Unwarranted assumptions: Claude Bernard and the growth of the *vera causa* standard

Raphael Scholl*

Draft of September 15, 2019

Contents

(Forthcoming in *Studies in History and Philosophy of Science*.)

*Department of History and Philosophy of Science, University of Cambridge,
e-mail: ras223@cam.ac.uk

Abstract

The physiologist Claude Bernard was an important nineteenth-century methodologist of the life sciences. Here I place his thought in the context of the history of the *vera causa* standard, arguably the dominant epistemology of science in the eighteenth and early nineteenth centuries. Its proponents held that in order for a cause to be legitimately invoked in a scientific explanation, it must be shown by direct evidence to exist and to be competent to produce the effects ascribed to it. Historians of scientific method have argued that in the course of the nineteenth century the *vera causa* standard was superseded by a more powerful consequentialist epistemology, which also admitted *indirect* evidence for the existence and competence of causes. The prime example of this is the luminiferous ether, which was widely accepted, in the absence of direct evidence, because it entailed verified observational consequences and, in particular, successful novel predictions. According to the received view, the *vera causa* standard's demand for direct evidence of existence and competence came to be seen as an impracticable and needless restriction on the scope of legitimate inquiry into the fine structure of nature. The Mill-Whewell debate has been taken to exemplify this shift in scientific epistemology, with Whewell's consequentialism prevailing over Mill's defense of the older standard. However, Bernard's reflections on biological practice challenge the received view. His methodology marked a significant extension of the *vera causa* standard that made it both powerful and practicable. In particular, Bernard emphasized the importance of detection procedures in establishing the existence of unobservable entities. Moreover, his sophisticated notion of controlled experimentation permitted inferences about competence even in complex biological systems. In the life sciences, the *vera causa* standard began to flourish precisely around the time of its alleged abandonment.

Mais la science n'avancerait jamais si l'on se croyait autorisé à renoncer aux méthodes scientifiques parce qu'elles sont imparfaites; la seule chose à faire en ce cas, c'est de les perfectionner.

— Claude Bernard, *Introduction à l'étude de la médecine expérimentale* (1865)

1 Introduction

The nineteenth century witnessed a clash between two different conceptions of legitimate scientific method. One side held that entities may only be invoked in scientific explanations if their real existence can be supported directly, in the paradigmatic case by observation. Such entities were "true causes" or *verae causae*, following Newton's terminology in the *Principia*. The other side thought it legitimate to invoke hypothetical entities in scientific explanations, so long as their existence could be supported indirectly by their observable consequences. I will refer to the first group as adherents of the *vera causa* standard, and to the second as proponents of a consequentialist theory of evidence.

On a received view of the history of scientific method, consequentialism started out as a minority view but eventually prevailed.[1] This shift was exemplified by the debate between John Stuart Mill and William Whewell, in which a central point of contention was the acceptability of the luminiferous ether as a physical substrate for the wave theory of light. Mill argued that the ether did not meet the evidential demands of the *vera causa* standard. Neither its existence nor its competence to produce the effects ascribed to it had been demonstrated by direct evidence. That the ether hypothesis explained the phenomena of light was no reason to accept it, since any number of unconceived alternative hypotheses might furnish such explanations as well. Whewell, by contrast, argued that the luminiferous ether did not just explain a wide range of phenomena, but had also predicted phenom-

[1] See the essays collected in Larry Laudan's *Science and Hypothesis* (1981), especially chapter 7, "Thomas Reid and the Newtonian Turn of British Methodological Thought", and chapter 8, "The Epistemology of Light: Some Methodological Issues in the Subtle Fluids Debate". See also Laura Snyder (1997b), and the discussion in Section 2.3.

ena different from the ones the hypothesis was originally conceived to explain. He thought that this provided genuine, albeit indirect, evidence for the ether's existence. On the received view, the demands of the *vera causa* standard came to be seen as limiting, and something like Whewell's consequentialism came to be widely accepted as the appropriate tool for probing the unobservable fine structure of nature. This historical account was articulated in the second half of the twentieth century and cohered well with the accounts of scientific epistemology of the time, which were predominantly consequentialist. They included Hypothetico-Deductivism, Falsificationism, and Inference to the Best Explanation.

The thesis of this paper is that the *vera causa* standard not only survived but even flourished in the second half of the nineteenth century. If we turn our attention to the life sciences, the notion of a decisive shift to consequentialism collapses. I will argue that far from being soundly rejected as an epistemic guidepost, the *vera causa* standard was turned from an unachievable ideal into a practicable methodology by researchers in biology and medicine. A testament to this fact are the writings of one of the leading nineteenth-century methodologists of the life sciences: the French physician and physiologist Claude Bernard (1813-1878). In his seminal and influential *Introduction to the Study of Experimental Medicine* (1865), Bernard articulated a sophisticated and powerful version of the *vera causa* approach.

The plan of the paper is as follows. Section 2 revisits the origins, development, and demise of the *vera causa* standard. Section 3 shows that Bernard's methodological and scientific writings advanced the *vera causa* standard towards a practicable method of inquiry. Unlike earlier articulations of the standard, Bernard's version was expansive in its ability to accommodate unobservable causes, and powerful in its provisions for controlling known and unknown confounders in experiments. Section 4 concludes the discussion.

2 The rise and fall of the *vera causa* standard

2.1 Origins

The term *vera causa* traces back to Newton's *Principia*.[2] At the beginning of the third book, Newton presented methodological principles that were labeled as "regulæ philosophandi", or rules of philosophizing, from the second edition onward.[3] The first rule reads:

> Causas rerum naturalium non plures admitti debere, quam quæ & veræ sint & earum Phænomenis explicandis sufficiant. (Newton 1713, 357)

We are to "admit no more causes of natural things, than such as are both true and sufficient to explain their appearances" (Newton 1729, 202).

To some, such as the Scottish Enlightenment philosopher Thomas Reid (1710–1796), Newton's first rule was "the true and proper test, by which what is sound and solid in philosophy may be distinguished from what is hollow and vain" (Reid 1863, 236). But the rule required considerable unpacking before it could be seen as a practicable methodological guide. Its demands can appear vacuous, since it is little help to be told that one should search for true rather than false causes, or sufficient rather than insufficient ones. Indeed, the editor of Reid's collected works considered the rule a "barren truism" (Reid 1863, 236, footnote).

As Reid interpreted the first rule, however, it made two substantial demands.[4] First, the appropriate contrast for a true cause was not a false cause, but a *speculative* cause, a cause that *might* explain the phenomena but whose existence is uncertain. Reid thought that explaining by speculative causes was too easy:

[2]An important resource on the history and philosophy of the *vera causa* standard is the unpublished Ph.D. dissertation by Kavaloski (1974). Hodge (1977, 1992) has written extensively about the standard as it relates to Darwin's *Origin of Species*. More recently, *veræ causæ* have enjoyed renewed interest from Novick (2016), Pence (2018), and Novick and Scholl (2018).

[3]For the editorial history of the *regulæ*, see Newton (1999, 794) and Ducheyne (2015). On recent debates concerning the status of the *regulæ*, see Biener (2018) and Di Fate (2011).

[4]My historical interpretation of Reid's views on *veræ causæ* is mainly indebted to Laudan (1981a), but also to Wood (1989) and Callergård (1999). For more on Reid's philosophy of science see also Ducheyne (2006) and Callergård (2006), and for an overview of Reid's works see Cuneo and Van Woudenberg (2004).

> If the hypothesis hangs well together, is embellished by a lively imagination, and serves to account for common appearances, it is considered by many as having all the qualities that should recommend it to our belief, and all that ought to be required in a philosophical system. (235)

But such reasoning had led some to think that the earth is supported by a huge elephant standing on a huge tortoise, or that the planets are moved by Cartesian vortices (234–235). To guard against speculative explanations, Reid interpreted the demand for a true cause as the demand for evidence that the cause really exists:

> If a philosopher … pretends to shew us the cause of any natural effect, whether relating to matter or to mind, let us first consider whether there is sufficient evidence that the cause he assigns does really exist. If there is not, reject it with disdain, as a fiction which ought to have no place in genuine philosophy. (236)

Thus, demonstrating the existence of a cause required something in addition to mere explanatory power. This is the *existence requirement* of the *vera causa* standard.

On Reid's interpretation, the second methodological demand of the first rule was that once a cause was known to exist, its ability to produce the effects ascribed to it needed to be supported by evidence. He wrote:

> If the cause assigned really exists, consider, in the next place, whether the effect it is brought to explain necessarily follows from it. (236)

To demonstrate that a cause is "sufficient to produce the effect" (250) is the *competence requirement*.

It is helpful to illustrate Reid's views by a concrete example. Reid extensively criticized David Hartley's (1749) *Observations on Man*, which proposed a "doctrine of vibrations" to explain the operation of the senses and the mind.[5] Hartley's fourth Proposition expressed one of the doctrine's fundamental assumptions:

[5]For an overview of Hartley's work, see Allen (2017).

