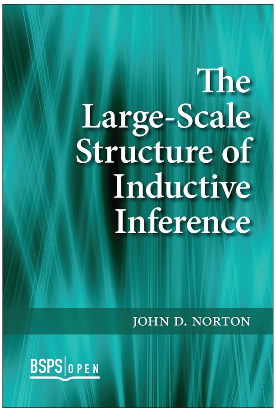
UNIVERSITY OF CALGARY
Press



THE LARGE-SCALE STRUCTURE OF INDUCTIVE INFERENCE

John D. Norton

ISBN 978-1-77385-541-7

Acknowledgement: *We acknowledge the wording around open access used by Australian publisher,* **re.press**, *and thank them for giving us permission to adapt their wording to our policy* http://www.re-press.org

4

The Uniqueness of Domain-Specific Inductive Logics

1. The Challenge Posed

According to the material theory of induction, the inductive relations within a mature science form a self-supporting structure.[1] That is, the propositions of the science derive their inductive support entirely from an extensive body of empirical evidence, such that each proposition in a theory is supported individually by this body of evidence through the mediation of other propositions. Those other propositions are supported in turn in the same way.

This raises a challenge: what assurance do we have of the uniqueness of the resulting relations of inductive support? We should not expect such an assurance for a developing science sustained only by a fragmentary body of evidence. In such cases, the evidence is too weak to determine unique relations. But what of the case of a mature science in which the body of evidence is sufficiently expansive to provide strong evidential support for all of the propositions of the science? Is such a science uniquely supported? Might there be a second science whose propositions contradict the first science but is as strongly supported in all of its parts by the same body of evidence?

Were there such cases, the result would be inductive anarchy, and it would be of an especially troublesome kind within the context of the material theory of induction. Each set of facts proposed by each science would support its own inductive logic. Since the facts disagree, the resulting logics would not agree on the bearing of evidence. One could find propositions in a science

1 My thanks to James Woodward for helpful comments on an earlier draft of this chapter.

supported inductively or not according to which of the inductive logics is employed.

Perhaps we can find reasons to expect such multiple systems. If we think of relations of support as analogous to relations of structural support in a building, then we can erect very different self-supporting systems of masonry on the same foundations. So why do we not have multiple systems of inductive logic?[2] The underdetermination thesis in its strongest form is the grim speculation that no body of evidence, no matter how extensive, can determine the content of a theory. Inductive pessimists who find this speculation appealing will expect multiple systems as a matter of course. My goal in this chapter is to refute this inductive pessimism by means of three arguments.

First, if the underdetermination thesis were true, then all sciences, even the most mature, would be awash in incompatible competitor sciences that enjoy comparable inductive support by the evidence. As a matter of history, this is not the case. Rather, as I will review briefly in Section 2, once a science achieves maturity, its competitors are discarded, and a single science prevails and endures. Since the underdetermination thesis is accepted in some literatures as a truism of evidence, in Section 7 I review briefly why it is really a poorly grounded speculation, better called the underdetermination conjecture.

Second, competing relations of support derive from competing theories that make incompatible factual assertions. As I will argue in Section 3, the empirical character of science requires that such factual differences be reflected in differences of empirical evidence, or they lie outside the scope of empirical sciences. It follows that empirical evidence can always decide for some and against others of the competing theories. (I will develop this concern in Sections 7 and 8 in further discussion of the underdetermination thesis.)

2 Here the analogy to buildings is weak and misleading, for we can imagine a flat terrain on which we can erect many great cathedrals of differing design, selected according to our whims. However, a body of evidence analogous to this featureless terrain is bereft of evidential value. It can sustain only the thinnest of inductive logics such as the relations of "completely neutral support" described in *The Material Theory of Induction* (Norton 2021). The analogy improves somewhat if we imagine building on a complicated and richly structured terrain that admits only specific modes of construction. The empirical foundation of our science should be structured richly enough to direct us to a fuller content of the science itself.

Third, as I will relate in Section 4, there is a natural mechanism peculiar to the material theory of induction that favors the emergence of uniqueness. Understood materially, the competition between scientific theories is dynamically unstable as long as continuing attention is given to the full exploration of the evidence. If one theory gains an evidential advantage over another, then that theory's inferential powers are enhanced. According to the material theory of induction, facts warrant inductive inferences. Thus, the evidentially strengthened theory has secured more facts and with them a strengthened warrant to infer inductively to still more facts. The competing theory is correspondingly weakened. If this process continues, then it amplifies the advantage in a positive feedback loop and leads one theory to dominate and eventually eliminate its competitors.

These instabilities are illustrated in Section 5 with several examples. Two later chapters provide more extended examples. In Chapter 14, "Stock Market Prediction: When Inductive Logics Compete," we see that there are multiple systems currently in use for predicting price movements in the stock market. The chapter shows that they are in unstable competition and that a proper pursuit and weighing of the evidence would lead to one dominating the others. In Chapter 13, "Dowsing: The Instabilities of Evidential Competition," I recount how the practice of dowsing emerged in the sixteenth century. Even then it was a controversial practice. Two views competed: the proponents of dowsing and skeptics who argued that the practice was ineffective. Over the ensuing centuries, the evidential case for the skeptics made self-reinforcing advances that successively undermined the scientific credibility of dowsing until it collapsed.

In concluding sections, I consider standard challenges in the literature to the uniqueness claimed in this chapter. What of challenges to any theory by unconceived alternatives? Does not the already mentioned underdetermination thesis preclude uniqueness? What of observationally equivalent theories? In Sections 6, 7, and 8, I discuss each question and argue that none supports a cogent challenge. In Section 9, I argue that a material approach to inductive inference fares better at accommodating the uniqueness of inductive support of mature science than do formal accounts. In Section 10, I provide a brief summary and conclusion.

