describes the episode as an explanation or a prediction.

Whewell's second test concerns the "prediction of fact of the same kind." The consilience of inductions, on the other hand, is about "predictions of a different kind." Perhaps Whewell is a historicist when it comes to consilience? But here, Whewell also describes two different manifestations of the test: (a) when the second induction is new, it is described as a novel prediction, and (b) when the second induction is not new, it is described as a (common causal) explanation. These are different historical manifestations of the *same* test, and the strength of the

---

[21] Thus, statements scattered throughout the literature to the effect that on the logical theory historical facts cannot be relevant to confirmation (e.g., Snyder 1994) are misleading, and require qualification.

test is just as high in the second case, so long as the explanation is correct (that is, so long as the phenomenon is *correctly* deduced from the theory). For example, Whewell frequently cites Newton's explanation of the precession of the Earth's equinoxes as an impressive example of consilience. Here, the historicist viewpoint does not fit because the precession was a well established phenomenon prior to Newton Whewell is therefore implying that the temporal order is not crucial *from a normative standpoint.*[22]

10. Auxiliary Assumptions and the Correction of Errors

For Whewell, the general proposition that results from an induction provides a *theoretical* redescription of the observations "corrected by their general tendency," Curve fitting provides a dramatic and well understood example of what Whewell's means by this. "*The Method of Least Squares* is a Method of Means, in which the mean is taken according to the condition that the sum of the squares of the errours of observation shall be the least possible which the law of the facts allows," and "by this method, thus getting rid at once, in a great measure, of the errours of observation, we obtain data which are *more true than the* individual *facts themselves.*" (Butts, ed., 1989, p. 223) "If we thus take the *whole mass of the facts*, and remove the errours of actual observation, by making the curve which expresses the supposed observation regular and smooth, we have the separate facts corrected by their general tendency. We are put in possession, as we have said, of something more true than any fact by itself is." (Butts, ed., 1989, p. 227.)

Therefore, the process of induction proceeds from error-infected facts (where 'facts' for Whewell are concept-laden all the way down) to an inductive conclusion that succeeds in unifying and correcting those facts. How successful the initial colligation is successful in capturing the 'general tendency' depends on the case at hand. Even the cases judged to be the best examples of science in hindsight, like Kepler's inference for Mars, lend themselves to correction. How can Newton treat Kepler's laws as facts for the purpose of arguing for his universal theory of gravitation, and then go back and correct Kepler's laws without undermining his initial inference?

Mill is cut-off from the error-correcting step because he requires that Kepler's laws from a secure and incorrigible basis for Newton's induction. Mill's abides by the principle "Garbage in, garbage out." The problem traces back to Mill's deductivism. If Kepler's Law K is merely a *description* of the data E_K, as Mill's empiricism implies, then why does Newton think that it is false. Perhaps it is a false description? If so, why does the falsity of K not *dis*confirm Newton's theory? A more modern hypothetico-deductivist is likely to say that the falsity of K proves the falsity of the auxiliary assumptions that are used to derive K from Newton's theory. However, Newton's argument for universal gravitation does not fit in with this story either. First, Newton already knew that the auxiliary assumptions were false—they were *deliberate* idealizations. So, he had no need to use the deduction as a modus tollens argument because he already knew that there was at least one false premise. More importantly, Newton clearly used it as a modus

---

[22] There are related controversies concerning logical theories of confirmation, which we do not discuss. For example, does evidence used in the construction of a hypothesis count less strongly than evidence that is novel (see Mayo, 1996, ch. 6 for a recent discussion). Whewell makes a distinction here (between test 1 and test 2), but says nothing about the relative strengths of these tests. Of course, our view is that test 2 is the stronger test (see section 5).

ponens argument, for he presented the deduction as positive part of the argument for his theory.[23] How can the deduction of a falsehood from a falsehood ever *confirm* a theory for a deductivist?