External Objects impressed upon the Senses occasion, first in the Nerves on which they are impressed, and then in the Brain, Vibrations of the small, and, as one may say, infinitesimal, medullary Particles. (Hartley 1749, 11)

According to Reid, however, the existence such vibrations was unproven. Hartley had argued that “no motion, besides a vibratory one, can reside in any part for a moment of time” (Reid 1863, 250). But Reid could refute this by naming other continuing motions, “such as rotation, bending or unbending of a spring, and perhaps others which we are unacquainted with” (250). The possibility of unconceived alternatives blocked Hartley’s consequentialist inference to the existence of vibrations.

Since the existence of the postulated vibrations was uncertain, the question of whether they were competent to produce sensations and ideas was not even amenable to inquiry. Reid asked:

[H]ow can we expect any proof of the connection between vibrations and thought, when the existence of such vibrations was never proved? … The existence of both must be known before we can know their connection. (252)

To understand this part of Reid’s critique, it is instructive to contrast the doctrine of vibrations with a similar explanation that Reid regarded as successful: the explanation of sound in terms of vibrations of the air. According to Reid, such vibrations were known to be competent to produce the effects ascribed to them:

[W]e know that, as the vibration is strong or weak, the sound is loud or low; we know that, as the vibration is quick or slow, the sound is acute or grave. … [A]nd all this is not conjectured, but proved by a sufficient induction. (253)

Reid did not specify in what sense the competence requirement in this case was fulfilled by a “sufficient induction”. He only wrote that the vibrations of air “tally exactly” (253) with the phenomena of sound. We will see below how later authors expanded upon this notion.

In summary, Reid interpreted Newton’s first rule in terms of two substantial demands. First, legitimate scientific explanations must invoke causes whose existence can be supported by good evidence. Crucially, explanatory success alone does *not* count as good evidence. Second, legitimate explanations must invoke causes that are demonstrably competent to produce the kinds of effects ascribed to them. Reid insisted that unless a causal explanation meets both conditions, it is "good for nothing" (236, 250). In this context, Reid cited with approval Newton's famous dictum "hypotheses non fingo" from the General Scholium of the *Principia* (236). Having shown the effects of the force of gravity in the planetary system, Newton could only have speculated about the underlying cause of gravity itself, and he refused to do so.[6]

2.2 Development

In the early nineteenth century, further developments concerned particularly the competence requirement. The contributions of two philosophers stand out: John Herschel and John Stuart Mill.

John Herschel

In the *Preliminary Discourse on the Study of Natural Philosophy*, Herschel (1830) understood the goal of natural philosophy to be the explanation of phenomena in terms of an "immediate producing cause" (144, §137). As science progresses, our knowledge of such causes accumulates:

> Experience having shown us the manner in which one phenomenon depends on another in a great variety of cases, we find ourselves provided, as science extends, with a continually increasing stock of such antecedent phenomena, or causes (meaning at present merely proximate causes) competent, under different modifications, to the production of a great multitude of effects, besides those which originally led to a knowledge of them. To such causes Newton has applied the term *veræ causæ*; that is, causes recognized as having a real existence in nature, and not being mere hypotheses or figments of the mind. (144, §138)

We see here the familiar components of the *vera causa* standard. *Veræ causæ* have "a real existence" and are not mere "hypotheses". They must be "com-

[6] On the meaning and the English translation of Newton's dictum, see Cohen (1962). For a recent analysis of the methodology of the *Principia*, see Smith (2002).

petent" to produce their effects.[7] Crucially, they provide explanations for phenomena that are different from the ones "which originally led to a knowledge of them".

Herschel illustrated the force of these distinctions by an example from actual science:

> The phenomenon of shells found in rocks, at a great height above the sea, has been attributed to several causes. By some it has been ascribed to a plastic virtue in the soil; by some, to fermentation; by some, to the influence of the celestial bodies; by some, to the casual passage of pilgrims with their scallops; by some, to birds feeding on shell-fish; and by all modern geologists, with one consent, to the life and death of real mollusca at the bottom of the sea, and a subsequent alteration of the relative level of the land and sea. (144–45, §138)

He went through these possibilities in quick succession. The plastic virtues and celestial influences did not meet the existence requirement since they were "figments of fancy" (145, §138). Fermentation as a cause of shells in rocks failed the competence requirement, "since no such thing was ever witnessed as one of its effects, and rocks and stones do not ferment" (145, §138). Casual transport met the existence and competence requirements but was "not extensive enough" to be responsible for the phenomenon (145, §138). Only the transport of actual organic remains by known geological processes met the criteria of existence, competence, and sufficiency in magnitude.

One of Herschel's key contributions to the *vera causa* standard was his articulation of more precise methods for determining competence. He began by outlining some of the characteristics of "that relation which we intend by cause and effect" (151, §145). According to Herschel, there is an "invariable connection" between cause and effect, and also an "invariable negation of the effect with absence of the cause", unless the effect can be produced by multiple causes (151, §145). Moreover, if it is possible to produce an "increased or diminished intensity of the cause", then there will also be an "[i]ncrease or diminution of the effect" (151–52, §145).

[7] I agree with Pence (2018) that a *vera causa* in Herschel's writings can be any entity known to exist, but I take the methodological *standard* to require evidence of both existence and competence.

Herschel enumerated a total of ten epistemic strategies for determining causal relationships, of which I will discuss a selection. In a nod to Newton, he described these strategies as "rules of philosophizing" (152, §146).[8] A first rule is that if there is an "attendant circumstance" which is absent from some instances of a type of effect, then that circumstance "cannot be the cause we seek" (152, §146).[9] A second rule states that if all instances of an effect agree in a circumstance, then this circumstance *may* be the cause we seek, or else a "collateral effect" of it (152, §146).

The fifth rule connects directly with Reid's discussion of vibratory theories. Herschel returned to the example of the connection between vibrations of the air and the pitch of sounds. He argued that we can impart impulses of equal force but increasing frequency to the air, and that this will be perceived first as "a rattling noise, then a low murmur, and then a hum, which by degrees acquires the character of a musical note" (153, §153). This is what allows us to infer the competence of the cause to produce the effect:

> [F]rom this correspondence between the pitch of the note and the rapidity of succession of the impulse, we conclude that our sensation of the different pitches of musical notes originates in the different rapidities with which their impulses are communicated to our ears. (153–54, §153)

Remember Reid's phrasing that the cause and effect in this case "tally exactly". Herschel's fifth rule explicates this notion in terms of a correspondence between the intensities of the cause and of its effect, which, as we saw above, he took to be characteristic of causal relationships.

Herschel considered his seventh rule to be among the most conclusive. It articulates the notion of varying one thing at a time while keeping all else equal:

> If we can either find produced by nature, or produce designedly for ourselves, two instances which agree *exactly* in all but one particular, and differ in that one, its influence in producing the phenomenon, if it have any, *must* thereby be rendered sensible.[10] (154, §156, original emphasis)

[8] The full complement of rules is discussed by Herschel (1830) in paragraphs §146–162; see also Ducasse (1960) and Cobb (2012).

[9] Assuming, again, that the effect is not produced by multiple causes.

[10] The rule as stated is unsound. If the two instances are exactly alike but an alternative

Herschel recognized that the difficulty of the rule lies in the conditional. It licenses an inference only *if* two instances agree in all but one particular, and such instances are difficult to find in nature. The conclusiveness of an experiment depends precisely on the comparability of its instances:

> [Two instances] become more valuable, and their results clearer, in proportion as they possess this quality (of agreeing exactly in all their circumstances but one), since the question put to nature becomes thereby more pointed, and its answer more decisive. (155, §156)

Having recognized the importance of the question, Herschel had little more to say about how two comparable instances are found, except that they were "easy to produce" with the aid of experiments (155, §156).

Finally, Herschel's ninth rule states that if a phenomenon is produced by several interacting causes ("concurring, opposing, or quite independent of each other", 156, §158), we can subtract the effects of all known causes so as to leave a "residual phenomenon" to be explained. Herschel considered this rule to be the one by which "science, in its present advanced state, is chiefly promoted" (156, §158).

John Stuart Mill

Next to Herschel, John Stuart Mill was the most prominent nineteenth-century proponent of the *vera causa* standard. His treatment of causation and causal inference followed Herschel's structure. Mill began book III of the *System of Logic* with chapters focusing on the nature of causation (Mill 1974, especially III.IV–III.VI). While Mill's causal metaphysics is outside the scope of this paper, it is important to see that Mill's discussion refined Herschel's. Mill discussed the notion of invariable antecedents in much greater detail, with particular attention to the complexity of most causal processes. While individual antecedent conditions were usually described as the cause of a phenomenon, Mill insisted that the cause of a phenomenon, properly speaking, is an "assemblage" of antecedent conditions. He also discussed the "composition of causes", distinguishing different ways in which causes can interact.

cause of the effect acts in *both*, then this may mask the action of the cause under investigation. Mill's version of this rule (see below) avoids this difficulty.