2. The Uniqueness of Mature Sciences

Once a science reaches maturity in its domain of application, it stabilizes and remains fixed. The effect is so familiar that we need to recall only a few instances. At the level of precision required for virtually all applications, Euclid's ancient geometry has sufficed up to the present day. Deviations from it arise, according to general relativity, only when we venture well beyond the realm in which Euclid's geometry found its evidential support: that is, we explore systems with intense gravity or those on cosmological scales. At the level of precision for even the most exacting dynamic systems, Newton's seventeenth-century mechanics has sufficed up to the present day. Deviations appear only in domains remote from those in which Newton's mechanics is well supported evidentially. Examples of these remote domains are systems moving close to the speed of light or those at atomic scales, where quantum effects are important. The chemistry of common materials is based on a system of elements secured in the nineteenth century, deriving from the work of Lavoisier and its codification in the periodic table of Mendeleev. The diversity of geological structures derives from Lyell's early-nineteenth-century uniformitarianism, and the variety of life forms derives from Darwin's mid-nineteenth-century theory of evolution. The examples can be multiplied. The uniqueness of mature sciences contradicts the proliferation predicted by the underdetermination conjecture.

It can be tempting to imagine that the dominance of one mature science does not derive from the weight of evidence. It is, we might speculate darkly, merely a reflection of local conditions such as external social factors or political pressures or even the concerted fraud of scientists. Of course, aberrations are possible when local conditions eclipse the proper weighing of evidence. When they arise, such aberrations do not survive changes of location and time. Lysenko's mid-twentieth-century corruption of biology in Soviet Russia depended on his political power and support. Lysenkoism failed when that support was lost. It was bad science, unsupported by evidence. What is distinctive about mature sciences is their uniformity across culture and time. The geometry of Euclid might have been codified in fourth-century BCE Alexandria, yet it long escaped its Alexandrian roots to become the geometry used internationally and for millennia, without serious challenge, until tiny corrections were required by general relativity in the twentieth century.

3. Competition Is Empirically Decidable

Competing systems of relations of evidential support derive from competing theories. They compete in the sense that they make incompatible factual claims about the world. Since science is empirical, such competition cannot be sustained indefinitely. The empirical character of science requires that the factual claims of a theory must be supported inductively by the evidence of observation and experiment.[3] To respect this empirical character, the competition among incompatible factual claims of competing theories must be resolvable by observation and experiment. If their factual differences are beyond observation or experiment, then whatever constitutes these differences lies outside empirical science.[4] It follows that there must be some possible observation or experiment capable of deciding among competing theories. The competition will be resolved as long as scientists are diligent and inventive enough in their pursuit of empirical evidence.

A radical, skeptical view holds that there are limited prospects for this sort of comparison. The worry is that observations are so theory laden that they are useless for comparisons of theories. Theories become, to use Kuhn's (1996, Chapters X, XII) expression for paradigms, "incommensurable" or, more simply, beyond cogent comparison. I do not share this skepticism. Theories can be compared on their adequacy to the empirical evidence and are routinely thus compared. The best account of this comparison is provided by Nora Boyd's (2018a, 2018b) empiricism, already mentioned in Chapter 2. Boyd shows that, if we are to decide between two theories on the basis of some item of evidence, the procedure is to wind back toward the provenance of the evidence. We continue until we have stripped away enough of the theoretical encumbrances to have freed the statements of evidence of entanglement with the theoretical presumptions of either theory.

3 To preclude confusion, the empiricism advocated here is what I call "small e" empiricism. It is the widely held view that we can learn our sciences only from our experiences. It is distinct from antirealist versions of "big E" Empiricism, such as van Fraassens' (1980) constructive empiricism in which *all* that we know of the world is *only* what we can or could experience directly.

4 Further analysis might be needed, however. The two theories might appear to be different only since they merely represent the same facts in different guises. Perhaps one or both theories contain content superfluous to the empirical successes of the theories.

These decisions need not be immediate. However, when empirical evidence favors one theory over another, it introduces an instability that must be resolved. Competing theories are responsible to all of the empirical evidence in their domains of application. A faltering theory can choose to ignore or discount unfavorable empirical evidence only temporarily while awaiting rescue from further evidence. Alternatively, the faltering theory can make internal adjustments to accommodate the unfavorable evidence. Such adjustments weaken the theory and make it more prone to further weakening.

These considerations would not apply to pairs of theories in one domain whose empirical content is so disjointed that they never disagree on what is observable while still retaining their identities as distinct theories. Although I grant this possibility in principle, I have had trouble finding real examples. Candidates can be sought in theories that treat some domain at very different scales both in size and in time. Perhaps neuroscience and psychology are cases in which both theories treat what is essentially just brain activities. They use different theoretical devices without intersecting or intersecting much empirically. Although this disjointed character is possible, neuroscientists in particular work energetically to breach it. I discuss another candidate briefly in Chapter 14, "Stock Market Prediction: When Inductive Logics Compete." There are different systems for predictions of moves in stock prices. Insofar as one system might make predictions only in the shorter term and another might make them over the longer term, it might be possible for them to proceed from disjointed factual bases. Although this is a possibility in principle, it does not seem to have been realized.

4. Inductive Competition Is Unstable

When one theory, in competition with another, gains a slight evidential advantage, it follows from the material nature of inductive inference that this advantage will be amplified. Facts warrant inductive inference, and the more facts a theory has secured the more it can infer inductively.

