This is a slightly revised version of a paper presented at the 2007 Eastern APA in the Symposium on Hermeneutics and Scientific Theory Choice.

Interpretation and Empirical Constraint in the Development of Science

Harold I. Brown

www.niu.edu/~hbrown
hibrown@niu.edu

The transformation in our understanding of science that has been developing since the late 1950s recognizes a major role for interpretation in the process of scientific research and in evaluating the outcomes of that research. Logical empiricism left little scope for interpretation in science. The differing view of interpretation is particularly salient when we look at the interaction between scientific theories and empirical evidence. For logical empiricists, uninterpreted sensory experience provides both the semantic and epistemic foundations of science. On the semantic side, terms of our language that can be directly associated with sense experience were considered non-problematic and were taken to constitute our basic vocabulary. The meanings of other terms must be established by means of some appropriate relation to this basic vocabulary—a thesis that is particularly problematic for terms introduced to refer to items we cannot detect with our senses.[1] On the epistemic side, all propositions whose content goes beyond a summary of our sensory experience must be justified on the basis of sensory experience.

In several respects, the new approach reversed the relation between sensory evidence and theory by attributing a primary role to theories. The meanings of theoretical terms, it was suggested, are determined by implicational relations among the terms in a theory, and sensory evidence does not play a role in science until it has been interpreted in terms of a theoretical language. The new view of meaning was not well developed by the best-known exponents of this new approach and, due to time constraints, I

---

[1] See Brown 2007, Sec. 3.5 for discussion of changing views of this relation among logical empiricists.

will not pursue this topic here.[2] Instead, I will focus on a key consequence of this view of meaning: that empirical evidence is, somehow, theory-laden. I will argue that while, properly understood, this thesis is important and correct, it is only half of the story. In science, theory-laden evidence always involves interpretation of some input that is independent of the theory in question and that limits the range of admissible interpretations. An adequate understanding of how science develops, and of the epistemic significance of science, require that we grasp the interplay between interpretation and constraint. Moreover, I will argue, a central and characteristic feature of science is the continual attempt to increase the range of constraints on our interpretations.

When we consider the literature on theory-laden evidence we find that several quite different theses are often lumped together under this rubric; I will consider five of these. The strongest version of this thesis is that our beliefs are so deeply implicated in what we perceive, that we end up perceiving only what we expect to perceive. As a result, the story goes, scientists committed to a theory fail to notice anomalies that would challenge that theory. This seems to be the claim that Kuhn is defending in his discussion of the anomalous-card experiment in which subjects who are given brief glimpses of unusual cards report seeing a normal card. For example, a red ace in the shape of a spade is described as an ace of spades (Bruner and Postman 1949). Yet Kuhn's use of this experiment as a "schema for the process of scientific discovery" (1962: 64; henceforth this book will be cited as *Structure*) is problematic in several

---

[2] I have pursued it in considerable detail in Brown 2007 building on work of Wilfrid Sellars that pre-dates the developments in philosophy of science that I am discussing. See especially Sellars 1953, the essays "Empiricism and the Philosophy of Mind," "Is There a Synthetic A Priori?" and "Reflections on Language Games," in his 1963a; for a later statement see Sellars 1974. C. I. Lewis provides an earlier account of meaning that is completely divorced from any dependence on sensory experience—especially his 1956, originally published in 1929.

respects.[3] Note first that a brief glimpse of a carefully contrived artefact hardly provides a good model for scientific observation. Indeed, given a bit more time, most subjects of the experiment correctly reported the item before them. Second, if one seeks to use this experiment as a model of how scientists collect evidence, we should look to the activities of the experimenters, not their subjects. The experimenters were not limited to glimpsing items that other experimenters had contrived in order to challenge familiar expectations. Third, the suggestion that scientists fail to notice anomalies runs contrary to the overwhelming body of Kuhn's discussion. One of Kuhn's most important theses is that anomalies are not always counter-examples, but his arguments for this thesis require that scientists recognize anomalous outcomes. Kuhn describes many such cases and explores various ways in which scientists deal with anomalies. Thus, while failures to notice what does not conform to our expectations no doubt occurs, it is hardly a basis for understanding evidence collection in science. Let us move on to some more viable respects in which evidence may be theory-dependent.

A more moderate version holds that our concepts and beliefs play a role in the constitution of our sensory experience, but also recognizes the existence of some other element that limits the range of our interpretations. To take a mundane example, there are many ways in which I can interpret the water in my glass, but I cannot turn it into cognac or prevent the glass from falling if I let it go, just by changing my interpretation. We might interpret the fall as motion to a natural place, or as a response to a gravitational force, or as motion on a local geodesic, but the glass still falls. Thinking differently will not turn it into a ballerina doing pirouettes. In other words, theory-laden perception is not theory-generated perception. I will return shortly to the nature and scope of these limitations as well as to another point: when I say that the glass falls I am using everyday English that we can understand independently of our current scientific beliefs and our knowledge of the history of science.

---

[3] See Brown 2005 for a more detailed discussion.