Like Herschel before him, Mill understood these preliminary considerations to be a foundation for a set of epistemic strategies: Mill's four methods of experimental inquiry (III.VIII). In a letter to Herschel, Mill wrote that the four methods constituted "the most important chapter of the book", but were also "little more than an expansion & a more scientific statement of what you had previously stated".[11]

Thus, Herschel's first and second rules became the *method of agreement*; the fifth rule became the *method of concomitant variation*; and the ninth rule became the *method of residues*. Most importantly, Herschel's seventh rule became the *method of difference*. Mill framed the method as a kind of inference schema:

> If our object be to discover the effects of an agent A, we must procure A in some set of ascertained circumstances, as A B C, and having noted the effects produced, compare them with the effect of the remaining circumstances B C, when A is absent. If the effect of A B C is *a b c*, and the effect of B C, *b c*, it is evident that the effect of A is *a*. (III.VIII.§2, 391)

The method of difference compared instances of occurrence with instances of non-occurrence, to see in what they differed.[12]

Like Herschel, Mill thought that the method of difference was particularly decisive. The other methods could reveal candidate causes, but it was only by the method of difference that "we can ever, in the way of direct experience, arrive with certainty at causes" (394, III.VIII.§3).[13] Mill agreed with Herschel that natural instances meeting the requirements of the method were rare, but that in experiments "a pair of instances such as the method requires is obtained almost as a matter of course" (393, III.VIII.§3). He argued that intervention ensures the comparability of instances:

> We choose a previous state of things with which we are well acquainted, *so that no unforeseen alteration in that state is likely to*

[11]The letter is dated May 1, 1843, and cited in Mill (1974, lxviii). Herschel himself traced the rules to Bacon's "prerogatives of instances" (§190–200).

[12]Mill's articulation of the difference-making idea does not suffer from the problem discussed in fn. 10. If an alternative cause were masking the effect of the cause under investigation, then we would see no difference between the instances and no inference would be warranted.

[13]But see Ducheyne (2008) on Mill's changing views about whether the method of difference affords certainty.

> *pass unobserved*; and into this we introduce, *as rapidly as possible*, the phenomenon which we wish to study; so that in general we are entitled to feel complete assurance that the pre-existing state, and the state which we have produced, differ in nothing except the presence or absence of that phenomenon. (393, III.VIII.§3, my emphases)

These notions are sensible but demanding. If no relevant alterations in the initial state are to "pass unobserved", then experimental inferences are only possible if all relevant confounders are antecedently known. Similarly, if the goal is to act swiftly "before there has been time for any change in the other elements" (393, III.VIII.§3), this requires that we know about the time scale at which all relevant confounders act. We will see below that this demanding notion of comparability led Mill to put significant limits on the scope of the method of difference.

To conclude, Herschel and Mill offered successive refinements of strategies for assessing whether causes that are known to exist are also competent to produce the phenomena ascribed to them. Both assigned great importance to the study of differences between comparable instances, but neither had a satisfying account of how comparable instances can be found or produced.

2.3 Demise

Historians of scientific method have argued that the *vera causa* standard came to be seen as an impediment to the progress of theoretical science in the course of the nineteenth century (see Laudan 1981a, 1981b, and references below). At issue were in particular explanations of phenomena in terms of the behavior of subtle fluids. Today, the best known theory of this kind is the explanation of light in terms of vibrations in a luminiferous ether. But at the time, similar theories were proffered for the explanation of heat, magnetism, electricity, and other processes.

We encountered an eighteenth-century instance of such theorizing in Section 2.1, in the form of Hartley's doctrine of vibrations. Reid criticized the hypothesis for failing to meet the existence and competence requirements. However, Hartley did not accept Reid's epistemological standards. He wrote that even if we suppose a postulated entity to be "destitute of all direct Evidence, still, if it serves to explain and account for a great Variety of Phenomena, it will have an indirect Evidence in its favour by this means" (Hartley 1749, 15). Thus, Hartley proposed a consequentialist theory of evidence, one that allowed the existence and properties of entities to be supported indirectly by their verifiable observable consequences. In effect, he tried to turn what had always been a necessary condition for scientific theorizing, to save the phenomena, into a sufficient condition.

Laudan agreed with Reid and other eighteenth-century methodologists that simple versions of consequentialism, such as Hartley's, were "unconvincing and inadequate" (1981, 111). But he also argued that Reid's own *vera causa* standard was far too limited:

> As Reid construes the first *regula*, it amounts to the claim that any putative causal explanation (a) must be sufficient to explain the relevant appearances and (b) must postulate entities and mechanisms whose existence can be *directly* ascertained. Condition (b) is the crucial one because it is meant to explicate the demand for 'true causes'. What this amounts to is the claim that *unobservable entities*, because we can have no *direct* evidence of their existence, *have no role to play in causal explanations*. (Laudan 1981a, 93, original emphases)

On this interpretation, scientific theorizing about unobservables entities was beginning to chafe against a restrictive methodological doctrine. However, the consequentialist alternative was still far too permissive to be a plausible replacement.

According to Laudan, more sophisticated versions of consequentialism were developed, and came to be accepted, as scientific theorizing about unobservables progressed during the nineteenth century. The famous debate between Mill and William Whewell has been seen as an expression of this shift. On one side, Mill defended the *vera causa* standard. On the other, Whewell defended a new consequentialism. One of the main points of controversy was the acceptability of the luminiferous ether as a physical substrate for the wave theory of light.[14]

[14] The most extensive recent study of the Mill-Whewell debate is by Snyder (2006, 1997b). Snyder contextualizes the debate about scientific epistemology as only one aspect of a much larger disagreement. For earlier views on the debate, see in particular Strong (1955), Buchdahl (1971), and Fisch (1985).

Whewell saw traditional interpretations of Newton's first rule as "an injurious limitation of the field of induction" (Whewell 1860, 186).[15] He proposed a different account of induction, which has been labeled "antithetical". It involved both an empirical, objective element and an ideal, subjective element.[16] Whewell believed that in inducing theories from phenomena, the mind has to furnish "conceptions" that appropriately "tie together" or "colligate" these phenomena. Such conceptions could be of various kinds. In the study of the shape of Mars's trajectory, Kepler's mind supplied the conception of an ellipse to colligate the observed positions of the planet. Newton later colligated the planetary trajectories by a new conception: that of a force acting according to an inverse square law to produce an elliptical path. Similarly, the conception of waves in a luminiferous ether colligated the phenomena of light. Once an appropriate conception to colligate phenomena was introduced, it was generalized and then tested by its empirical consequences. Whewell's account was thus consequentialist, even though, as we will see below, there are significant differences between it and its more recent cognates.

Mill disputed the reliability of consequentialist theory testing. He did not reject hypotheses as such, since in the initial stages of inquiry hypotheses about possible causes of a phenomenon were "allowable, useful, and often even necessary" (Mill 1843, 21, III.XIV.§6). However, referring directly to Newton's first rule, he insisted that eventually "[the cause's] existence should be capable of being detected, and its connexion with the effect ascribed to it should be susceptible of being proved, by independent evidence" (496, III.XIV.§5). The luminiferous ether, although widely accepted, did not meet these requirements:

> The possibility of deducing from its supposed laws a considerable number of the phenomena of light, is the sole evidence of its existence that we have ever to hope for; and this evidence cannot be of the smallest value, because we cannot have, in the case of such an hypothesis, the assurance that if the hypothesis

[15] Whewell discussed Newton's rules at length in his *Philosophy of Discovery* (1860, Ch. 18). See also the exchange between Robert Butts (1970; 1973) and David Wilson (1973) on Whewell's attitude to true causes.

[16] Whewell (1840; 1847, especially book XI). On Whewell's epistemology, see Snyder (2006, Ch. 1), and see Yeo (1993) for additional historical context. The term "antithetical" derives from Fisch (1985).

> be false, it must lead to results at variance with the true facts. (22, III.XIV.§6)

Like Reid before him, Mill thought it insufficient to save the phenomena, because unconceived alternative hypotheses might save them just as well.

Unlike earlier consequentialists, however, Whewell did not recommend the acceptance of hypotheses merely because they accounted for a wide range of phenomena. He articulated three further "Tests of Hypotheses" (Whewell 1847, XI.V.III, Art. 10–13). First, an hypothesis needed to "foretel" or predict phenomena. Second, it needed to explain phenomena of a different *kind* from the ones which originally led to the hypothesis. Whewell called this the "consilience" (or "leaping together") of inductions. Third and finally, hypotheses needed to cohere over time and to tend towards simplicity. Whewell argued that the luminiferous ether met these criteria, since experiments had confirmed the optical phenomena predicted by Young and Fresnel. Moreover, the theory's further explanatory successes, such as polarization and double refraction, were instances of consilience.