The role of hypotheses in a developing science can make this process of amplification potent. As we have seen, when the body of evidence supporting a science is meager, or the import of the existing evidence has not yet been fully explored, the scientists proceed in their investigations by positing hypotheses of suitable strength to warrant their inferences. These hypotheses must eventually be given suitably strong evidential support. During the preliminary period, it is possible to sustain multiple systems of facts and the inductive logics that they induce. Systems in competition will be distinguished by their employment of incompatible hypotheses. The viability of these multiple systems is fragile and unstable. If one system gains a small advantage through the import of novel evidence or a novel interpretation of existing evidence, then that small gain strengthens the system, in particular lending more support to its founding hypotheses. The competing systems are correspondingly weakened. This momentary advantage can persist and be amplified, or a weakened system itself can find new evidence that restores its support. However the competition might play out, its dynamics is unstable and overall tends to favor further strengthening of the system that has gained a small inductive advantage. The tendency then is for the advantaged system to be strengthened still further, whereas those in competition find it harder to recover. The dynamics drives toward the dominance of one system and the elimination of others.

5. Illustrations of Instability

A detailed examination of the competition described in Section 4 in particular cases would be lengthy. In later chapters, I provide such examinations in the cases of competing systems of stock market prediction and the historical competition between proponents and skeptics of dowsing. Here I can describe other cases only briefly. To do so, I draw from the convenience provided by Chapter 9 of *The Material Theory of Induction* (Norton 2021). As part of its analysis of the argument form "inference to the best explanation," the chapter reviews pairs or sets of theories in competition. We can see in these examples how each theory gains an evidential advantage while disadvantaging its competitors. Here I will not recount the details of the competing theories but only the dynamics of the competition. I refer readers to this chapter in *The Material Theory of Induction* for further details and citations of the pertinent literature.

5.1. Darwin's *Origin of Species*

In his *Origin of Species*, Darwin developed his theory of the origin of diverse biological forms through natural selection. It is portrayed throughout as in competition with the proposal that this diversity arises from the independent creation of each of these forms. Darwin argued that advantageous features of organisms arise through one process, their selection by nature. However, independent creation must attribute each new feature to a new decision by a Designer to create each organism just as it is. More telling are examples of organisms with features that have no apparent advantage. Why do terrestrial geese, for example, have webbed feet when webbing is useful only in water? Darwin gives an evolutionary account: terrestrial geese evolved from aquatic geese. Independent creation can attribute the webbed feet only to a capricious decision by the Designer.

With each successful accounting of advantageous and otherwise anomalous attributes, Darwin's original hypothesis of natural selection gains evidential support. Each of these successes weakens the competing hypothesis of independent design, which accumulates a growing burden of independent and capricious design decisions. The accumulation of these successes amplifies the evidential advantage of natural selection. It is moved from plausible speculation to a well-supported proposition while its competitor, independent creation, languishes.

5.2. Lyell's *Principles of Geology*

Uniformitarian geology asserts that present-day geological features were produced slowly by processes still acting in the present. Lyell's *Principles of Geology* made the case for it. Lyell was in a polemical dispute with competing catastrophist theories. They accounted for the same features by processes not currently acting and often of great violence. The initial advantage of the catastrophists was that it is natural to imagine great mountains and deep valleys as created by sudden, momentous events. Lyell chipped away at this advantage by showing how one geological feature after another can arise from currently acting processes. To use an example that he promoted, a competing account of fossils is that they arise in stone from a "plastic virtue, or some other mysterious agency." Lyell, however, accounted for them in terms of the fossilization of ordinary living things.

The evidential dynamic is similar to that of Darwin's case for natural selection.[5] With each uniformitarian success, Lyell's uniformitarian hypothesis is strengthened and its evidential advantage amplified, whereas support for special and even mysterious catastrophist processes is weakened.

5 That is not surprising since Lyell's work was an inspiration for Darwin.

5.3. Thomson's Cathode Rays

J.J. Thomson's paper in 1896 on "Cathode Rays" is celebrated as establishing that cathode rays consist of negatively charged particles, soon to be known as "electrons." Thomson, at the time, was embroiled in a debate with Philipp Lenard over the nature of these cathode rays. Thomson advocated for a particle account. Lenard defended the competing view that they are radiative, which then meant that they were waves propagating in the ether. Lenard had argued against a matter theory of cathode rays akin to Thomson's by noting that the rays persist even when the cathode ray tubes are completely evacuated. That is, there is no matter in the tubes to comprise the rays. Only ether remains. The rays, he concluded, had to be processes in the ether. Thomson's analysis depended on his experimental results that cathode rays are deflected by magnetic and electric fields exactly as if they are charged particles in rapid motion. Lenard struggled to accommodate these items of evidence in his account of ether. He could only speculate that Thomson's magnetic field had somehow disturbed the ether so that the rays would bend. This vagueness further weakened his retreating theory.

Thomson pressed his advantage with a coup de grâce. Waves in the ether bend because their velocity varies from place to place. This is how light is refracted by media of differing optical densities. A uniform magnetic field would disturb the ether in the same way in every place. Thus, elementary wave optics precludes it from bending cathode rays. However, uniform magnetic fields do bend the rays. Thus, the evidence that gave strong support to Thomson's particle theory is the same evidence that undid Lenard's ether wave theory.

The evidential advantage of Thomson's hypothesis is amplified by its accommodation of further evidence. For example, a metal vessel catching cathode rays becomes negatively charged, as one would expect if the rays are streams of negatively charged particles. An ether wave theorist might seek to dismiss this as an accidental artifact of the experimental arrangement. That escape ceases to be plausible once the charged particle hypothesis has an evidential advantage.