According to a third version of the theory-dependence thesis, in order for observational evidence to be relevant to the evaluation of a theory, that evidence must be described using the language of that theory. I think this claim is correct, but its significance must not be exaggerated. The need to describe evidence using the concepts of a theory provides no impediment to recognizing and reporting evidence that either supports or challenges that theory. Moreover, when the need arises—such as when advocates of competing theories are disputing the significance of an observational result—it is generally possible to describe that result without adopting the language of the theories in question. Galileo, for example, did this when he discussed whether a stone dropped from the mast of a moving ship would land at the foot of the mast or towards the rear of the ship. Such description does not require the use of a theory-free observation language. It requires only that our overall linguistic resources allow for a description that does not presuppose any of the competing theories. All of us come into a scientific situation with a body linguistic resources that goes well beyond what is needed to discuss the case at hand. It will be useful for us to have some terminology for describing such cases. Adapting a suggestion from Basu (2003), I will distinguish between *data* and *evidence*. *Evidence* is used to evaluate theories and must be characterized using the language of the theory in question. *Data*, as I will use the term, is a description of the result of some interaction with nature that does not use the language of any theory *under evaluation.* In cases of theory comparison, proponents of competing theories will often be able to agree on the data even if they disagree on its role as evidence for or against a particular theory. Even when there is dispute over whether a particular outcome occurred, sufficient language will typically be available to allow the disputants to clearly state what it is that they are disagreeing about.

A fourth version of theory-dependence holds that which observations scientists undertake are determined by accepted theory. This is surely correct and important. The range of items we can explore is too large to be pursued without some guidance. If I set out to describe the room we are in without some notion of what is worth describing, I would not finish in my lifetime. But once we decide what we are going to look at, our expectations do not pre-determine what we will find. The history of science reminds us that the persistent pursuit of data often turns up results that challenge the very theory that led us to seek this data in the first place. As Kuhn frequently suggests, working within a framework sharpens one's sensitivity to anomalies (*e.g.*, *Structure* 24, 52, 64-65). In addition, the pursuit of data sometimes provides unexpected results that cannot be accommodated by any existing theory. Classic examples include the unanticipated discovery of sperm in the seventeenth century, and the successive discoveries of X rays in 1895 and radioactivity the following year. Encounters of this sort provide one major impetus for generating new concepts and new theories. Even outside of science nature often impinges on us in ways that we did not anticipate and that we would prefer to avoid; earthquakes provide a particularly dramatic example. The same holds in human interactions. For example, societies have been disrupted by incursions from other societies that have a superior military technology, and this occurs even though the losers in these encounters previously believed that no such societies existed.

My final version of theory-dependence holds that evidence gathered to evaluate a comprehensive theory presupposes that very theory in a way that prevents a fair test of that theory. As a result, it is claimed, empirical tests of major theories are viciously circular. But we have already noted that this is not true; scientists working within a theory regularly recognize situations that pose a challenge to that theory. Early astronomers recognized that the motions of the planets challenge the thesis that all celestial motions are circular; Aristotle recognized that projectile motion poses a problem for his account of terrestrial motion. Further examples can be amassed at great length and Kuhn's *Structure* provides a convenient starting point for compiling such a list. Moreover, I have argued elsewhere (1993, 1994) that even when a complex evidence-gathering procedure presupposes a theory in a central way, it may still be possible for the outcome of the procedure to contradict that very theory. Whether this can occur depends on the details of the procedure and it is a mistake to declare that a circle has been found without examining these details. Shogenji (2000) has shown that in many such cases the theory being evaluated is “presupposed” only in the way that we use a theory in setting up an hypothetico-deductive test or a *reductio ad absurdum* argument.

Skeptics about the significance of theory testing will often move, at this point, to the Duhem-Quine thesis and thereby bring in another way in which interpretation enters into scientific procedures. They note that all scientifically significant predictions and explanations of what we observe require multiple premises; when faced with a problematic outcome it is always possible, as a matter of logic, to protect a favored belief by challenging one or more different proposition in this set. However, those who invoke this thesis are already agreeing that we do encounter evidential results that contradict theory-generated expectations, and that this situation requires modification somewhere in the set of beliefs that generated the now-problematic result. In addition, the Duhem-Quine thesis goes hand-in-hand with the claim that our theories form an holistic structure in which changes in one member have consequences for other members of the set. But, as Greenwood (1990) has noted, changes anywhere in such a set may generate consequences for what we will observe in other situations. As long as we are committed to getting the observational results correct—and without this commitment we are not doing science—then we cannot make just any change we wish in order to accommodate unwanted evidence. In particular, the oft-repeated claim that any specific thesis can be protected come what may is seriously open to question. It would be interesting, for example, to explore the changes in the current body of science that would be needed to protect the thesis that the earth is at the center of the universe and made of different materials than the moon and planets. There is no *a priori* guarantee that this can be done given the full range of inter-locked beliefs evidence. While there is flexibility in our interpretation of the import of a body of evidence, this flexibility occurs only within limits set by powerful empirical constraints.

Quine, at one point, suggests that if sufficiently desperate we can reject an observation by declaring that the observer is hallucinating. But this would work only if evidence in science consisted of brief glimpses by epistemically isolated individuals, and we have already noted that this is not the case.