On the other hand, Whewell explains that hypotheses may be "of service to science" when they are in error in the following way:

> A maxim which it may be useful to recollect is this,--that *hypotheses may often be of service to science, when they involve a certain portion of incompleteness, and even of errour*. The object of such inventions is to bind together facts which without them are loose and detached; and if they do this, they may lead the way to a perception of the true rule by which the phenomena are associated together, even if they themselves somewhat misstate the matter. The imagined arrangement enables us to contemplate, as a whole, a collection of special cases which perplex and overload our minds when they are considered in succession; and if our scheme has so much of truth in it as to conjoin what is really connected, we may afterwards duly correct or limit the mechanism of this connexion. (Butts, ed., p.149)

While Whewell is pointing to the *psychological* dangers of otherwise perplexing and overloading our minds, he is also claiming that the scheme may have "so much truth in it as to conjoin what is really connected." That is, the idealized assumptions may have the advantage of isolating the information that relevant to connecting the moon's motion to the motion of terrestrial projectiles. In particular, Kepler's and Galileo's laws succeed in capturing an aspect of the their respective phenomena that is due to the gravitational attraction of a single body. The motion of a projectile may deviate from Galileo's law because of air resistance, but this cause has no influence on the moon's motion. The moon's motion deviates from Kepler's laws because of its asymmetric shape, and again this has no observable effect on projectile motions. The consilience of inductions points *only* to the *common* causes. Once the common cause is correctly identified and measured, scientists may separate the effects of the common cause from the effects of other causes. In Newton's case, when the effect of the earth's gravity was correctly quantified and measured, the residual effects were more clearly defined, which allowed Newton to focus on their explanation. In this way, despite their incompleteness and error, Kepler's Laws played an essential role in Newton's argument for universal gravitation.

11. Summary and Conclusions

Mill and Whewell agree that the conceptualization of empirical data, or the colligation of facts, is a source of subjectivity in science. Concepts are added by the mind of scientists and therefore are potentially subject to personal prejudice and bias. Both thinkers wish to defend science as a source of objective knowledge. Mill's approach is to adopt an empiricist view of colligation, according to which concepts are determined by the empirical facts from the very beginning. In his view, the potential sources of prejudice and bias may never take root if the empirical methods of science are diligently enforced throughout the process. Whewell's viewpoint is importantly different: Conceptions are subjective elements that play a positive role in the progress of science

---

[23] Glymour's bootstrapping account of confirmation (Glymour 1980b) is a hypothetico-deductivist account of testing designed to confront the problem of circularity. His solution may be to say that Newton's Laws, L and initial auxiliary assumptions A_1, were such that the evidence E failed to entail an instance of K with the help of $L\&A_1$. Later, E did entailed an instance of K' with the help of $L\&A_2$, where K' is a corrected version of K, and A_2 is a more realistic set of auxiliary assumptions. According to Glymour's notion of confirmation, E disconfirmed K relative $L\&A_1$ and E confirmed K', relative to $L\&A_2$. We do not claim that Glymour's theory cannot answer our question, but we fail to see how K can confirms L when K is false. We suspect that it was not Glymour's intention to capture the kind of confirmation and correction claimed by Newton and condoned by Whewell.

towards objective truth, but only if they are tested by their generalization to facts of a kinds different from the ones they were designed to explain. It is not enough to endlessly test the predictions of a law of phenomenon against new instances. We need to look further afield at the *relationship* between these and other phenomena. This may be an unending process of successive generalization, but it is only through this process that conceptions are confirmed to be free of the bias and prejudices of their inventors.

It would be nice if Mill's picture of science were correct. But it is an empirical question whether science is really like that or not. Our examination of some of the key examples suggests that Mill's picture is wrong even in the so-called exact sciences. Kepler's ellipse was not determined by the empirical data. It was a conjecture in need of further verification, which Newton provided. Even if Mill were right and concepts are determined from the data, it does not follow that the *identity* of two concepts is determined from the data as well. Newton assumed without comment that the conception of mass used to colligate beam-balance phenomena is identical to the conception used to explain the behavior of springs. Yet these phenomena do provide independent measurements of the Newtonian mass of an object, and the agreement of these measurement confirms that the concepts represent some relational features of the evidence. If a Millian scientist had introduced beam-balance mass and spring-mass as separate concepts, then the disunified version of Newton's theory would have fitted the data just as well, and maybe better (Forster, 1986). The kind of cross-situational invariance (Forster 1984; Hooker 1987) exhibited by a unified theory such as Newton's, which is confirmed by the consilience of inductions, is strictly stronger evidence for a theory than the mere fact that a theory accommodates the total set of observations (section 5). Mill's subsumption under general laws is nothing more than a test of accommodation (section 5), while Whewell's consilience of inductions is strictly stronger.