5.4. Einstein and the Anomalous Motion of Mercury

In November 1915, an exhausted Einstein was putting the finishing touches to his general theory of relativity. In that month, he found to his great joy that his new theory accounted exactly for a long-standing anomaly in the orbit of Mercury that so far had resisted explanation. His theory's success with Mercury was immediately recognized as an evidential triumph. The history does not follow the pattern of one theory gaining a slight evidential edge, which is then amplified. The accounts competing with Einstein's theory had all been discredited by the time of his completion of general relativity. However, if we consider the logical relations among the competing theories, independently of their order of emergence historically, then we see the same pattern of competition and amplification of slight evidential advantages.

The natural competitor to Einstein's theory is that the anomalous motion of Mercury arises from gravitational effects fully within Newtonian theory. It results from the perturbative effects of further, unrecognized matter. The "further matter" hypothesis has an initial advantage. It had become routine for astronomical anomalies to be resolved by the identification of further matter. For example, irregularities in the orbit of Uranus could be accounted for by the mass of a more distant, unrecognized planet. That led to the discovery of the planet Neptune. General relativity, however, is an exotic theory of extraordinary complexity mathematically. That it happens to return precisely the anomalous motion of Mercury is interesting. But it is hardly decisive evidence for the theory when standard Newtonian theory has a proven track record of accommodating just such anomalies by prosaic means.

However, these prosaic means falter. The various formulations of the favored, further matter hypothesis successively fail when evidence capable of separating the competing formulations is accommodated. If the further matter was located in a planet, "Vulcan," then its position was calculable, but no planet was observed there. Further possibilities located the matter in a slightly flattened Sun or in a dispersed cloud of matter surrounding the Sun that produces the zodiacal light. Neither proved to be viable. With each failure of the further matter hypothesis, the fortunes of Einstein's theory rose. Another possibility was an adjustment to the exponent in Newton's inverse square law of gravity. Although that exponent can be adjusted to accommodate the anomalous motion of Mercury, it fails to fit well with the motions of the remaining planets. Einstein's theory, however, has no adjustable parameters. It cannot accommodate any other motion of Mercury. Seen against this accumulation of failures of competitors, Einstein's theory rises as the only viable alternative.

5.5. Big Bang and Steady State Cosmology

In the mid-twentieth century, the prominent decision for cosmology was between the big bang and steady state theories. Later textbook accounts point to Penzias and Wilson's announcement in 1965 of their discovery of cosmic background radiation. It was, they said, the observational fact that confirmed the big bang theory and refuted the steady state theory. We are led to imagine the competition as ending abruptly.

That is not what happened. There was no immediate decision favoring big bang cosmology. It did gain a small advantage since the big bang cosmologists of the time — notably Dicke's group at Princeton University — had predicted something like it. However, the big bang cosmologists of the 1960s were reluctant to claim a definitive victory in print and with good reason, for the import of the evidence was still equivocal. Rather, it took roughly three decades for the decision between the two to be definitive.

Three developments were needed during these decades. First, considerably more observational work was needed. We now report Penzias and Wilson as observing thermal radiation of a cosmic origin of 2.7K. However, to affirm that a radiation field is thermal requires measurements across the spectrum. Penzias and Wilson had only measured one wavelength, 7.4cm. Many more measurements were needed and in fact were undertaken in the decades following. The incontrovertible evidence of a thermal spectrum was provided by NASA's COBE satellite in 1989.

Second, big bang cosmology needed to establish that it did indeed predict such thermal radiation. This required the development of precise cosmological models. In them, the radiation that we now measure is the remnant of radiation in a hot early universe that decoupled from matter when the cosmic fireball had cooled to 3,000K. That decoupled radiation is cooled to 2.7K by the expansion of the universe. Many components of this big bang account have to work correctly. The most troublesome is establishing that the early cosmic fireball was an equilibrium thermal system to which a temperature can be assigned in the first place. One could simply assume thermal equilibrium from the outset. It would be better, however, if cosmic processes in the early universe would produce this equilibrium. That was precluded in the cosmological models popular in the 1960s and 1970s by the so-called horizon problem. It showed that matter in those models was expanding so fast that it could not interact enough to achieve thermal equilibrium. The standard solution has been to invoke an early inflationary phase in the expansion of the universe.

The ready acceptance of this inflationary account illustrates the amplification of earlier successes. Until a big bang cosmology has some strong support, the inflationary addition is merely a speculative supplement to an already speculative theory. Once the big bang dynamics is supported, however, an inflationary phase is easy to accept as its natural completion.[6]

Third, it needed to be shown that steady state cosmology cannot accommodate the cosmic background radiation. This is by no means obvious, for thermal radiation can be acquired cheaply by theorists. All they need is some system to come to thermal equilibrium. Steady state theorists sought this through various avenues. One was that there is a slight opacity to space itself. Radiation from the continuous process of creation of steady state cosmology would be absorbed and reradiated through this slight opacity, thereby arriving at a thermal equilibrium. This proposal failed since the amount of opacity needed would be too great to allow observation of distant radio sources. Other efforts by steady state theorists, such as iron whiskers to thermalize starlight, also failed. This illustrates how an evidentially disadvantaged theory is further weakened by the need for successively more far-fetched repairs.

These three developments led to the decision in favor of big bang cosmology. That decision came slowly. Big bang cosmology enjoyed only a slight advantage at the outset. It grew steadily as observational results and theoretical developments favored it while efforts by steady state theorists to accommodate the same evidence faltered.