It would go beyond the scope of this paper to trace the further thrusts and parries of the debate in detail. Whewell argued that consilience gave a theory "a stamp of truth beyond the power of ingenuity to counterfeit" (Whewell 1849, 61). Mill countered that even false theories, if they are good enough to account for some phenomena, should naturally be expected to account for some more. Famously, he wrote that only the "ignorant vulgar" were impressed by successful predictions.[17] It is safe to say that the debate found no resolution at the time.

Modern commentators, however, have generally favored Whewell's position on the existence criterion.[18] We have seen that Laudan held the *vera causa* standard to be limited to observables, and he argued that consequentialist approaches such as Whewell's were necessary for inquiries into unobservables. In her more recent study of the Mill-Whewell debate, Snyder (1997b, 2006) arrived at similar conclusions. To be sure, she corrected earlier authors (especially Laudan) who had assimilated Whewell too closely to twentieth-century consequentialist theories of evidence. While modern accounts allow for the free invention of hypotheses, Whewell's epistemology

[17] In later editions, the ignorant vulgar became the merely "uninformed" (Mill 1974, 500, III.XIV.§6).

[18] But see Peter Achinstein's (1992) review of Lipton's *Inference to the Best Explanation*.

demanded an inductive path from phenomena to theories (Snyder 1997a, 2006, 2009). Only theories generated by appropriate inductions could be supported by consequentialist evidence, contra Laudan (1981c).[19] Nevertheless, Snyder agreed with Laudan that "Mill's inductive methodology, unlike Whewell's, does not allow inferences to *any* hypotheses about unobservable properties or entities" (188 Snyder 1997b, original emphasis). Thus, she concluded that Whewell's views conformed more closely than Mill's "to the practice of scientists such as Kepler, Newton, and Fresnel, who do attempt to discover laws involving unobservables" (195).

In addition to criticizing the existence criterion, Whewell also criticized Mill's competence criteria. He granted that the four methods of experimental inquiry would allow inferences if they could be applied, but he objected that the difficulty was "the reduction of the phenomena to formulæ such as are here presented to us" (Whewell 1849, 44, §42). How were we to find the required combinations of conditions and phenomena? Whewell thought that the methods "take for granted, the very thing which is most difficult to discover" (44, §42). Mill had failed to show that the four methods had actually played a role in significant discoveries in the history of science:

> Who will carry these formulæ through the history of the sciences, as they have really grown up; and show us that they these four methods have been operative in their formation; or that any light is thrown upon the steps of their progress by reference to these formulæ? (Whewell 1849, 45, §42)

Mill's answer to this historical challenge spanned a mere paragraph, and his preference for toy examples (such as "dogs bark" to illustrate the method of agreement or "fire burns" to illustrate the method of difference) did little to reveal the methods' actual usefulness in scientific practice (Mill 1974, 431–2, III.IX.§6).[20]

Surprisingly, perhaps, Mill himself took the scope of the method of difference, in particular, to be severely limited. We have already seen that he had a very demanding notion of when two instances were sufficiently comparable to allow an inference. In keeping with this view, he judged the method to be "entirely unavailing" in the complex systems studied by

[19] Cobb (2012) made a similar argument for the case of Herschel.

[20] However, Cobb (2011) argued that Mill's historical examples do not entirely deserve their bad reputation.

physiologists or social scientists (Mill 1974, 451, III.X.§8). As an example, he discussed the question of whether mercury can cure a particular disease. We might try to find an answer by comparing a patient both before and after the administration of mercury. But Mill thought this could not work:

> The mercury of our experiment being tried with an unknown multitude (or even let it be a known multitude) of other influencing circumstances, the mere fact of their being influencing circumstances implies that they disguise the effect of the mercury, and preclude us from knowing whether it has any effect or not. (Mill 1974, 450, III.X.§8)

On Mill's account, the method of difference demanded not only that we exclude all unknown antecedent conditions, but also that we suppress all known causes of the phenomenon, or at least that we "make them such that we can compute and allow for their influence" (450). This made inferences impracticable:

> Unless we already knew what and how much is owing to every other circumstance, (that is, unless we suppose the very problem solved which we are considering the means of solving,) we cannot tell that those other circumstances may not have produced the whole of the effect, independently or even in spite of the mercury. (Mill 1974, 450, III.X.§8)

Since direct inferences to causes in complex systems were impossible, Mill thought that only basic regularities were to be determined experimentally. Most of science would proceed deductively from these basic regularities (Mill 1974, 481–83, III.XIII.§7).

Recent commentators have generally agreed that Mill's methods can demonstrate competence only within narrow limits. We saw above that Snyder took Mill's methods to be restricted to observable causes. Similarly, Peter Lipton (2004, 126–28) argued that Mill's method of difference was only a first-pass description of many scientific inferences. Like Synder, he thought it was limited to observables. Moreover, he argued that the method gave us no satisfying guidance on how to judge the comparability of two instances – on how to distinguish between relevant differences and abundant irrelevant ones. He suggested that his account of inference to the best explanation could repair these defects. Even though we had no direct access to whether a sole unobservable difference existed between instances, we could make the subjunctive judgment that *if* such a difference existed, then it *would* explain a difference in experimental outcomes. In effect, he held that Mill's methods needed to be subsumed by a consequentialist framework (for discussion, see Scholl 2015).

3 Claude Bernard on the epistemology of experimental medicine

We saw that the *vera causa* standard is supposed to have failed as a viable epistemology of science around the middle of the nineteenth century, and that this failure has been explained as a result of the standard's epistemic limitations. The existence requirement has been interpreted as limiting scientific inquiry to observables, and the competence requirement as being too demanding for most practical purposes. In this section I will argue that far from disappearing as a regulative ideal, the *vera causa* standard was at the core of the developing epistemology of the nineteenth-century life sciences.

I will refer to the work of the French physiologist Claude Bernard (1813–1878) as a guide to the leading edge of nineteenth-century methodological thinking in experimental biology and medicine. When his *Introduction to the Study of Experimental Medicine* was published in 1865, Bernard was fifty-two years old and a world-renowned physiologist (Olmsted and Harris Olmsted 1952; Grmek 2008). He had contributed a series of significant studies, particularly on digestion and hepatic glycogenesis (Holmes 1974) and on the mode of action of various poisons, including carbon monoxide and curare (Grmek 1973). With the *Introduction*, Bernard established himself, in addition, as a methodologist and philosopher of science (Gayon 2015).

Bernard's *Introduction* never mentioned the term *vera causa* or Newton's first rule. It may thus seem contrived to situate it within the *vera causa* tradition. However, we will see below that Bernard's methodological thought aligns very closely with the tradition. One way to explain this correspondence would be to argue that Mill and others correctly described aspects of scientific practice, and that Bernard, although perhaps unfamiliar with the methodological writers, was well trained in that practice. But there is also abundant evidence of direct influence. Even though Mill is not mentioned in the *Introduction*, Bernard was familiar with his writings. In an undated note published posthumously, Bernard explained that he concerned himself with scientific invention, something that, in his view, had been neglected "by philosophers and even by Mill".[21] This suggests that Bernard held Mill in particularly high esteem. Bernard's exchange of letters with Marie Raffalovich also testifies to Bernard's knowledge of Mill. In the years 1873 and 1874, he inquired about Raffalovich's views on Mill and mentioned him as a corrective to other authors, and Raffalovich even considered translating a recent biography of Mill into French (Bernard 1950, 174, 176). None of these references to Mill can be dated to before the appearance of the *Introduction* in 1865. But as we will see below, the *Introduction* itself strongly suggests that Bernard was already familiar with Mill's views on the nature and scope of experimental methods. Furthermore, Bernard was acquainted with the methodological views of August Comte. In his *Cours de philosophie positive*, Comte (1838) argued that experimental methods had only limited power in biology (see especially pp. 320ff.). His views agreed closely with those of Mill, and not coincidentally. The *Cours* is known as a significant influence on Mill at the time of the composition of the *System of Logic* (Bourdeau 2018), and Comte and Mill corresponded extensively for several years (Haac 1995). Canguilhem (1967, 30–31) suggested that Bernard often contrasted his own methodological views with Comte's, even though, like Mill, Comte was not mentioned by name.[22]

Bernard's innovations within the *vera causa* tradition likely had a significant influence on subsequent methodological thinking. However, the story of the reception of Bernard's methodological thought is (unsurprisingly) complicated. In a survey of the reception of the *Introduction*, Grmek (1973, 8–16) noted that French scientists paid little attention to it in the first decades after its publication, while American and British scientists read it avidly. The book's philosophical reception was the reverse. Philosophers discussed it in France, but not in the United States, Britain, or Germany. But even if the *Introduction* was unevenly received, Bernard's scientific contributions were widely appreciated, and they may have provided a methodological tem-

[21] "[J]e m'occupe de l'invention scientifique qui a été negligée par les philosophes et même par Mill" (Bernard 1937, 32).