The retrograde motions of the planets, the motion of projectiles, the evidence for radioactivity, and many other examples, all involve persistent, intersubjectively available items. If these are hallucinations, they are persistent and wide-spread. On Quinean grounds, the claim that these are hallucinations is one more claim to be included in the web of belief and is subject to evaluation along with other such claims. The fact that scientists do not typically engage in such maneuvers supports the conclusion that such claims do not stand up to Quinean pragmatic evaluation.

To achieve a better understanding of the ways in which interactions with nature constrain scientific theorizing we must look at empirical evidence more carefully. It is especially important to wean ourselves from a central thesis of traditional empiricism: that empirical evidence consists of sensory experiences so that ultimately the range of evidence we can acquire is limited to the range of items we can sense. Kuhn, and many other exponents of historical philosophy of science, never made this step. They shifted from the view that empirical evidence consists of observations that are independent of any theory, to the view that observations are theory-laden. But they never doubted that there is a sensory core at the basis of all evidence and that this core delimits the scope of the evidence we can acquire. As a result, they failed to note that—as I shall argue shortly—the theory-dependence thesis points the way to a richer understanding of empirical evidence. The situation is somewhat ironic in the case of those who urge that philosophers look at what scientists actually do.[4] I want to note some of the lessons we should learn about evidence from engaging in this endeavor.

I will focus on the role that instruments play in the procedures by which scientists gather evidence. While the systematic use of instrumentation enters into scientific research with Galileo's use of the telescope, there is at least one even more striking instrument that entered into common usage at an earlier date: the magnetic compass, which allows us to see the direction of the earth's magnetic field even though we have no senses that respond to magnetism. One of the most striking features of the

---

[4] Shapere 1982 is a particularly important exception; see also Shapere 1977.

development of science is the postulation of items that we cannot detect with our senses. This was deeply troubling to logical empiricists. Hempel, for example, asserted that science aims at "establishing explanatory and predictive order among . . . the phenomena that can be 'directly observed' by us" (1965: 177), and then noted that:

> It is a remarkable fact, therefore, that the greatest advances in scientific systematization have not been accomplished by means of laws referring explicitly to *observables*, i.e., to things and events that are ascertainable by direct observation, but rather by means of laws that speak of various *hypothetical*, or *theoretical*, *entities*, i.e., presumptive objects, events, and attributes which cannot be perceived or otherwise directly observed by us. (1965: 177)

A very different attitude, one that is typical of naturalist realism, was expressed by Sellars: "It would be odd if the only qualitative dimensions of the world were those which are tied to the sensory centers of the human brain" (1963b: 149). Theory construction provides the means by which we endeavor to describe such items, and the kind of instrumentation that concerns me provides the means by which we attempt to interact with these items in order to test our theories about them.[5] The theory-guided attempt to interact with items that may exist in nature provides the central theme in the account of empirical evidence that I want to sketch.[6] The outcomes of such attempts, whether they are successful or unsuccessful, provide

---

[5] Although Galileo's telescope shows only items that are, under appropriate circumstances, observable by unaided perception, it shares one important feature with the kind of instrument that concerns me: it is inserted into the causal chain between the scientist and the items under study. Still, Galileo did not use his telescope to test claims about non-observables.

[6] For a more detailed development see Brown 1987 and the substantial corrections to that account in Brown 1995.

constraints on our theories. I think that this will make more sense if we have an example before us, so I want to outline the solar neutrino experiments that play a central role in an ongoing research project. I will begin with some historical background on neutrinos.

The existence of neutrinos was originally proposed in the early 1930s as a desperate attempt to solve a set of problems that arose in the study of beta decay, one of the three types of radioactivity. Beta decay occurs when a nucleus emits an electron, and one of the problems was that energy seemed to be missing, which would violate conservation of energy. Attempts to solve this problem included the suggestion by Bohr (and others) that energy conservation is a statistical effect and does not apply to individual interactions, Dirac's proposal that energy conservation is a low-velocity effect that breaks down at relativistic velocities, and Pauli's proposal that an additional undetected particle involved in beta decay carries the missing energy (and also solves the other problems). Pauli was extremely skeptical about his proposal, which he made in a letter that included several reasons why the proposal was probably wrong (Pais 1986: 315). Nevertheless, within a couple of years Fermi developed Pauli's proposal into a serious mathematical theory. There was, however, a key problem with Fermi's theory: there seemed to be no way to detect neutrinos, and thus no way to test the theory. The problem arose because neutrinos, as postulated, have no electric charge and no rest mass. In a classical framework this would mean that neutrinos cannot interact with other matter, and therefore that they could not trigger a detector. But these developments were taking place within the framework of quantum theory which implies that interactions between neutrinos and other matter are not impossible, only highly improbable. Given a combination of a sufficiently rich source of neutrinos and a sufficiently sensitive detector, neutrinos might be detected. These requirements were met over the next twenty years: nuclear reactors provided a rich source of neutrinos and during the 1950s Reines and Cowan succeed in detecting them (actually, anti-neutrinos). By 1956 physicists were confident that they could measure the rate at which neutrinos are produced by a reactor (Reines and Cowan 1953, Cowan *et al.* 1956). I now want to move to another scientific field.

