The Relation of Theories to Experience

Coherence and operations in the epistemology of science

by Hywel Griffiths

Department of Philosophy,

Central European University

In partial fulfilment of the requirements for the degree of Master of Arts

Supervisor: Hanoch Ben-Yami

Budapest, Hungary

2010

TABLE OF CONTENTS

Introduction

To begin where we must begin: in the middle. The world of our everyday experience of nature is one of mid-sized objects and mid-sized time-scales. We cohabit the world with other people, from whom we learn a language, and with whom we communicate. At any stage of inquiry, whether during an individual life or through history, we already have a conceptual or descriptive system of some kind, with an indefinite origin, having passed through many modifications, and an indefinite future, with the expectation of further modification to come. This thesis will be an investigation into what justifies us in accepting a conceptual system as an adequate description of nature. It is clear that our experience of nature must be a final arbiter in any such justification. Our experience of nature must therefore be among the foundations of the project of discovering the truth about the world. The question then becomes one of asking in exactly what way our conceptual or theoretical systems can be related to our experience of nature so that we may be justified in accepting them. The ways in which our conceptual systems are actually related to our experience of nature are the various scientific methods themselves. An investigation of these is therefore in order. The particular conceptual systems I will be focused on in this thesis are our physical theories, though much of what I say will be applicable more generally to other areas of science.

The thesis will consider two problems which have been seen as challenging the testability of scientific theories, and hence the rationality of accepting them. These are the problems of theory-dependence, and of holism. I give arguments for accepting both of these problematic theses, and then develop an account which demonstrates, despite these problems, the rationality of the scientific project, and hence of the acceptance of scientific results.

Terminology

I will now give a rather schematic account of how I will be using certain terms in the thesis, and their relations to each other. Finer points regarding the concepts denoted by these terms will be introduced and, I hope, become clear as the thesis progresses.

Theories are sets of interrelated statements, forming a consistent logical structure. A geometry would count as a theory on this definition. Physical, scientific, or empirical theories are distinguished from purely abstract structures by the context in which they are applied and tested. (A particular Riemannian geometry with variable curvature may be considered a geometry that has been applied and tested and thereby become a part of physical theory.)

The testing of a theory requires a statement known as a prediction to be derived from it (though not necessarily from it alone), and this prediction needs to be tested for consistency with observation. An observation is an experience involved in the testing of scientific theories. In logical terms, consistency is a relation between statements whereby their conjunction does not imply a contradiction. Since consistency is a relation between statements, it is necessary that a statement be produced from an observation (which is an experience and not a statement) in order to test the theory. A statement produced from an observation (though not necessarily from it alone) I will call an observational statement, and the process by which this occurs I will call an experiment. Predictions and observation statements are not always conceived of as being different, but I wish to make this distinction since it allows for a clearer presentation of the problems and their solution. I will also occasionally use the term conceptual scheme, whereby I mean the general background knowledge within which a theory is embedded, and which may be more or less explicitly articulated.

Chapter 1: Analysis of the problems

The problem of theory-dependence

The obvious importance of our experience in justifying our theoretical systems leads naturally to a linear conception of justification whereby scientific theories require a self-justifying or basic foundation of statements based directly on observations underlying scientific theories. I will be arguing against the existence of such a foundation. The criticism of such a view will apply equally whether the foundation is considered to be phenomenal, as in Carnap's *Der logische Aufbau der Welt* (1928), or physicalist, as articulated by Hempel (1952). The criticism will be based on the theory-dependence of observation, a thesis called by Hacking (1983) 'idealist-leaning' since it 'makes the very content of the feeblest scientific utterances determined by how we think, rather than mind-independent reality.' I will argue for a strong form of theory-dependence, such that there are no statements justified exclusively by experiences. My views have the consequence that our feeblest scientific utterances are unavoidably always, at least partly, determined by our theories, and I will show how this is compatible with the rationality and objectivity of science.

If a prediction derived from theoretical statements and an observational statement produced by an experiment are to be such as to permit logical relations between them, they must share the same vocabulary, and use the same concepts. (The problem should really be called 'the theory-dependence of observation statements,' since an observation, on its own, is scientifically useless since it cannot have logical relations to predictions.) The first sentence which comes into one's head when observing a phenomenon may therefore be inadequate as an observation statement which can have logical relations with a prediction. For this reason, not every way of describing an observation is adequate as an observation statement, and what is adequate will depend on the context of the experiment and the theory that one wishes to test. An evocative passage from Pierre Duhem's *The*

Aim and Structure of Physical Science illustrates the point. It should be read in a mysterious, whispered tone:

> Go into this laboratory; draw near this table crowded with so much apparatus: an electric battery, coils, a small iron bar carrying a mirror. An observer plunges the metallic stem of a rod, mounted with rubber, into small holes; the iron oscillates and, by means of the mirror tied to it, sends a beam of light to a celluloid ruler, and the observer follows the movement of the light beam on it. There, no doubt, you have an experiment; by means of the vibration of this spot of light, this physicist minutely observes the oscillations of the piece of iron. Ask him now what he is doing. Is he going to answer: 'I am studying the oscillations of the piece of iron carrying this mirror?' No, he will tell you that he is measuring the electrical resistance of a coil. If you are astonished, and ask him what meaning these words have, and what relation they have to the phenomena he has perceived and which you have at the same time perceived, he will reply that your question would require some very long explanations, and he will recommend that you take a course in electricity. (Duhem 1962 [1906], 145)

Duhem does not seem to distinguish types of theory-dependence explicitly, so at this point I should make the clarification that there are at least three types of theory-dependence, which can be termed *perceptual*, *semantic*, and *epistemic*. At least the first two of these are to be found in Kuhn (1962).