[22] For a study of Bernard's relationship to philosophy, and to Comtean Positivism in particular, see Virtanen (1960, Ch. 2 and 3).

plate even to those who were unfamiliar with Bernard's explicitly methodological works. What is more, a long series of international visitors passed through Bernard's laboratory and attended his lectures (Grmek 2008, 27). Visitors from England and the United States, in particular, spread Bernard's findings and methodology in their home countries (Olmsted 1935a, 1935b; Warner 2003).

With these preliminaries in place, let us turn to Bernard's methodological views. Like the traditional proponents of the *vera causa* standard, Bernard believed that the goal of experiment was to "connect natural phenomena with their necessary conditions or, in other words, with their immediate causes" (57).[23] He argued that the fundamental principles of reasoning about proximate causes were the same across the sciences, but that biology required "certain special principles of experimentation" (57).

The first sentence of the first chapter of the *Introduction* names the core issues of Bernard's methodological program:

> Only within very narrow boundaries can man observe the phenomena which surround him; most of them naturally escape his senses, and mere observation is not enough. (5)

As Bernard expanded this comment, it amounted to two claims: First, scientific reasoning is limited to observable causes only in the simplest cases, and scientists usually need to rely on intricate detection processes that make remote entities accessible. Second, "mere observation" only reveals the existence of phenomena. To elucidate causal relationships, sophisticated experimental practices are required. I will discuss the first point in Section 3.1, and the second in Section 3.2.

3.1 Detection and intervention

Bernard took an expansive view of the kinds of entities for which the existence requirement could be met. He wrote that scientists had to increase the power of their organs of sense "by means of special appliances", even as they equipped themselves with instruments that allowed them "to penetrate inside of bodies, to dissociate them and to study their hidden parts"

[23] Except where otherwise indicated, quotations from the *Introduction* are from the English translation by Henry Copley Greene (Bernard 1949), first published in 1927.

(5). Bernard identified a "necessary order" of research from simple investigations, which study objects that can be examined with unaided senses, to complex investigations, which "bring within our observation, by various means, objects and phenomena which would otherwise remain unknown to us forever" (5).

One way in which hidden causes could be made observable was by dissecting dead and living organisms. For instance, a gastric fistula gives access to the interior of the stomach for the study of digestion. If this was combined with experimental interventions on gastric nerves, the effects of those nerves on the production of gastric juices could be studied (9).

However, rendering observable but unobserved processes visible by dissection was only a first step. Unobservable components of organisms also needed to be rendered observable by appropriate means of detection. According to Bernard, in the life sciences "discovery of a tool for observation or experiment is much more useful than any number of systematic or philosophic dissertations" (171). To illustrate, he referred to his own research on hepatic glycogenesis, which depended on his ability to detect the presence of sugars in the blood vessels leading away from the liver. Such research could only be conducted "after chemistry gave us reagents for recognizing sugar, which were much more sensitive than those we had" (171). In his original publication on a "new function" of the liver, Bernard discussed at length the copper reduction test he employed, and he compared it in detail to other available methods for detecting sugars (Bernard 1853, 16–30). To Bernard, the transfer of such techniques represented a significant mode of disciplinary integration. "Chemistry is most useful to physiologists in giving them means of separating and studying individual compounds", he wrote (Bernard 1949, 73). The task of the physiologist was to use the access that chemistry granted to these compounds in order to study their role in living organisms.

Successful experimental inference usually required not only the detection of unobservable components, but also intervention upon them. In principle, observation could furnish appropriate contrasting instances for determining causal roles. Most of the time, however, it would be necessary to bring about the required contrasts by intervention:

> [E]xperimenters must be able to touch the body on which they wish to act, whether by destroying it or by altering it, so as to

> learn the part which it plays in the phenomena of nature. (9)

As in the case of detection, we should not conceive of "touching" an object merely in macroscopic terms. While the vivisectionist's scalpel was certainly one way of suppressing or altering organic processes, Bernard understood dissection broadly as "a displacing of a living organism by means of instruments and methods capable of isolating its different parts" (105). For this purpose, macroscopic vivisection was limited:

> Our instruments for vivisection are indeed so coarse and our senses so imperfect that we can reach only the coarse and complex parts of an organism. (104)

To reveal and intervene upon unobservable causes, more fine-grained interventions were necessary. Sometimes this would mean dissecting under a microscope, although this procedure was limited to animals sufficiently small to fit under the instrument. Even more subtle interventions were often needed:

> [W]hen we reach the limits of vivisection we have other means of going deeper and dealing with the elementary parts of organisms where the elementary properties of vital phenomena have their seat. We may introduce poisons into the circulation, which carry their specific action to one or another histological unit. (104)

Bernard's own work on toxic substances exemplifies the approach of using poisons as means of dissection (Bernard 1857, 1878). He wrote:

> I have particularly considered toxic agents as kinds of physiological instruments that are more delicate than our mechanical means, and that are destined to dissect, so to speak one at a time, the properties of the anatomical elements of the living organism.[24]

Bernard conducted a long series of studies on the physiological properties and mechanism of action of the arrow poison curare (Bernard 1857, Ch. 16–23). He already had one key instrument at his disposal: he was able to

[24] Bernard (1857, v, my translation).

stimulate nerves or muscles by electrical currents, that is, by "galvanizing" them. In the normal state, the stimulation of motor nerves would lead to muscular contractions, as would direct stimulation of the muscles themselves. Crucially, if curare was administered to animals ranging from frogs to dogs, electrical stimuli applied to motor nerves no longer had an effect, although the muscles continued to contract if stimulated directly. Meanwhile, sensation was not impaired. Bernard concluded that the poison affects only the motor nerves while leaving other physiological systems intact. He took this to explain how curare causes death: The poison eventually paralyzes the motor nerve leading to the diaphragm, thus stopping breathing. To demonstrate the correctness of the explanation, Bernard kept animals alive by forced ventilation (using a caoutchouc bladder to inflate their lungs) until the effects of curare passed, at which point they continued to live unimpaired.

Having found that curare affects the motor nerves but not the musculature, Bernard could now use the poison in further studies as an agent blocking motor nerves selectively. For example, he used it to study heart rate regulation (Bernard 1857, Ch. 25). Using curare as an instrument of (physiological) dissection allowed Bernard "to penetrate into the most hidden corners of our constitution".[25]

Bernard had succeeded in connecting death from curare poisoning with proximate causes. However, he cautioned that these causes were not some absolute limit of science. They were a limit only "with regard to our too feeble current means of investigation".[26]

Bernard's emphasis on gaining access to remote or unobservable causes by suitable instruments was not without precedent in the *vera causa* tradition. Reid, for example, remarked that "telescopes, microscopes, camera obscuras, [and] magic lanthorns … give just and true information, and the laws of nature by which they are produced, are of infinite benefit to mankind" (Reid 1863, 338).[27] Similarly, Herschel thought that instruments permitted scientists to extend their observational abilities and to expand the

[25] Bernard 1878, 298, my translation.

[26] Bernard 1878, 301, my translation.

[27] Wood (1989) has argued that Reid's demand for evidence of an entity's existence was not, *contra* Laudan, a demand for direct perception (see also Callergård 1999). Notably, Reid endorsed Benjamin Franklin's theory of the electrical fluid as an hypothesis to be pursued further.

store of entities that are known to exist. In one striking passage, he wrote that the telescope "has conferred upon [man], if not another sense, at least an exaltation of one already possessed by him that merits almost to be regarded as a new one" (Herschel 1830, §284, 256–57).

In the *Principles of Geology*, Charles Lyell (another prominent exponent of the *vera causa* standard) wrote that we must become "sensible of our natural disadvantages" and extend our senses:

> We are called upon, in our researches into the state of the earth, as in our endeavours to comprehend the mechanism of the heavens, to invent means for overcoming the limited range of our vision. We are perpetually required to bring, as far as possible, within the sphere of observation, things to which the eye, unassisted by art, could never obtain access. (Lyell 1830, 83)

He linked progress in geology to magnifying instruments, and in the true *vera causa* spirit he contrasted those who were "persevering in the attempt to improve their instruments" with those who were "engaged in the indolent employment of framing imaginary theories" (84).

Even Mill, in later editions of the *System*, granted that the luminiferous ether was "not in its own nature entirely cut off from the possibility of direct evidence in its favour" (Mill 1974, 499, III.XIV.§6). Evidence that the ether may be offering resistance to the motion of celestial bodies promised at least the possibility of detection. If confirmed, Mill thought that this might help the luminiferous ether to make "a considerable advance towards the character of a *vera causa*" (499, III.XIV.§6).

In brief, Bernard's emphasis on instruments for detection and intervention reflects a sophisticated version of the existence criterion. His position is not, however, a radical departure from earlier authors in the *vera causa* tradition, who consistently allowed for the detection of unobserved or unobservable entities by instruments. It would exceed the scope of this paper to study whether proponents of the tradition (including Bernard) ever gave a coherent account of how we establish access to unobservables without, at bottom, relying on some sort of consequentialist reasoning. But we can firmly reject the notion that the existence criterion was taken to bar unobservable entities from causal explanations.