5.6. Arp and Bahcall on the Origin of Galactic Red Shifts

While the publicly more visible debate between big bang and steady state cosmologies proceeded, a narrower, less visible debate unfolded among astrophysicists and astronomers on the observational foundations of these cosmologies. Both big bang and steady state cosmologies assumed an expansion of the universe. Its evidential support lay in the finding by astronomers, starting most prominently with Hubble in 1929, that the galaxies are receding from our galaxy with a velocity that, on average, increases linearly with their

6 However, doubts linger about whether a period of inflation really does solve the horizon problem or whether it merely relocates it into the need to fine-tune initial conditions in a still earlier phase of cosmic expansion.

distances from our galaxy. (Hubble's analysis in 1929 is the subject of Chapter 7, "The Recession of the Nebulae.") An inference to a distance-dependent velocity of recession proceeded from the observation that light from the galaxies is uniformly shifted to the red end of the spectrum, with the shift increasing linearly with distance. This red shift was interpreted as deriving from a velocity of recession.

That the red shift in a galaxy's light resulted from its velocity of recession was disputed energetically by Halton Arp, a well-established astronomer. His case against this association grew in the 1960s and was regarded as sufficiently serious to merit a direct confrontation at the meeting of the American Association for the Advancement of Science (AAAS) on December 30, 1972, in Washington, DC. There Halton Arp faced John Bahcall, an astronomer at the Institute for Advanced Study in Princeton, New Jersey, there to defend the standard view.

I need not here rehearse the details of the debate. I have recounted them elsewhere (Norton, 2023) and refer readers to this source for elaborations. What matters for my purposes here is that the confrontation of Arp and Bahcall provides another illustration of the unstable dynamics of competition among theories. Is the red shift of light from galaxies the result of their motion of recession, as Bahcall affirmed? Or is it the result of some other source, as Arp argued? Each laid out his case.

Bahcall based his case on the evidence, available in multiple forms, that the red shift of light from the galaxies varies roughly linearly with the distances to those galaxies. Establishing that linear dependence was his major concern. The connection to a velocity of recession was provided by the then favored expanding universe cosmologies: they all required a linear relation between the velocity of recession of a galaxy in our vicinity and its distance from us.

Arp's case depended on his own extensive observations of galaxies. Arp had amassed an extensive collection of cases of galaxies that appeared to be physically connected but had very different red shifts. A physical connection would mean that the associated galaxies must be at roughly the same distance from us. Their marked differences in red shift could not then derive from a linear dependence of red shift on distance.

The two views in competition were sufficiently strong to merit serious examination at the AAAS meeting in 1972. However, the competition was unstable. Bahcall's view was already the recognized view. As his position strengthened subsequently, Arp's dissenting view was correspondingly weakened.

We can trace this instability in the competition in three areas. The first area was new astronomical data, which continued to conform to Bahcall's view. Arp's view, however, was weakened by investigations indicating that the physical associations so central to his case were merely fortuitous alignments in our sky of objects separated by great distances.

The second area was the connection to cosmology. Bahcall's view conformed to then standard cosmologies. If one applies general relativity to the sorts of matter distributions observed by the astronomers, then a dynamic cosmology ensues. It might be contracting or expanding. However, a static universe, such as Einstein had originally proposed in 1917 and Bahcall needed, was unstable and thus precluded.

Just as Bahcall's view was supported by then standard cosmology, so too his view of the linear dependence of red shift and distance provided support for the cosmology. It was the observational basis of the expansion of the universe. The outcome was a magnification of his evidential advantage. His evidential success strengthened support for expanding universe cosmologies, and their strengthened support then further enhanced his position.

Arp's view, however, found no support in existing cosmology. If the red shift was not derived from a velocity of recession, then the ensuing cosmology was one of an overall static mass distribution that lay outside standard cosmology. To preserve the viability of his critique, Arp needed to presume a static cosmology for which there was no real independent support. The evidential processes that enhanced support for Bahcall's view simultaneously weakened support for Arp's view.

The third area in which the instability manifested was in the physical basis of the red shift. Bahcall's standard view could employ a simple basis, ready to hand. The velocity of recession of galaxies in an expanding universe cosmology led directly to it. With that source precluded, Arp had no correspondingly established physics from which to derive the red shifts. He resorted briefly to speculation, such as "tired light."

Quasars proved to be a decisive test. They are luminous bodies with great red shifts. In the standard view, they must be very distant from us and thus have enormous intrinsic luminosity. Initially, in this view, it was hard to explain the enormous energies that it supposed for these bodies. Arp's alternative was that they are merely nearby objects, highly red shifted, but not of such great intrinsic luminosity. Quasars were subsequently identified as the enormously energetic nuclei of a galaxy, likely holding a supermassive black hole. Once again the evidential success of Bahcall's standard view was magnified. The view supported the immense energy and distance of quasars, and establishing a physical basis for their immense energy then enhanced support for Bahcall's standard view. Arp, however, was unable to provide a cogent physical basis for the high red shift of quasars if they are assumed to be nearby objects.

As Bahcall's standard view went from strength to strength, Arp's dissident view faltered and was dropped from serious consideration.

5.7. More Illustrations

Chapter 9 of *The Material Theory of Induction* (Norton 2021) recounts two more competitions: oxygen versus phlogiston theories in the late eighteenth century and corpuscular versus wave theories of light in the nineteenth century. The details of their competition are too involved to admit compact summaries. I can extract one result, however.

At the crudest level, oxygen theory prevailed over phlogiston theory when Lavoisier's experiments required that oxygen must be attributed to a conserved weight. Phlogiston theory faltered since the same experiments required that phlogiston be attributed to a dubious negative weight, levity. Similarly, a major factor in the decision on theories of light came with Fizeau and Foucault's measurements of the speed of light in air and water. The corpuscular theory required the speed to increase in a denser medium, whereas the wave theory required it to decrease in a denser medium. The experiments found a decrease in the speed.

We see here that theories in competition are responsible to the same experiments and that careful exploration can find experiments that only one of the theories can accommodate. Although we might doubt that just one experiment can be decisive, that responsibility still plays a major role in the dynamics that leads one theory to prevail over its competitors.