Mill is also unable to explain how general laws are supported by the deduction of false or incomplete phenomenological laws (section 10). If Kepler's laws are false, and Newton's theory entails Kepler's laws, then Newton's theory fails a hypothetico-deductive test. So how can it be that this is part of Newton's argument for universal gravitation? Whewell's answer is that Kepler's Laws serve the purpose of representing the cause of the lunar motion that is also one of the causes of projectile motions, viz., the gravity of the earth. Though Kepler's Laws are false, they serve this purpose well. Once the common cause is verified by the consilience of inductions, it can be more accurately measured, and the residues, or errors, of the lower level laws can be explained by the introduction of new causes. These will be verified as common to other phenomena, thereby establishing a complex nexus of interrelationships amongst the phenomena.

Whewell does not conflate justification and discovery, or prescription and description, or proof and invention (as Mill alleges). His account of science is not only the better description, but it is also the better prescription if Whewell is right that concepts are initially underdetermined by the local evidence, but are less underdetermined by the more global, relational, facets of the total phenomena.

Whewell's account of scientific progress embraces the concept-ladenness of observation and reason throughout the process, while Mill tries to eliminate it from an essential role in the normative evaluation of science. But Mill has since lost the battle of minds to authors like Kuhn (1970, p.94), who use the concept-ladenness of observation to deny that there is any sense in which science evolves towards the truth: 'As in political revolutions, so in paradigm choice-there is no standard higher than the assent of the relevant community'. Yet, for Whewell, Kuhn's conclusion does not follow. Far from precluding the possibility of progress towards the truth, the

intrusion of subjective elements plays a positive role in the testing and correction of scientific hypotheses according to a standard that is often unnoticed and quickly forgotten by scientists, and neither receives nor requires their assent.

References

Achinstein, Peter (1994): '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.

Bell, John S. (1964). 'On the Einstein-Podolsky-Rosen Paradox', *Physics* **1:** 195-200.

Belot, Gordon (1998): 'Understanding Electromagnetism.' *British Journal for the Philosophy of Science* **49**: 531-555.

Blake, R. M., Ducasse, C. J., and Madden, E. H. (1960). *Theories of Scientific Method: The Renaissance through the Nineteenth Century*. University of Washington Press, Seattle.

Bogen, James & James Woodward (1988): 'Saving the Phenomena.' *The Philosophical Review,* vol. **XCVII:** 303-352.

Brown, James R. and J. Mittelstrass (eds.) (1989): *An Intimate Relation*. Dordrecht: Kluwer Academic Publishers.

Browne, Michael (2000): 'Cross-validation methods.' *Journal of Mathematical Psychology* **44**: 108-132.

Buchdahl, Gerd (1971): 'Inductivist *versus* Deductivist Approaches in the Philosophy of Science as Illustrated by Some Controversies Between Whewell and Mill,' *The Monist* **55**: 343-367.

Busemeyer, J. R. and Yi-Min Wang (2000): 'Model comparisons and model selections based on generalization test methodology,' *Journal of Mathematical Psychology* **44**: 177-189.

Butterfield, Herbert (1949) *The Origins of Modern Science*, 1300-1800 (London, 1949).

Butts, Robert E. (1973). 'Whewell's Logic of Induction,' in Giere & Westfall (eds.) *Foundations of Scientific Method: The Nineteenth Century*. Bloomington: Indiana University Press.

Butts, Robert E. (ed.) (1989). *William Whewell: Theory of Scientific Method.* Hackett Publishing Company, Indianapolis/Cambridge.

Cartwright, Nancy (1983): *How the Laws of Physics Lie.* Oxford: Oxford University Press.

Forster, Malcolm R. (1984): *Probabilistic Causality and the Foundations of Modern Science.* Ph.D. Thesis, University of Western Ontario.

Forster, Malcolm R. (1986): 'Unification and Scientific Realism Revisited.' In A. Fine and P. Machamer (eds.), *PSA 1986.* E. Lansing, Michigan: Philosophy of Science Association. **1:** 394-405.

Forster, Malcolm R. (1988): 'Unification, Explanation, and the Composition of Causes in Newtonian Mechanics.' *Studies in the History and Philosophy of Science* **19:** 55 - 101.

Forster, Malcolm R. (2000): 'Key Concepts in Model Selection: Performance and Generalizability,' *Journal of Mathematical Psychology* **44**: 205-231.