By the early 1960s the problem of how stars—including our sun—generate power was considered to have been solved. Nevertheless, it was also recognized that the accepted account had an untested consequence: the processes that produce the energy should produce neutrinos at quite specific rates and with specific energies. Given the new-found confidence that neutrinos can be detected and counted, around 1962 Raymond Davis proposed an experiment that would measure the rate at which neutrinos produced in the sun arrive at the earth. Davis' detector has been described as "conceptually straightforward, but technologically very difficult" (Trimble and Reines 1973: 1). The heart of the detector is a large tank containing a common cleaning fluid, C_2Cl_4, about a mile underground at the bottom of a gold mine in South Dakota. The earth above the detector will filter out any particles from space other than neutrinos—although a small number of neutrinos, and other particles that can mimic the neutrino signal, will be produced by radioactive trace elements in the surrounding earth. Occasionally a neutrino will be captured by a chlorine nucleus which will be transformed into an argon atom. The argon that is produced will occur in the tank as a dissolved gas that can be periodically flushed from the tank (by a fairly complex set of chemical procedures). This gas is radioactive with a well-established half-life, so the amount produced in a given time period can be determined by recording its radioactive decays. I have skipped many details but the point I want to note is that the first published data (Davis, Harmer, and Hofman 1968) indicated that the neutrino flux was about one-third of the calculated amount. Something is clearly wrong somewhere. "The critical problem is to determine whether the discrepancy is due to faulty astronomy, faulty physics, or faulty chemistry" (Trimble and Reines 1973: 1). In other words, the question is whether the problem lies in the presumed theory of neutrinos, the theory of stellar-energy production that was explicitly being tested, or the chemistry involved in removing very small amounts of argon from the tank. There are other options as well since many initial conditions, such as the sun's composition, internal temperature, and magnetic fields, are involved in calculating a predicted neutrino flux.[7] Scientists involved in the experiment and in the ensuing discussions took these options very seriously. For example, the low neutrino account enhanced concerns about the role of radioactive elements in the rock surrounding the detector and led to systematic measurements at various depths, to an increase in shielding around the detector (the cavity was flooded with water), and to a redesign of the instrument used to determine the amount of argon produced.[8] The low neutrino count was sustained and led physicists to seek ways of overcoming limits inherent in Davis' experiment. In particular, by the time the experiment was done two different types of neutrinos had been identified, and a third type was discovered in 1975. The Davis experiment could detect only one of these types, electron neutrinos. Stellar theory also predicts that different reactions involved in stellar-energy production produce neutrinos of different energies. Davis' experiment could detect only the highest-energy electron neutrinos, considerably less than one-percent of those predicted. Scientists responded by developing new experiments that could detect neutrinos with a wider variety of energies, different types of neutrinos, the direction from which a particular neutrinos came, and more.

I will pause to underline some central points:

1. The researchers recognized the existence of an anomaly.
2. Without the body of theory that guided the design of the apparatus and the interpretation of the data, no anomaly would have been found.
3. While it is clear that something is wrong in the complex of theory and initial conditions that guide this experiment, there are many candidates for the location of the problem.

---

[7] In the early 1960s an experiment was proposed along the same lines as Davis' experiment but conceived of as determining the internal temperature of the sun.

[8] For detailed discussions see Bahcall 1990, Franklin 2001.

4. Whatever the eventual diagnosis of this problem, the experiment has yielded a body of outputs that provide constraints on future theories in this field.

5. While descriptions of the experimental results may make use of the language of the theories that are assumed in the design of the experiment—*e.g.*, the claim that a specific number of neutrinos was detected—this does not prevent the outcome from being recognized as a challenge to those very theories.

6. Our senses play only a tangential role in the collection and analysis of this data. To be sure, we depend on our senses to read the neutrino counters, but this is not at all what traditional empiricists were after when they advocated a fundamental role for our senses in the acquisition of knowledge. For example, the color, size, and shape of readings on digital counters are irrelevant to the role of this output as evidence for theory evaluation. It would not matter if the neutrino count were displayed in Roman numerals, or channeled through a voice synthesizer and detected aurally, or printed out in Braille. Our senses play a pragmatic role since this is the route by which information finds its way into our brains, but the items we can sense have no foundational role in this process.[9]

---

[9] See note 12 for a possible exception. Sometimes one encounters the claim that a complex instrumental chain necessarily lowers the probability that the outcome is accurate. But anyone who wears eye glasses knows that this is not correct. See Brown 1990 for discussion of the supposed probability calculation.

Returning to our story, by the late 1990s a consensus was reached in the physics community that the astrophysical theory originally under test was ok, and that the problem lay with the theory of neutrinos—in particular with the thesis that neutrinos have no rest mass. Why is this relevant? While these various experiments were going on, alternative theories were developed in which neutrinos have rest mass, and these theories predict that neutrinos change type as they move along (the technical term is that they *oscillate*). Thus the neutrino deficit can result from some of the electron neutrinos originally produced changing type, which makes them undetectable by Davis' detector. While other types of neutrinos will become electron neutrinos, these will have energies that are too low to be registered by Davis' detector. This modified view of neutrinos, which may well have significant consequences for other features of high-energy physics, is currently being tested in new experiments.