6. Unconceived Alternatives

The instability illustrated in these examples arises in the competition between two theories. Is that enough to make the case? Might we worry that there is a third, fourth, or fifth, as yet unimagined or unarticulated, theory lurking in the wings, such that evidence cannot separate one of them from our favorite theory? The possibility of such further theories has been defended, notably by Stanford (2006), as "unconceived alternatives."

They do not provide the sort of threat that one might imagine. They open the possibility that our current best theory might not be the one best supported by the evidence. That is not the question here. The question is whether the best supported theory is unique. That can be the case even when the theory that we happen to favor most strongly is not the best supported.

For unconceived alternatives to challenge uniqueness, they must provide us with a theory challenging our favored theory that is equally well supported, assuming that our favored theory is the best supported given the evidence, or the challenger theory must provide us with two unconceived alternatives equally well supported and still better supported than our favored theory.

The analysis already given indicates that such an achievement lies beyond what unconceived alternatives can supply. As long as these alternative theories differ in some factual claim, their difference must be open to adjudication by observation and experiment, even if that adjudication might not be practical immediately. Otherwise, their differences lie outside empirical science.[7]

7. The Underdetermination Conjecture

If one seeks literature to contradict this chapter's claim of uniqueness, the natural reference is the "underdetermination thesis."[8] Loosely speaking, it asserts that no body of evidence, no matter how extensive, can pick out a theory uniquely as the one best supported inductively. The thesis is then used to advance the tendentious claim that our commitment to any theory, even those of the most mature sciences, always relies on the addition of other factors, possibly social, psychological, pragmatic, or conspiratorial. The thesis is mislabeled as a "thesis" insofar as theses are commonly taken to be propositions for which we have good evidence. It is, as I will now argue, merely a conjecture that has never secured proper support. It can be stated for my purposes here as

7 The closest that the literature can provide, for these theories, balanced perfectly evidentially, arise as the observationally equivalent theories used to support the underdetermination thesis. In Section 7, I explain why these examples fail in their purposes.

8 For an introduction, see Stanford (2017).

> *Underdetermination conjecture*: any body of empirical evidence, no matter how extensive, will provide inductive support for multiple, mutually exclusive sets of propositions such that no one set is distinguished as enjoying the strongest support.

This conjecture should be distinguished from the weak, de facto claim that at some definite moment the extant evidence for a theory might fail to determine it. This circumstance commonly arises in newly emerging sciences. If the science matures, then it is merely a transient shortcoming. Otherwise, it is not.

The full conjecture is remarkably strong in its pessimism. It applies to all bodies of evidence and theory. Thus, it is astonishing that the conjecture has never advanced beyond what for many is merely a comfortable hunch. For them, the conjecture seems to be plausible and welcome. If one is inclined to it, then easy but inadequate examples might be enough motivation. The evidence can tell us of a correlation between children who watch cartoons and children who behave violently in the playground.[9] That evidence leaves undetermined which causes which, or if there is a common cause of both, or if the correlation itself is mere happenstance. The example merely illustrates de facto underdetermination. Randomized control trials can decide among the possibilities.

Once it has been mentioned enough in the literature, the plausibility of the conjecture for some makes it easy to lose sight of the fact that there is no cogent demonstration of the conjecture. The arguments offered in favor of the underdetermination conjecture have been subjected to repeated analysis and have failed scrutiny. The arguments can be shown to neglect much of the existing work in inductive inference and to make dubious claims concerning observationally equivalent theories. Laudan and Leplin (1991) and Norton (2008) explore these failures, too extensive to be developed in detail here.

The simplest and most common demonstration of the conjecture rests on an inadequate account of inductive inference. A single body of empirical evidence can be entailed by many different sets of hypotheses, with suitable boundary conditions and auxiliary assumptions. With a naive hypothetico-deductive account of confirmation, it would then follow that they are all equally well supported inductively. This naive account has long been

9 This example is from the opening paragraph of Stanford (2017).

subjected to criticism from many perspectives. Consider the standard geological and evolutionary account of the origin of fossils. Compare it with a revisionary theory claiming that the Earth and its rock strata were created five minutes ago, complete with an intact fossil record. Since both entail the same evidence, we would have to say that both are equally well supported. The standard response in the literature is sketched in Section 5, "Hypothetical Induction," in Chapter 1. It is that bare hypothetico-deductive confirmation must be supplemented by further conditions to enable discrimination in such cases. We might be told, for example, to assign greater support to the more explanatory hypothesis or to the simpler one.

Within the material theory of induction, merely entailing the evidence by itself does not confer inductive support for a hypothesis or theory. The entailment must happen in the right way: each of the parts of the propositions in the theory must be supported inductively in accordance with the requirements of the material theory. The supposition that the creation occurred exactly five minutes ago, as opposed to ten or fifteen minutes or a millennium ago, must be supported. The revisionary theory can provide no discriminating evidence. In comparison, standard geology does provide extensive evidence for its chronology of the formation of the Earth.

The transition from hypotheses that merely entail the evidence to an evidentially well-supported body of propositions is difficult and can take a long time. We see in Chapter 12, "The Use of Hypotheses in Determining Distances in Our Planetary System," that, in spite of sustained and ingenious efforts, a system of orbital sizes for the planets of our solar system was not firmly established until the eighteenth and nineteenth centuries. Indeed, at the most general level, the nature of inductive inference is sufficiently irregular, according to the material theory of induction, that there can be no sufficiently expansive framework sufficiently precise to admit a cogent demonstration of the conjecture.

After I drafted this chapter, Sam Mitchell sent me his book published in 2020. It also seeks to undo the skepticism concerning the reach of evidence associated with Duhem and Quine. His concern is specifically to respond to the claim that the import of evidence is always holistic. We cannot be assured that contradicting evidence refutes any specific hypothesis, the inductive pessimists insist. They suppose that any such judgment requires auxiliary hypotheses that can be the real culprits in the contradiction. Mitchell disagrees.