Forster, Malcolm R. and Elliott Sober (1994): 'How to Tell when Simpler, More Unified, or Less *Ad Hoc* Theories will Provide More Accurate Predictions.' *The British Journal for the Philosophy of Science* **45:** 1 - 35.

Friedman, Michael (1974). 'Explanation and Scientific Understanding.' *The Journal of Philosophy* **LXXI** 5-19.

Glymour, Clark (1980a). 'Explanations, Tests, Unity and Necessity.' *Noûs* **14:** 31-50.

Glymour, Clark (1980b). *Theory and Evidence.* Princeton University Press, Princeton.

Hanson, N. R. (1973). *Constellations and Conjectures,* W. C. Humphreys, Jr. (ed.) D. Reidel: Dordrecht-Holland.

Harman, Gilbert (1965). 'Inference to the Best Explanation.' *Philosophical Review* **LXXIV**, pp.88-95.

Harper, William L. (1989): 'Consilience and Natural Kind Reasoning' in J. R. Brown and J. Mittelstrass (eds.) *An Intimate Relation:* 115-152. Dordrecht: Kluwer Academic Publishers.

Hooker, Cliff A. (1987): *A Realistic Theory of Science*. Albany: State University of New York Press.

Jeong, Byung-Hoon (1991): 'William Whewell's Theory of Scientific Discovery,' *Hanguk Kwahak-Sa Hakhoe-Ji* **13**: 56-72.

Kitcher, Philip (1976): 'Explanation, Conjunction and Unification.' *Journal of Philosophy* **LXXIII***, pp. 207-212.

Kuhn, Thomas (1970): *The Structure of Scientific Revolutions,* Second Edition. Chicago: University of Chicago Press.

Mayo, Deborah G. (1996): *Error and the Growth of Experimental Knowledge*. Chicago and London, The University of Chicago Press.

Metcalfe, John F. (1991): 'Whewell's Developmental Psychologism: A Victorian Account of Scientific Progress.' *Studies in History and Philosophy and History of Science* **22:** 117 - 139.

Mill, John Stuart (1874), *A System of Logic, Ratiocinative and Inductive: Being a Connected View of the Principles of Evidence and the Methods of Scientific Investigation* (New York: Harper & Row).

Moulton, Forest Ray (1970): *An Introduction to Celestial Mechanics* (Second Revised Edition), New York, Dover Publications Inc.

Musgrave, Alan (1974): 'Logical Versus Historical Theories of Confirmation.' *The British Journal for the Philosophy of Science 25:* 1-23.

Popper, Karl (1959): *The Logic of Scientific Discovery.* London: Hutchinson.

Reichenbach, Hans (1956): *The Direction of Time*. Berkeley: University of California Press.

Ruse, Michael (1975): 'Darwin's Debt to Philosophy: An Examination of the Influence of the Philosophical Ideas of John W. Herschel and William Whewell on the Development of Charles Darwin's Theory of Evolution,' *Studies in the History and Philosophy of Science*, **6**: 159-181.

Ruse, Michael (1976): 'The scientific methodology of William Whewell,' *Centaurus*. Vol. 20, no.3, 227-257.

Snyder, Laura (1994): 'Is Evidence Historical?,' in Peter Achinstein and Laura Snyder (eds.) *Scientific Methods: Conceptual and Historical Problems*, Malabar, Fla.: Krieger Publishing Company.

Snyder, Laura (1997): 'The Mill-Whewell Debate: Much Ado about Induction,' *Perspectives on Science*. Vol. 5, no.2, 159-198.

Stone, M. (1977): 'An Asymptotic Equivalence of Choice of Model by Cross-Validation and Akaike's Criterion.' *Journal of the Royal Statistical Society B* **39**: 44-47.

van Fraassen, Bas (1980): *The Scientific Image.* Oxford: Oxford University Press.

van Fraassen, Bas (1982). 'The Charybdis of Realism: Epistemological Implications of Bell's Inequality.' *Synthese* **52**, pp.25-38.

Whewell, William (1834): 'On the Nature of Truth of the Laws of Motion,' *Transactions of the Cambridge Philosophical Society*, V, pp. 149 (read 17 Feb., 1834).

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

Whewell, William (1858): *The History of Scientific Ideas*, 2 vols. (London, John W. Parker).