All of these experiments provide constraints that must be met by new theories in these fields—and sometimes by theories in other fields not originally under discussion. For example, in 1987 several neutrino detectors around the world responded to some phenomenon for which they had not been designed. It was quickly determined that they were responding to a supernova that was also detected by more traditional means. This led to the recognition that neutrinos provide an additional, previously untapped, source of data about stellar processes with the result that neutrino telescopes are now under construction. These will operate along with the optical telescopes that have been in play since Galileo, and the radio, X-ray, and other telescopes that have been developed during the last few decades. As a result of these developments, fields of study that had previously been pursued independently of each other are now linked together so that the range of theses that can be altered in response to a problematic outcome has grown. But the impact of a specific change has also grown so that we have an increase in both the range of interpretations and the constraints on acceptable interpretations.

The growing variety of telescopes illustrates a central feature of scientific research: a continual drive to extend the range of our interactions with nature—and thereby increase the constraints that theories must meet. It may well be that no body of constraints, no matter how rich, will ever dictate a single theory in any field of scientific research. But these constraints do limit the range of theories that can be reasonably defended. Elimination of theories from the class of serious contenders is one way in which we make progress in learning about nature. It is an open question just how far we can constrain the set of admissible theories in a particular domain.

I have been working towards a new way of thinking about empirical evidence; I now want to sketch that alternative view.[10] I will focus on the process of testing theories, leaving aside cases such as radioactivity where nature impinges on us in an unanticipated way, and cases in which scientists poke at nature to see what comes up.[11] Scientific theories make claims about some aspects of nature; tests of a theory require that one endeavor to interact with those aspects. This requires that the theory, along with currently accepted auxiliary theories, are sufficiently well developed to allow researchers to do two things: find a means of interacting with the items in question, and predict the outcomes of those interactions. The solar-neutrino experiment is an example of a contemporary testing procedure. The experiment provided a new test of the extant account of stellar processes by interacting with a star in a way that had never been attempted before. The procedure is highly indirect and involves multiple theories, but those theories provided the grounds for attempting this kind of interaction. Note especially that such procedures involve *attempts* to interact with a presumed aspect of nature. A failed attempt to

---

[10] See Brown 1995 for a more detailed account.

[11] As an example of the latter situation, here is Bruno Rossi's description of the decision to put an X-ray detector on a rocket—an experiment that led to the development of X-ray astronomy: "The initial motivation of the experiment which led to this discovery was . . . a subconscious trust of mine in the inexhaustible wealth of nature, a wealth that goes far beyond the imagination of man. This meant that, whenever technical progress opened up a new window into the surrounding world, I felt the urge to look through this window, hoping to see something unexpected" (Rossi 1977: 39).

carry out an intended interaction may be just as important an outcome as a result in which the attempt succeeds but does not fit the detailed expectations. Suppose, in the solar-neutrino case, that no neutrinos at all had been detected. This would have been at least as important an outcome as detecting a neutrino flux that differs from than that which had been predicted—it is simply another case in which the outcome of a procedure differs from the predicted outcome. But this possibility is worth emphasizing because it underlines an important point: in a sufficiently well-developed theoretical context, finding nothing can be an important result, while finding nothing without appropriate background beliefs will typically have no significance at all. The theories that provide the basis for designing evidence-gathering procedures and interpreting the outcomes of these procedures thus have an *enabling* function: without these theories we would not be able to design procedures that may result in a challenge to these very theories, and that provide new constraints on future theorizing. In other words, the introduction of theory-guided procedures for collecting evidence permits the introduction of a richer array of tests than can be developed without such theory guidance.

Since the involvement of theories in the process of acquiring evidence serves as a source of skepticism in much philosophical literature, I want to dwell for a moment on the enabling role that theories plays in our attempts to study nature. Without the guidance of theory we would never have looked for evidence of the motion of the earth through the ether, for solar neutrinos, for the Higgs boson, and for much else that is central to the development of science. One of the key lessons of this development is that nature is full of items that we cannot detect with our unaided senses—viruses, radioactivity and toxins in our environment, and much more. We can discover and study these—but only with the guidance of theories that indicate their existence and properties. And these theories can be tested, improved, and sometimes replaced as a result of the testing procedure. So the introduction and testing of theories plays a central role in the discovery of unanticipated features of nature, and in providing new constraints on theorizing—on both new theories and on theories that are already in play.

This can occur in many ways. Taking a new example, in a pre-Copernican framework studies of our sun would not be relevant to our understanding of the stars. But once we have come to view the sun as a fairly typical star, we can study stars by studying the sun (and conversely). As a result, our understanding of stellar-energy production is constrained by our understanding of solar-energy production. In general, the use of theory-guided means of acquiring evidence opens up new means of learning about the world, while also introducing new epistemic risks. The risks are an unavoidable price that we pay as we seek a richer understanding of nature.