3.2 Comparison and control

To demonstrate that a cause is competent to produce the effects ascribed to it, Bernard, like Herschel and Mill, took comparison to be key. He wrote that "science can be established only by the comparative method" (2), since it compares facts and "tests one by another" (5).[28] There are at least two senses in which Bernard took comparison to be fundamental to the methodology of science in general and of the life sciences in particular. He labeled these as "counterproof" and "comparative experiment".

Bernard regarded the counterproof as the main principle of experimental reasoning. Even if we already have "proof that a given condition always precedes or accompanies a phenomenon", this does not allow us to infer that the condition is really the cause of the phenomenon (55). What is required is a demonstration that removing the condition also removes the phenomenon. Bernard wrote that "the only proof that one phenomenon plays the part of cause in relation to another is by removing the first, to stop the second" (56). The counterproof was "a necessary and essential characteristic" of experimental reasoning which "decides whether the relation of cause to effect, which we seek in phenomena, has been found" (55–56). And Bernard cautioned that "proof, in science, never establishes certainty without counterproof" (56). The counterproof, then, was a test of causation, closely related to Herschel's seventh rule, or Mill's method of difference.

The proximity of Bernard's views to the *vera causa* tradition, and his rejection of consequentialism, are evident in his discussion of hypotheses. Like Mill, Bernard did not reject hypotheses outright. He thought that they were "auxiliaries to the [experimental] method, indispensable as scaffolding is necessary in building a house" (51). Anticipating a version of the discovery-justification distinction, he wrote that when "devising experiments or imagining means of observation … we must give free rein to our imagination" (24). But hypotheses must be "regulated and given a criterion" (24). That criterion was the counterproof:

> [W]e may say that in all experimental reasoning there are two possibilities: either the experimenter's hypothesis will be disproved [*infirmée*] or it will be proved [*confirmée*] by experiment. When experiment disproves his preconceived idea, the experi-

[28] Note well that in the French original, "to test" is "contrôler" (Bernard 1865, 12).

> menter must discard or modify it. But even when experiment fully proves [*confirme pleinement*] his preconceived idea, the experimenter must still doubt; for since he is dealing with an unconscious truth,[29] his reason still demands a counterproof. (52)

The differences between Bernard's methodology and consequentialism are glaring. Consequentialists would take an hypothesis to be judged by the agreement or disagreement of its consequences with observable phenomena – perhaps with the proviso that a subset of consequences counts disproportionately or exclusively, such as novel predictions on Whewell's account or falsifying instances on Popper's.[30] But Bernard insisted instead on the extraordinarily strict criterion of the counterproof, which requires us to detect and intervene upon a suspected cause in order to show that its suppression or alteration affects the phenomenon ascribed to it. Thus, Bernard agreed with Mill that hypotheses were allowable, useful, and often even necessary as scaffolding. But acceptance required direct evidence of existence and competence.

Bernard's account of method went far beyond those of the well-known exponents of the *vera causa* tradition in terms of detail and practicability. He recognized that the notion of a counterproof described only the ideal core of a reasoning strategy:

> Counterproof has not the slightest reference to sources of error that may be met in observing facts; it assumes that they are all avoided and is concerned only with experimental reasoning; it has in view only judging whether the relation established between a phenomenon and its immediate cause is correct and rational. (127)

In practice, then, the difficulty was to obtain suitable instances for comparison:

> In animals, and especially the higher animals, experimentation is so complex and liable to so many sources of error, both fore-

[29]Bernard referred to truths concerning the external world, as opposed to truths concerning subjective experience, as "unconscious" truths [*une véritée inconsciente*], cf. Bernard (1949, 28–29) and Bernard (1865, 51).

[30]See Malherbe (1981) for the altogether astounding claim that Popper's methodology is a formalization of Bernard's.

seen and unforeseen, that we must proceed most circumspectly to avoid them. (126)

Recall that Whewell doubted the power of Mill's experimental methods not because they were invalid as inference schemes, but because nature rarely affords us suitable instances to which the methods are applicable. Mill himself thought the methods inapplicable to complex systems because of the difficulty of finding comparable instances.

Bernard's answer to these difficulties was that the counterproof needed to be joined with a related but separate tool: comparative experimentation. While the counterproof was about reasoning from given facts about differences occurring upon intervention, comparative experimentation related to "ascertaining a fact and to the art of disengaging it from circumstances or from other phenomena with which it may be entangled" (127). Bernard believed that progress in physiology would come not only from progress in instruments and procedures, but most of all from the "reasoned and well-regulated use of *comparative experimentation*" (Bernard 1865, 221, my translation and original emphasis).

Bernard argued that comparative experimentation was "not exactly what philosophers have called the method of difference".[31] According to Bernard, the method of difference suggests that we separate all the parts of a system (Mill's conditions A B C) and assign to each part its proper effect (Mill's phenomena *a b c*). In biology, however, it is rare that we can enumerate parts comprehensively, let alone assign individual effects to each of them. Mill had concluded from this that direct experimentation was not feasible in sufficiently complex systems. But Bernard diagnosed the problem differently. He thought that Mill's articulation of the method of difference demanded too much to be practicable, but also more than was needed for an inference to causality.

Comparative experimentation did not require the comprehensive enumeration of all the conditions that influence a phenomenon. Instead, the task was only "completely to isolate the one phenomenon on which our studies are brought to bear, separating it … from all surrounding complications" (128). Bernard wrote:

[31] "L'expérimentation comparative n'est pourtant pas précisément ce que les philosophes one appelé la méthode par différence" (Bernard 1865, 222, my translation). Although no philosophical writers are cited, it is likely that this is a reference to Mill's *System of Logic* (see the discussion on p. 19).

> Comparative experimentation reaches this goal by adding to a similar organism, used for comparison, all our experimental changes save one, the very one which we intend to disengage. (128)

If, for example, we wished to know the effect of ablating a deep-seated organ, there is a danger of confusing the effects of the ablation itself with the effects of the operative procedures required for gaining access to the organ in the first place. The operation needed to be performed twice, once in its entirety and once with the omission of the actual ablation:

> We thus have two animals in which all the experimental conditions are the same, save one,—ablation of an organ whose action is thus disengaged and expressed in the difference observed between the two animals. (128)

Bernard's crucial suggestion was that we can draw conclusions from experiments on complex systems even when we are ignorant of a great many conditions that influence the effect under study. We do not need to know about them because they are, as it were, cancelled out by a suitably chosen comparison. Bernard claimed boldly that the result of comparative experimentation is "to eliminate by a single stroke all known and unknown sources of error" (127).

Bernard's experimental method is thus neither identical to Mill's method of difference nor an alternative to it. Both methods deploy the same criterium for diagnosing causal relationships: that the effect must disappear upon removal of the cause, other things being equal. This is what Bernard called the counterproof. But the two authors differ in their treatment of the *ceteris paribus* requirement. Mill articulated very strict prerequisites for experimental inferences that could rarely, if ever, be met in practice. Bernard's comparative experimentation relaxed these prerequisites. It was a way to realize counterproofs even in complex systems about which much remained unknown.[32]

[32]My interpretation differs from Jutta Schickore's in her recent book *About Method* (2017, Ch. 7–9). I agree that Bernard's comparative experimentation was not merely an "application" of Mill's method of difference. However, Schickore took Bernard's counterproof to be identical to the method of difference, and argued that Bernard rejected the counterproof in favor of comparative experimentation for the purposes of biological inquiry (see especially pp. 124–131). But Bernard understood the counterproof as indispensable to all experimental reasoning, and comparative experimentation not as an alternative to it, but merely as a means "in complex circumstances, to simplify phenomena and to forearm oneself against

Bernard refrained from discussing the theory of comparative experimentation at great length and instead referred the reader to concrete examples. These examples reveal at least four ways in which comparative experimentation helped to avoid errors. First, as already noted, comparative experimentation implements the counterproof. Bernard told the story of his researches into the fate of sugars in metabolism. Early on, he had fed sweetened milk soup to a dog and then found sugar in the hepatic vessels. The natural assumption was that the sugar he had found derived from the soup. However, he conducted a comparative experiment on another dog who had been fed only meat. That is, the suspected cause of the sugar in the hepatic vessels, the dietary sugar, was removed. "Great was my astonishment", wrote Bernard, "at finding that the blood of the animal which had not eaten any also contained sugar" (182). The counterproof showed that dietary sugar was *not* necessary for the presence of sugar in the hepatic vessels. Eventually, this led Bernard to the recognition that the liver synthesized sugar regardless of the composition of the dog's diet.[33]

Second, comparative experimentation contributes to the test of the detection procedures involved in the experiment. In the case discussed above, Bernard had used the copper reduction test to detect sugar in the hepatic vessels. While copper was known to react with sugar, Bernard hypothesized that this "empirical characteristic" might also be shown "by substances still unknown in the bodily economy" (182). A positive result in the comparative experiment could have been an indication that the copper reduction tests produced such false positives.