His analysis agrees on many points with the one developed here and is most welcome.

8. Observationally Equivalent Theories

Theories that have the same observable consequences are frequently displayed in the literature on the underdetermination thesis as "observationally equivalent theories" or "empirically equivalent theories."[10] They serve to illustrate the underdetermination thesis since, it is asserted erroneously, no evidence can favor one over the other, and they are used in an attempt to make the case for the underdetermination thesis.

Do these observationally equivalent theories pose a threat to the uniqueness urged in this chapter? Here I will recount briefly why they do not. I will use a simple example of a pair of observationally equivalent theories. For a more expansive inventory of examples and for more detailed, critical analysis of the underdetermination thesis along the lines below, see Norton (2008).

In the early seventeenth century, purely astronomical observations of the relative positions of the Sun, Moon, and planets could not discriminate two systems. The first was the familiar Copernican heliocentric system. The second was the Tychonic geocentric system. The observational equivalence followed assuredly from the simple fact that the Tychonic system could be generated merely by relocating the point of rest in the Copernican system from the Sun to the Earth but otherwise preserving all relative motions.

This example and the others like it fail to sustain any interesting conclusions about the limited reach of evidence for two reasons. First, if the competing theories differ in something factual, then the empirical character of science requires that the difference should manifest in something observable. The Copernican and Tychonic systems differ in whether the Earth or the Sun is at rest. Purely astronomical facts about the relative positions of the Sun, Moon, Earth, and other planets cannot decide, for they provide no notion of rest. They can be separated, however, if we ask after the physical forces acting among the bodies of the solar system. Newton's later physics distinguished bodies moving inertially from those that accelerate. Inertial motion becomes

10 Here I resist this latter expression because of its vagueness. If two theories have identical observational consequences, it does not follow that they are supported equally by observations. That is, one can still be favored empirically over the other, as I argued in the preceding section.

the Newtonian surrogate for rest. At most, either the Earth or the Sun can be in inertial motion. When we seek the gravitational forces acting between the bodies of the solar system, that body must be the Sun and not the Earth. We decide in favor of the Copernican system.[11]

This decision was possible because subsequent investigations in a broader domain, that of gravitational physics, provided the further evidence needed to separate the systems. This possibility remains for every case of observationally equivalent theories. Insofar as they differ on anything factual and they lie within empirical science, we cannot preclude new evidence separating them. Indeed, we should expect determined investigators to find such evidence.[12]

Second, if we set aside the possibility of new evidence, then there is a second failing of all the cases of observationally equivalent theories in the literature. If the case is to be presented in the literature, then it must be possible to demonstrate in the confines of tractable publication that the two theories really are observationally equivalent. For example, there is a simple recipe for converting the Copernican system into the Tychonic system. We take the motions of the Copernican system and simply subtract vectorially from them the motion of the Earth. The result is a system of motions with the Earth at rest but agreeing with the Copernican system in all relative motions.

When such a translation is available, we cannot preclude the possibility that the two theories do not differ in anything factual. Rather, they are merely different presentations of the same theory. If we restrict considerations only to the relative positions of bodies in the solar system, then this is the case for the Copernican and Tychonic systems. They differ only in the designation of which body is at rest. But that designation lies outside the body of facts pertinent to our restricted domain. It is, as far as they are concerned, merely an empty stipulation.

---

11 As an exercise, one might like to contemplate whether some distribution of masses might enable the Tychonic system to conform to Newtonian gravitation theory. One would require, for example, that the Earth must be much more massive than the Sun so that the Sun orbits the Earth and not vice versa. We can then no longer account for the motion of Venus, whose maximum elongation from the Sun is between 45° and 47°. It would be pulled out of its orbit around the Sun by the far greater attraction of the Earth or, failing that, display significant perturbations because of the Earth's attraction.

12 Here the historical sciences might provide an exception. The totality of evidence recoverable from some archaeological site, for example, might leave questions about the site unanswered. The failure is not the result of a lack of power of inductive inference but merely the paucity of evidence.

This possibility threatens all cases of observationally equivalent theories. That they can be interconverted opens the possibility that they are merely the same theory. They differ only in their descriptions and in superfluous posits of no factual import. It is possible and sometimes enticing to mistake these posits as having factual import, even though they manifest in nothing observable. The most familiar example in real science concerns a suitably refined version of Lorentz's ether-based electrodynamics and the relativistic electrodynamics that Einstein introduced in 1905 with his special theory of relativity. The two are observationally equivalent, and, as far as experimentation was concerned in the first decade of the twentieth century, they were treated as the same theory. However, Lorentz insisted that the ether factually has a state of rest, contrary to Einstein's principle of relativity. The difficulty was that nothing observable — no experiment — could determine just which of the infinity of inertial states of motion was that ether state of rest. The mainstream of physics soon came to discount the ether state of rest as fictional.

9. Formal Accounts

Since the material theory of induction can meet the challenge, it is proper to ask whether formal accounts of inductive inference can also meet it. They do not do well with it and for reasons associated directly with their formal character.