This view of theory-testing is largely a traditional hypothetico-deductive account—except that it drops the special role of sensory experience that has usually been associated with this account. Note that I have avoided the term ‘observation’, with all its historical baggage, in the above sketch, and it is worth emphasizing that the qualitative features of the instrumental output we detect with our senses will often play no particular epistemic role.[12] Moreover, evidence-gathering procedures are designed to produce outcomes that are stable over time and intersubjectively available—and it is usually not important whether different researchers access an output in different ways. What is important is the significance of the output in the theoretical context—which holds even when the output is nothing at all.

Let us now reconsider the three of versions of theory-dependence that survived our earlier discussion. The second version reads:

---

[12] These features may play a role in some cases, such as when we are studying perception. In the solar-neutrino case, the amount of argon produced was determined by counting radioactive decay events. When the instrumentation detected an event, the signal was channeled to an oscilloscope and the oscilloscope display was photographed so that physicists could study the shape of the display to help distinguish genuine argon-decays from similar events. In this case, a perceived shape was significant—although with sufficiently sophisticated instruments that shape might have been detected by touch rather than by vision or analyzed by a computer without anyone ever examining the photograph.

> 2. Our concepts and beliefs play a role in the constitution of our sensory experience although there is some other element that limits the range of interpretations.

On the basis of our recent discussion this should be rewritten as:

> 2′. Our concepts and beliefs play a role in the constitution of empirical evidence, although there is some other element that limits the range of relevant concepts and beliefs.

The additional element is provided by the outcomes of our interactions with nature. The fact that these outcomes are subject to varying interpretations often points the way to further interactions that yield new constraints. This is just a variation on the traditional requirements that theories should be subject to continual testing, and that competing theories should imply different results for at least some further tests. It is worth repeating that instrumental outputs need not be described in language that is independent of all theoretical interpretations. They need only be described in language that is independent of the theories being compared.[13] Such situations are common in the history of science. In addition to examples already introduced, consider current debates over the interpretation of quantum theory. Bohm's account of what is going on at the microlevel is quite different from that of the Copenhagen account but there is agreement on the experimental outcomes. In fact, both accounts yield the same experimental predictions, and this is a major factor in making the choice between them quite intractable.

Our third form of theory-dependence now becomes:

> 3′. For evidence to be relevant to the evaluation of a theory, that evidence must be described using the language of that theory.

---

[13] This relatively neutral language is not required when only one theory is concerned.

Note, however, that finding an interpretation of a body of data in the conceptual framework of a particular theory may not be straightforward, and the failure to find such an interpretation may pose a serious challenge to a theory. So the need to find such an interpretation becomes another way in which empirical outcomes constrain theories.

Finally, the fourth version becomes:

> 4′. Which evidence scientists seek to collect is determined by accepted theory.

As I have already noted, deciding to look in a particular place does not determine what we will find. Moreover, as Kuhn sometimes emphasized, the fact that a theory picks out a particular portion of nature for detailed exploration enhances the opportunities for unanticipated discoveries.

I now want to emphasize another feature of the interaction between interpretation and constraint that was noted briefly above. The outcome of any well-done experiment provides a permanent constraint on future theorizing in its domain, although the outcome may take on a very different interpretation than was adopted by the original experimenters.[14] After the development of special relativity, the outcome of the Michelson-Morley experiment was interpreted as showing the absence of an ether, but this was not how it was originally interpreted. The original purpose of the experiment was to compare two different views of motion through the ether: a view which held that the earth moves freely through the ether, and one which held that the earth drags ether along as it moves. In its original context, the outcome supported the latter view. Indeed, to suggest that it indicated the absence of an ether would have undermined the experiment which was developed on the basis of the wave theory of light. At the time, no one could make sense of this theory without an ether. Still, it is possible to describe the outcome of the experiment without using the language of either relativity or its predecessor theories, and any new theory in this domain will have to account for this data.

---

[14] For discussion of this theme see Ackerman 1985, Franklin 1986, 2002, Galison 1987, 1997.

I turn now to another theme that pervades the recent literature on theory-choice in science: that the existence of different theories with different conceptual structures generates problems of communication. Many people (including Kuhn some of the time) write as if each of us is locked into a single language and is thus at a cognitive loss when faced with a different language that cannot be directly translated into our own language. But, as Kuhn eventually recognized, many people are multilingual and can shift between languages with no great difficulty. In a similar way, a contemporary physics professor may teach classical mechanics in the morning and quantum theory in the afternoon while also doing research in general relativity—and will rely on classical notions of space and time when planning a vacation or a tryst. Physics students also manage to study all of these without confusion. Moreover, while treating a theory as a language is illuminating in some contexts, it is quite misleading when taken to extremes.[15] Natural languages are much richer than scientific theories, and there is no particular difficulty in discussing various theories in a language that is shared among advocates of competing theories. The ability to deal with multiple theories is especially striking when we look at major innovators. Galileo, Newton, Darwin, and Einstein, for example, were complete masters of the theories they endeavored to supersede. Because of this, they were able to present results in ways that were understandable to all interested parties and that might even encourage others to do the work necessary to learn how they arrived at their results. There is really no great difficulty in noting where a stone dropped from the top of the mast of a moving ship lands, whatever difficulties may remain about why it landed where it did.