Third, experimentation on comparable experimental objects reduces the risk of confounding causes. The crucial question, of course, was whether two experimental objects are in fact comparable. Bernard treated this as an empirical matter. For instance, he was interested in the effect of various substances on the glycogen content of the liver. However, even on similar

unforeseen sources of error" (Bernard 1949, 56). Notice that Bernard often referred to his own comparative experiments as furnishing counterproofs (Bernard 1949, e.g., 153, 156, 164).

[33]This eventually led to the discovery of the process of hepatic glycogenesis, and it has been studied in depth on the basis of extant laboratory notebooks (Grmek 1968, Holmes 1974, Ch. 18–19). While Bernard's summary in the *Introduction* condenses and rearranges the chronology of the process of discovery, it appears to describe the role of the counterproof accurately. In this respect, Bernard's methodological pronouncements matched his actual experiments.

diets, Bernard failed to find animals that were comparable in this respect. He wrote:

> According to their age, sex, plumpness, etc., animals bear starvation better or worse and destroy more glycogen or less, so that I could never be sure that the differences I found were the result of differences in diet. (183)

Bernard therefore conducted the experiment on a single animal. He removed one part of the liver before the experimental intervention (an injection of food), and a second part afterwards, thus ensuring comparability. Similarly, to study the effect of contraction on respiration in frog muscles, the experiments needed to be performed on two limbs of the same animal, "because in this respect two frogs are not always comparable" (183). Thus, the appropriate comparison depended on the phenomenon under study. If a phenomenon varied markedly due to unknown differences between organisms, then the comparison needed to occur within-organism.

Fourth and finally, comparative experimentation mitigates the risk of fat-handed interventions. A fat-handed intervention is one that influences the effect under study by a causal pathway that does *not* run through the suspected cause that is the target of the intervention. For instance, Bernard reported an experiment by Magendie, in which the object was to determine the functions of cerebrospinal fluid. Experiments suggested that the removal of cerebrospinal fluid resulted in a characteristic disturbance of motions. However, an accident revealed that the preparatory steps of the operation sufficed to cause the disturbance. "Comparative experimentation", Bernard wrote, "would obviously have solved the difficulty" (182–83).

To conclude, Bernard's comparative experimentation allowed inferences even in contexts where known and even unknown confounders were in play. In this respect it improved significantly on earlier articulations of the method of difference such as Mill's.

4 Conclusions

Claude Bernard did not consider himself a philosopher. He thought that the truth is found by improving our techniques of investigation, and that "the best philosophic system consists in not having any" (221). He rejected the builders of philosophical systems as vehemently as the positivists who opposed them. He polemicized that in research, those who knew Bacon and his modern-day successors best achieved the least. Nevertheless, Bernard must be understood, as I have done here, as a significant contributor to nineteenth-century philosophical reflection on the nature and methods of science.

On a received view of the history of scientific method, the *vera causa* standard's influence declined around the middle of the nineteenth century. With its insistence on direct evidence for the existence and competence of causes, the standard came to be regarded as a cumbersome restriction on scientific inquiry into unobservables, and it was therefore abandoned in favor of a more powerful consequentialist alternative. The famous debate between Mill and Whewell has been taken to exemplify this shift, with Mill defending the moribund *vera causa* standard against Whewell's consequentialism. Thus the old standard gave way to a version of the consequentialism that twentieth-century philosophers themselves favored.

In Bernard's writings, however, we find a sophisticated version of the *vera causa* standard that addresses many of the epistemic limitations usually ascribed to the approach. Bernard argued that unobservable causes needed to be made accessible by suitable chemical and physical means for detecting and intervening upon them. In addition, he developed a flexible account of comparative experimentation that allowed him to infer causal relationships even in complex systems in which known and unknown confounders existed. Bernard's methodology was no doubt highly demanding in terms of the evidence that it required for claims to be established, but it could not be accused of limited scope.

None of this is to say that Bernard single-handedly solved all of the outstanding problems of the *vera causa* tradition. Although his methodology marked a powerful advance over previous articulations of the standard, it faced its own limitations. For example, Bernard thoroughly rejected statistical approaches. He argued that when all the conditions that are relevant to an effect have been found, the effect occurs "always without exception" (137), and so statistical analysis is superfluous in mature science. But his successors of course came to rely heavily on probabilistic and statistical tools in order to enable inferences in situations where full-fledged deterministic understanding was far out of view. Thus, just as Bernard relaxed

Mill's very strict prerequisites for valid experimental inferences, his successors would find ways to relax Bernard's own methodological demands further, and to expand the scope of experimental inquiry again. In this and other respects, Bernard is an important link rather than an endpoint in the history of scientific epistemology.

It would be merely facetious to declare Bernard the out-of-competition winner of the debate between Mill and Whewell. But Bernard was a significant exponent of a non-consequentialist tradition of scientific reasoning and practice, whose scope and power have been underappreciated. Arguably, this tradition, with its focus on establishing *veræ causæ*, still underpins much of experimental biology today and for this reason alone must be better understood. But there are broader philosophical implications. It was fittingly Laudan (1995) who identified consequentialist theories of evidence as a curse on all of philosophy. If he was right, then there is still much to be learned from the tradition whose tenets Bernard helped to articulate.

Acknowledgments

For detailed comments on an earlier draft of this manuscript, I thank William Bechtel, Tim Lewens, Laurent Loison, Aaron Novick, and an anonymous reviewer for *Studies*. It was Aaron who originally made me aware of Charles Lyell's discussion of the value of instruments. For comments on early versions of this material, I also thank audiences at the Oberseminar Wissenschaftsgeschichte in Munich, at the 8th Quadrennial International Fellows Conference of the Pittsburgh Center for Philosophy of Science, held in Lund, at the University of Geneva's Department of Philosophy, at the University of Cambridge's Department of HPS, and at the IHPST in Paris. This work was supported in part by a grant from the Swiss National Science Foundation (grant number P300P1_154590).

References

Achinstein, P. 1992. Inference to the best explanation: Or, who won the Mill-Whewell debate? *Studies in History and Philosophy of Science* 23 (2): 349–364.

Allen, R. 2017. David Hartley. In *The Stanford Encyclopedia of Philosophy*, Summer 2017, ed. by E. N. Zalta. Metaphysics Research Lab, Stanford University.

Bernard, C. 1853. *Nouvelle fonction du foie*. Paris: J. B. Baillière.

——— 1857. *Leçons sur les effets des substances toxiques et médicamenteuses*. Paris: J. B. Baillière et fils.

——— 1865. *Introduction à l'étude de la médecine expérimentale*. Paris: J. B. Baillière et fils.

——— 1878. Le curare. In *La science expérimentale*, 237–315. Paris: J. B. Baillière et fils.

——— 1937. *Pensées. Notes détachées.* Ed. by L. Delhoume. Paris: J.-B. Baillière et fils.

——— 1949. *An Introduction to the Study of Experimental Medicine*. Translated by Henry Copley Greene. New York: Henry Schuman.

——— 1950. *Lettres beaujolaises*. Ed. by J. Godart. Villefranche-en-Beaujolais: Cuvier.

Biener, Z. 2018. Newton's Regulae Philosophandi. In *Oxford Handbook of Isaac Newton*, ed. by E. Schliesser and C. Smeenk. Oxford University Press.

Bourdeau, M. 2018. Auguste Comte. In *The Stanford Encyclopedia of Philosophy*, Summer 2018, ed. by E. N. Zalta. Metaphysics Research Lab, Stanford University.

Buchdahl, G. 1971. Inductivist versus deductivist approaches in the philosophy of science as illustrated by some controversies between Whewell and Mill. *The Monist* 55 (3): 343–367.

Butts, R. E. 1970. Whewell on Newton's rules of philosophizing. In *The Methodological Heritage of Newton*, ed. by R. E. Butts and J. W. Davis, 132–149. University of Toronto Press.

——— 1973. Reply to David Wilson: Was Whewell interested in true causes? *Philosophy of Science* 40 (1): 125–128.

Callergård, R. 1999. The hypothesis of ether and Reid's interpretation of Newton's first rule of philosophizing. *Synthese* 120 (1): 19–26.

——— 2006. An Essay on Thomas Reid's Philosophy of Science. PhD thesis, Acta Universitatis Stockholmiensis.

Canguilhem, G. 1967. Théorie et technique d l'expérimentation chez Claude Bernard. In *Philosophie et méthodologie scientifiques de Claude Bernard*, ed. by E. Wolff. Masson & Cie.

Cobb, A. D. 2011. History and scientific practice in the construction of an adequate philosophy of science: revisiting a Whewell / Mill debate. *Studies in History and Philosophy of Science* 42 (1): 85–93.

——— 2012. Inductivism in practice: Experiment in John Herschel's philosophy of science. *HOPOS: The Journal of the International Society for the History of Philosophy of Science* 2 (1): 21–54.

Cohen, I. B. 1962. The First English Version of Newton's *Hypotheses non fingo*. *Isis* 53 (3): 379–388.