First, as I argued at some length in *The Material Theory of Induction* (Norton 2021), the rules of various formal systems are poorly articulated, and an ambiguity in their import is inevitable. Consider, for example, the use of arguments by analogy to infer the properties of light. Light is analogous to sound in that both have a wave character. The pitch of sound is analogous to the color of light. However, sound needs a medium in which to propagate, the air, analogous to the discredited nineteenth-century luminiferous ether. This difficulty does not arise in a different analogy. In it, light is taken as analogous to rapidly moving corpuscles. Then light, like corpuscles, can propagate in vacuo without the support of a medium. Yet the corpuscles of the nineteenth-century and earlier theories have no wave-like properties. Just how are we to weigh the conflicting successes and failures of these different analogies? The general rules in the literature are too vague and hedged to give us a definite answer.[13]

Second, there are multiple formal schemes for inductive inference and no clear guides to which to use in any application. Take, for example, argument by analogy and inference to the best explanation. Neither of the analogies of light to sound and light to rapidly moving corpuscles recovers the phenomenon of light polarization. Sound waves are longitudinal, whereas polarization derives from the transverse character of light waves. That is, neither of the familiar analogies provides an explanation of polarization. Rather, the best explanation of polarization is that light is not analogous to either sound or corpuscles.[14]

Which formal scheme should be applied where? In particular cases, we might use prudence to decide and have things work out tolerably well. However, we do so in the absence of unambiguous metalogical rules.

Third, the Bayesians are confident that they have a solution. Their scheme, they believe, embraces and explains all others and can recover uniqueness through various limit theorems. This confidence can be sustained only as long as they ignore the enduring and insoluble problem of the priors. The Bayesian system is not and cannot be self-contained. The selection of prior probabilities must be made outside the normal processes of conditionalization by Bayes' theorem. Yet these priors can be so selected as to protect almost any bias. The simplest illustration arises when we have two theories T_1 and T_2 that deductively entail the same evidence E. Then we have equal likelihoods: $P(E|T_1) = P(E|T_2) = 1$. An application of Bayes' theorem then tells us that

$$P(T_1|E) \,/\, P(T_2|E) = P(T_1) \,/\, P(T_2)$$

That is, our comparative assessment of the relative support afforded to the two theories by the evidence, the ratio of posterior probabilities $P(T_1|E) / P(T_2|E)$, is determined entirely by whatever external judgments led us to the ratio of prior probabilities $P(T_1) / P(T_2)$. Bayesians face an unwelcome dilemma: either set these priors arbitrarily so that the final judgment is arbitrary, or seek

13 See Chapter 2 of this text and Chapter 4, "Analogy," in *The Material Theory of Induction* (Norton 2021).

14 Might we try the analogy to waves propagating along a flexible rope since they are waves of transverse displacement? This analogy fails to recover the behavior of polarized light in polarizing filters. The best explanation of the behavior is that, when it comes to polarizing filters, light is not analogous to waves on a flexible rope.

guidance from other accounts of inductive inference. This problem troubles all formal calculi of inductive inference, or so I have argued in Chapter 12, "No Place to Stand: The Incompleteness of All Calculi of Inductive Inference," in *The Material Theory of Induction* (Norton 2021). None can be self-contained but can only return nontrivial results insofar as nontrivial inductive content is introduced from outside the scope of the calculus.

It is fortunate that scientists do not try to conform their judgments of inductive support algorithmically to these conflicting and ambiguous formal schemes, for that would induce inductive anarchy.

10. Conclusion

In this chapter, I have sought to establish that the threat of multiple and equally well-supported systems of inductive inference has been parried. The escape derives from the empirical character of science. Competing systems of inductive logic derive their competing factual warrants from different theories within science. When these warranting facts differ, their differences must manifest in something accessible to possible observation, or they lie outside empirical science. When the pertinent observations are secured, they will strengthen one of the theories while weakening its competitors.

This escape is enhanced by the close integration of the facts of a science and its relations of inductive support, asserted by the material theory of induction. The integration promotes a positive feedback dynamic that accelerates the strengthening of one system of relations of support at the expense of its competitors. As more of the factual claims of a science are sustained by the evidence, the growing body of supported fact authorizes stronger inductive inferences within the domain of the science. That in turn leads to inductive support for still further facts. As one theory ascends, even if haltingly, its competitors will fall. When sufficient evidence is available, the accumulation of these processes will lead to the dominance of one science and its associated relations of inductive support while its competitors are eliminated. The uniqueness and inductive solidity of mature sciences in their domains are expected and explained.

REFERENCES

Boyd, Nora. 2018a. "Scientific Progress at the Boundaries of Experience." PhD diss., University of Pittsburgh.

———. 2018b. "Evidence Enriched." *Philosophy of Science* 85: 403–21.

Kuhn, Thomas S. 1996. *The Structure of Scientific Revolutions*. 3rd ed. Chicago: University of Chicago Press.

Laudan, Larry, and Jarrett Leplin. 1991. "Empirical Equivalence and Underdetermination." *Journal of Philosophy* 88: 449–72.

Mitchell, Sam. 2020. *Fault Tracing: Against Quine-Duhem*. Berlin: De Gruyter.

Norton, John D. 2008. "Must Evidence Underdetermine Theory?" In *The Challenge of the Social and the Pressure of Practice: Science and Values Revisited*, edited by Martin Carrier, Don Howard, and Janet Kourany, 17–44. Pittsburgh: University of Pittsburgh Press.

———. 2021. *The Material Theory of Induction*. BSPS Open Series. Calgary: University of Calgary Press.

———. 2023. "Inductive Inferences on Galactic Redshift, Understood Materially." In *Current Debates in Philosophy of Science*, edited by C. Soto, Synthese Library 477, 227–46. Cham, Switzerland: Springer.

Stanford, Kyle. 2006. *Exceeding Our Grasp: Science, History, and the Problem of Unconceived Alternatives*. New York: Oxford University Press.

———. 2017. "Underdetermination of Scientific Theory." In *The Stanford Encyclopedia of Philosophy,* winter 2017 ed., edited by Edward N. Zalta. https://plato.stanford.edu/archives/win2017/entries/scientific-underdetermination/.

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