In addition, translation into my current language is not the only way of learning a new language. There is a theme that is central to *Structure* but drops out of much of Kuhn's later work, although it reappears in some his last writings. When we examine scientific research—that is, the process, instead of just on the product—we find that assessment and communication depend on a wide variety of human

---

[15] See Sankey 1994, Ch. 3 for a helpful discussion.

cognitive skills. Languages and theories are created by humans who have such skills for the use of other humans who share these skills. We deploy these skills when we endeavor to understand a language that embodies a different conceptual system than the one with which we are already familiar. Discussing the case of the historian (who must often deal with larger conceptual gaps than does the scientist involved in actual theory choice), Kuhn writes:

> Faced with untranslatable statements, the historian becomes bilingual, first learning the lexicon required to frame the problematic statements and then, if it seems relevant, comparing the whole older system (a lexicon plus the science developed with it) to the system in current use. Most of the terms used in either system will be shared by both, and most of these shared terms occupy the same positions in both lexicons. Comparisons made using those terms alone ordinarily provide a sufficient basis for judgment. (1989: 77)

A few years earlier Kuhn wrote:

> Translation is, of course, only the first resort of those who seek comprehension. Communication can be established in its absence. But where translation is not feasible, the very different processes of interpretation and language acquisition are required. These processes are not arcane. Historians, anthropologists, and perhaps small children engage in them every day. (1983: 53, *cf.* 1993: 238)

Kuhn also tells us that "anything which can be said in one language can, with imagination and effort, be understood by a speaker of another. What is prerequisite to such understanding, however, is not translation but language learning" (1989: 61). And, "with sufficient patience and effort, [one can] discover the categories of another culture or of an earlier stage of one's own" (1991a: 220). Kuhn has also backed off from his earlier metaphor of a scientific revolution as a gestalt shift (although this may still be an appropriate analogy for particular historians). "To speak, as I repeatedly have, of a community's undergoing a gestalt shift is to compress an extended process into an instant, leaving no room for the microprocesses by which the change is achieved" (1989: 88).

But once understanding has been achieved, two further questions remain: the grounds for accepting one of a set of competing theories, and the grounds for believing that the theory we accept is true. It is important that we separate these questions. Kuhn emphasized this point when he distinguished between the *grounds for belief* and the *grounds for incremental change of belief* (1991b: 112-13). Even when the debate is between two false theories, there remains a difference between an intellectually appropriate process of choice and one that is arbitrary or based on inappropriate considerations. Within the limits of my rapidly vanishing time allotment I can make only a few remarks on each of these large themes.

With regard to the basis for theory choice, it is important that philosophers get past the "all or nothing" approach of holding that we have either logically conclusive grounds for accepting a theory or no grounds at all. Scientists have gotten over this long ago. Even Descartes recognized that proof beyond a shadow of a possible doubt was not the appropriate criterion for his own extensive empirical research. And, of course, none of us accept this demand in our daily lives. Having less than conclusive grounds for a decision is not the same as having no grounds at all. An adequate account of theory choice requires that we consider more than just formal relations (although these are an important part of the story). It also requires that we consider the cognitive skills that people bring to bear in carrying out evaluations. This thesis goes strongly against the grain of standard analytic epistemology. Working under the influence of Frege, analytic philosophers sought to eliminate any psychological consideration from epistemology. This was not the view of the classical empiricists for whom our cognitive abilities are at the center of an account of human knowledge. These themes have reemerged in naturalistic epistemology—especially in work that has gotten beyond Quine's behaviorism (see Kitcher 1992 for a useful introduction to these developments).

Along with the role of cognitive skills, there is another theme that was important in *Structure*, but dropped into the background in many of Kuhn's later writings: the role of epistemic communities in scientific decision making. For many epistemologists, introducing a social role into epistemic decisions is tantamount to giving up a reasoned basis for our choices. But the use of social resources, along with evidence and reasoning, can enhance our epistemic power. We all acknowledge this when we consult manuals, use computers and software designed by others, and ask our colleagues for comments on drafts of our papers. This theme has also begun to move into so-called "mainstream" philosophy under the rubric "social epistemology." I urge that the next step towards a deeper understanding of the grounds for epistemic assessment will come from further work on these two Kuhnian themes.

With regard to the grounds for believing that a scientific claim is true—or, more appropriately—that science is capable of moving towards true beliefs in at least some domains—I submit that evidential constraints provide the only basis that we have for such beliefs. We have better grounds for such beliefs in the case of such contemporary theories as relativity, quantum theory, plate tectonics, and the evolution of species than our ancestors had for their beliefs because the newer theories have faced much more demanding tests than their predecessors. Of course, the process of testing and accumulating constraints continues, and it is an open question whether we will ever accumulate sufficient constraints in a particular field to converge on a final theory. But even if we do not, even if theory remains permanently underdetermined by evidence and there are always competitors, this must not be confused with the claim that all proposals are equally legitimate. Vast arrays of older theories have—by the scientifically appropriate standards—been ruled out of contention. So have even vaster arrays of logically possible theories that no one has been foolish enough to propose. Just what can be accomplished in a particular domain is itself an open question, and the best approach we have to answering this question is continued development of theories, accumulation of constraints, and their interplay.