Comte, A. 1838. *Cours de philosophie positive, Tome troisième, contenant la philosophie chimique et la philosophie biologique*. Paris: Bachelier.

Cuneo, T., and R. Van Woudenberg, eds. 2004. *The Cambridge Companion to Thomas Reid*. Cambridge University Press.

Di Fate, V. J. 2011. Is Newton a radical empiricist about method? *Studies in History and Philosophy of Science Part A* 42 (1): 28–36.

Ducasse, C. J. 1960. John F. W. Herschel's methods of experimental inquiry. In *Theories of Scientific Method: The Renaissance Through the Nineteenth Century*, ed. by R. M. Blake, C. J. Ducasse, and E. H. Madden, 153–82. University of Washington Press.

Ducheyne, S. 2006. Reid's adaptation and radicalization of Newton's natural philosophy. *History of European Ideas* 32 (2): 173–189.

——— 2008. J. S. Mill's Canons of Induction: from true causes to provisional ones. *History and Philosophy of Logic* 29 (4): 361–376.

——— 2015. An editorial history of Newton's regulae philosophandi. *Estudios de Filosofía* 51: 143–164.

Fisch, M. 1985. Necessary and contingent truth in William Whewell's antithetical theory of knowledge. *Studies in History and Philosophy of Science* 16 (4): 275–314.

Gayon, J. 2015. Les réflexions méthodologiques de Claude Bernard: structure, contexte, origines. In *L'épistemologie française, 1830–1970*, ed. by M. Bitbol and J. Gayon, 217–236. Éditions matériologiques.

Grmek, M. D. 1968. First Steps in Claude Bernard's Discovery of the Glycogenic Function of the Liver. *Journal of the History of Biology* 1 (1): 141–154.

——— 1973. *Raisonnement expérimental et recherches toxicologiques chez Claude Bernard*. Genève: Droz.

——— 2008. Claude Bernard. In *Complete Dictionary of Scientific Biography*, 2:24–34. Charles Scribner's Sons. http://link.galegroup.com/apps/doc/CX2830900401/GPS?u=cambuni&sid=GPS&xid=ba410523.

Haac, O. A. 1995. *The Correspondence of John Stuart Mill and Auguste Comte*. Transaction Publishers.

Hartley, D. 1749. *Observations on Man, His Frame, His Duty, and His Expectations*. London: S. Richardson.

Herschel, J. 1830. *Preliminary Discourse on the Study of Natural Philosophy*. London: Longman, Rees, Orme, Brown, / Green / John Taylor.

Hodge, M. 1977. The structure and strategy of Darwin's "long argument". *The British Journal for the History of Science* 10 (3): 237–246.

——— 1992. Darwin's argument in the *Origin*. *Philosophy of Science* 59 (3): 461–464.

Holmes, F. L. 1974. *Claude Bernard and Animal Chemistry*. Harvard University Press.

Kavaloski, V. C. 1974. The *Vera Causa* Principle: A historico-philosophical study of a metatheoretical concept from Newton through Darwin. PhD thesis, The University of Chicago.

Laudan, L. 1981a. Thomas Reid and the Newtonian turn of British methodological thought. Chap. 7 in *Science and Hypothesis: Historical Essays on Scientific Methodology*, 86–110. The University of Western Ontario Series in Philosophy of Science.

——— 1981b. The epistemology of light: Some methodological issues in the subtle fluids debate. Chap. 8 in *Science and Hypothesis: Historical Essays on Scientific Methodology*, 111–140. The University of Western Ontario Series in Philosophy of Science.

——— 1981c. Why was the Logic of Discovery abandoned? Chap. 11 in *Science and Hypothesis: Historical Essays on Scientific Methodology*, 181–191. The University of Western Ontario Series in Philosophy of Science.

——— 1981d. *Science and Hypothesis: Historical Essays on Scientific Methodology*. The University of Western Ontario Series in Philosophy of Science.

——— 1995. Damn the consequences! *Proceedings and Addresses of the American Philosophical Association* 69 (2): 27–34.

Lipton, P. 2004. *Inference to the Best Explanation*. Routledge.

Lyell, C. 1830. *Principles of Geology: An Attempt to Explain the Former Changes of the Earth's Surface, by Reference to Causes Now in Operation*. Vol. 1. London: John Murray.

Malherbe, D. J.-F. 1981. Karl Popper et Claude Bernard. *Dialectica* 35 (4): 373–388.

Mill, J. S. 1843. *A System of Logic*. London: John W. Parker.

——— 1974. A System of Logic. In *Collected Works of John Stuart Mill*, ed. by J. M. Robson, vol. 7–8. University of Toronto Press / Routledge & Kegan Paul.

Newton, I. 1713. *Philosophiae naturalis principia mathematica, Editio secunda auctior et emendatior*. Cornelius Crownfield. https://doi.org/10.3931/e-rara-1237.

——— 1729. *The Mathematical Principles of Natural Philosophy*. Translated by Andrew Motte. In Two Volumes. London: Motte.

——— 1999. *The* Principia. *Mathematical Principles of Natural Philosophy. A New Translation by I. Bernard Cohen and Anne Whitman assisted by Julia Budenz. Preceded by a Guide to Newton's* Principia *by I. Bernard Cohen*. University of California Press.

Novick, A. N. 2016. Metaphysics and the vera causa ideal: The nun's priest's tale. *Erkenntnis*. https://doi.org/10.1007/s10670-016-9863-1.

Novick, A., and R. Scholl. 2018. *Presume It Not: True Causes in the Search for the Basis of Heredity*. Forthcoming in *The British Journal for the Philosophy of Science*. http://dx.doi.org/10.1093/bjps/axy001.

Olmsted, J. M. D., and E. Harris Olmsted. 1952. *Claude Bernard & the Experimental Method in Medicine*. New York: Henry Schuman.

Olmsted, J. 1935a. The influence of Claude Bernard on medicine in the United States and England: Part I. *California and Western Medicine* 42 (2): 111–113.

——— 1935b. The Influence of Claude Bernard on medicine in the United States and England: Part II. *California and Western Medicine* 42 (3): 174–176.

Pence, C. H. 2018. Sir John F. W. Herschel and Charles Darwin: Nineteenth-century science and its methodology. *HOPOS: The Journal of the International Society for the History of Philosophy of Science* 8 (1): 108–140.

Reid, T. 1863. Essays on the Intellectual Powers of Man. In *The Works of Thomas Reid*, ed. by W. Hamilton. Edinburgh: Maclachlan / Stewart.

Schickore, J. 2017. *About Method: Experimenters, Snake Venom, and the History of Writing Scientifically*. University of Chicago Press.

Scholl, R. 2015. Inference to the Best Explanation in the Catch-22: How Much Autonomy for Mill's Method of Difference? *European Journal for Philosophy of Science* 5 (1): 89–110.

Smith, G. E. 2002. The Methodology of the *Principia*. In *The Cambridge Companion to Newton*, ed. by I. B. Cohen and G. E. Smith, 138–73. Cambridge: Cambridge University Press.

Snyder, L. J. 1997a. Discoverers' induction. *Philosophy of Science* 64 (4): 580–604.

——— 1997b. The Mill-Whewell debate: Much ado about induction. *Perspectives on Science* 5 (2): 159–198.

——— 2006. *Reforming Philosophy: A Victorian debate on science and society*. University of Chicago Press.

——— 2009. Hypotheses in 19th century British philosophy of science: Herschel, Whewell, Mill. In *The Significance of the Hypothetical in the Natural Sciences*, ed. by M. Heidelberger and G. Schiemann, 59–76. Walter de Gruyter.

Strong, E. W. 1955. William Whewell and John Stuart Mill: their controversy about scientific knowledge. *Journal of the History of Ideas* 16 (2): 209–231.

Virtanen, R. 1960. *Claude Bernard and His Place in the History of Ideas*. Lincoln: University of Nebraska Press.

Warner, J. H. 2003. *Against the spirit of system: The French impulse in nineteenth-century American medicine*. Johns Hopkins University Press.

Whewell, W. 1840. *The Philosophy of the Inductive Sciences, Founded Upon Their History*. London: John W. Parker.

——— 1847. *The Philosophy of the Inductive Sciences, Founded Upon Their History*. 2nd ed. London: John W. Parker.

——— 1849. *Of Induction, with Especial Reference to Mr. J. Stuart Mill's System of Logic*. London: John W. Parker.

——— 1860. *On the Philosophy of Discovery, Chapters Historical and Critical*. London: John W. Parker.

Wilson, D. B. 1973. Butts on Whewell's view of true causes. *Philosophy of Science* 40 (1): 121–124.

Wood, P. 1989. Reid on hypotheses and the ether; a reassessment. In *The Philosophy of Thomas Reid*, ed. by M. T. Dalgarno and E. H. Matthews, 433–446. Springer.

Yeo, R. 1993. *Defining Science: William Whewell, Natural Knowledge and Public Debate in Early Victorian Britain*. Cambridge University Press.