References

Ackerman, R. (1985), *Data, Instruments, and Theory*, Princeton: Princeton University Press.

Bahcall, J. (1990) *Neutrino Astrophysics*, Cambridge: Cambridge University Press.

Basu, P. (2003) “Theory-ladenness of Evidence: A Case Study from the History of Chemistry,” *Studies in History and Philosophy of Science* 34A: 351-368.

Brown, H. (1987) *Observation and Objectivity*, New York: Oxford University Press.

---(1990) “Prospective Realism,” *Studies in History and Philosophy of Science*, 21: 211-42.

---(1993) “A Theory-Laden Observation Can Test the Theory,” *British Journal for the Philosophy of Science*, 44: 555-9.

---(1994) “Circular Justifications,” in D. Hull, M. Forbes and R. Burian (eds.) *PSA 1994* vol. 1, East Lansing MI: The Philosophy of Science Association, 406-14.

---(1995) “Empirical Testing,” *Inquiry*, 38: 353-99.

---(2005) “On the Epistemology of Theory-Dependent Evidence,” in A. Raftopoulos (ed.) *Cognitive Penetrability of Perception*, Hauppauge NY: Nova Science Publishers, 49-71.

---(2007) *Conceptual Systems*, London: Routledge.

Bruner, J. and Postman, L. (1949) “On the Perception of Incongruity: A Paradigm,” *Journal of Personality* 18: 206-223.

Cowan, C., *et al.* (1956) "Detection of the Free Neutrino: A Confirmation," *Science* 124: 103-4.

Davis, R., Harmer, D., and Hofman, K. (1968) "Search for Neutrinos From the Sun," *Physical Review Letters* 20: 1205-09.

Franklin, A. (1986) *The Neglect of Experiment*, Cambridge: Cambridge University Press.

---(2001) *Are There Really Neutrinos? An Evidential History*, Cambridge MA: Perseus.

---(2002) *Selectivity and Discord: Two Problems of Experiment*, University of Pittsburgh Press.

Galison, P. (1987) *How Experiments End*, Chicago: University of Chicago Press.

---(1997) *Image and Logic*, Chicago: University of Chicago Press.

Greenwood, J. (1990), "Two Dogmas of Neo-Empiricism: The `Theory-Informity' of Observation and the Quine-Duhem Thesis," *Philosophy of Science* 57: 553-574.

Hempel, C.(1965) "The Theoretician's Dilemma, in *Aspects of Scientific Explanation*, New York: Free Press, 173-226 .

Kitcher, P. (1992) "The Naturalists Return," *Philosophical Review*, 101: 53-114.

Kuhn, T. (1962) *The Structure of Scientific Revolutions*, Chicago: University of Chicago Press.

---(1983) "Commensurability, Comparability, Communicability," in Kuhn (2000), 33-57.

---(1989) "Possible Worlds in History of Science," in Kuhn (2000), 58-89.

---(1991a) "The Natural and the Human Sciences" in Kuhn (2000), 216-223.

---(1991b) "The Trouble with Historical Philosophy of Science," in Kuhn (2002), 105-120.

---(1993) "Afterwords," in Kuhn (2000), 224-252.

---(2000) *The Road Since Structure*, Chicago: Chicago University Press.

Lewis, C. I. (1956) *Mind and the World Order*, New York, Dover Books.

Pais, A. (1986) *Inward Bound*, New York: Oxford University Press.

Reines, F., and Cowan, C. (1953) "Detection of the Free Neutrino," *Physical Review* 92: 830-31.

Rossi, B. (1977) "X-Ray Astronomy," *Daedalus* 106, no. 4: 37-58.

Sankey, H. (1994) *The Incommensurability Thesis*, Aldershot, UK: Avebury.

Sellars, W. (1953) "Inference and Meaning," *Mind* 62: 313-38.

---(1963a) *Science, Perception and Reality*, New York: Humanities Press.

---(1963b) "Theoretical Explanation," in B. Baumrin (ed.) *Philosophy of Science: The Delaware Seminar*, vol. 2, New York: Wiley, 61-78.

---(1974) "Meaning as Functional Classification," *Synthese*, 27: 417-37.

Shapere, D. (1977), "What Can the Theory of Knowledge Learn from the History of Knowledge?" *The Monist* 60: 488-508.

---(1982) “The Concept of Observation in Science and Philosophy,” *Philosophy of Science*, 49: 485-525.

Shogenji, T. (2000) “Self-Dependent Justification Without Circularity,” *British Journal for the Philosophy of Science*, 52: 287-98.

Trimble, V., and Reines, F. (1973) "The Solar Neutrino Problem--A Progress (?) Report," *Reviews of Modern Physics* 45: 1-5.