Perceptual theory-dependence

I will discuss perceptual theory-dependence first only in order to show why it is irrelevant for the purposes of this thesis. Kuhn's account of perceptual theory-dependence is based on experiments such as Bruner & Postman's (1949), in which anomalous playing cards such as a black five of hearts are shown briefly to subjects, who fail to recognise them as anomalous, interpreting them instead as ordinary playing cards, such as a red five of hearts. Brewer & Lambert's (2001) discussion of these types of cases taking into account more recent results concludes that prior theory or belief does influence perception, but that this only occurs in a strong sense when the stimulus is ambiguous or degraded, e.g., as in the case of the card experiment, when the stimulus is of very short duration. A couple of cases from the history of science illustrate this issue. The first is the controversy between Rutherford and Petterson concerning the emission of protons from elements such as carbon or silicon undergoing bombardment by alpha particles (described by Steuwer 1985). Petterson's lab found a positive result in this experiment, whereas Rutherford's did not. The results were attained by discriminating flashes on a scintillation screen through a microscope. This dispute was not solved by attempting more careful perceptual discriminations, but by testing Petterson's assistants themselves. It was found that they reported proton flashes just as frequently whether the apparatus was operating normally as when it was operating in conditions where no protons could possibly have been detected. Petterson's results were thereby uncontroversially discredited.

Another historical example is the observation of Saturn's rings (discussed in Brewer & Lambert 1993, citing Van Helden 1974). Prior to Huygens' hypothesis that Saturn was encircled by rings, astronomers including Galileo made drawings representing their observations of Saturn as having satellites or handles rather than rings. After Huygens' hypothesis they came to see Saturn as having rings, a hypothesis which has afterwards been vindicated by superior methods of observations which do not leave open the interpretative possibility that what is perceived are satellites and not rings. Again, this is a case of a dispute arising simply from the stimulus being ambiguous or degraded. Therefore, it may be the case that this kind of theory-dependence is not a particularly significant threat to a linear view of justification, since it can be straightforwardly eliminated by improving the conditions under which perception takes place, or by involving a control condition in the experiment. In any case, there are other forms of theory-dependence which seem to pose a more substantial problem and which will be the focus of the remainder of this paper.

Semantic theory-dependence

To turn then to semantic theory-dependence, this is the view that the meaning of observational statements is dependent upon the meaning of theoretical ones, and they are subject to be modified or rejected in the face of changes in the theoretical statements. Semantic theory-dependence occurs because observational statements need to contain the same concepts as the predictions derived from theoretical statements in order to be logically related to them.

The meanings of terms in abstract systems of postulates can be given syntactically by the role they play in the system. This is a familiar point which is illustrated by the different meanings of the terms 'point', 'line', 'triangle', and so on, in Euclidean and non-Euclidean geometries respectively, due only to changes in the structure of one of the postulates which define the respective geometries, and not to the introduction of any new terms. The relevance of this purely abstract fact to the case of science is that theoretical and observational sentences must be part of one inferential framework in order to have any relation at all to one another.

Physical theories are interrelated statements forming a consistent logical structure, and therefore have what can be called an *abstract meaning* as the result solely of syntactic interrelations between the symbols in the theory. The terms in physical theories are thus subject to implicit definition and change in meaning due solely to changes in the structure of the theories, in the same way as the terms in the geometrical postulates. In addition, abstract physical theories gain what can be called *physical meaning* from the experimental and manipulatory context in which the abstract theoretical structure is embedded. This is akin to the Wittgensteinian idea that meaning follows use, but applied more narrowly to the use of a concept within the experimental and explanatory confines of physical science. Through this embeddedness in a context of experiment and application, the variables involved in the statements of the theories come to represent physical quantities. Observations of the results of measurements, which are supposed to give a numerical value to these theoretically designated quantities, are what are required in order to test the theoretical statements. The meaning of the statements describing these observations is therefore dependent upon what we conceive these physical quantities to be, and this conception is dependent upon the structure of the theory in which the quantities are posited. This is not to say that what is meant by terms in observational statements is *determined* by the theoretical structure (the actual details are more complicated, as will be discussed below), but it nevertheless means that changes in the theoretical structure can result in changes in the meaning of observational terms.

Bogen (2009) has pointed out that the above conception of semantically theory-dependent observation statements is based on a linguistic model of both theories and observations as sets of sentences, whereas records of observation can take many different representational forms, including pictures, graphs, tables of numbers, and so forth. However, whatever their format, representations of the results of experiment are always highly dependent upon the theoretical framework, explicit or implicit, in which the observation is carried out. What is the scientific significance of a photograph of the tracks in a bubble chamber without the theoretical framework of the standard model of the behaviour of subatomic particles? - one is tempted to say 'nothing'. So whether we talk of 'meaning' in a restricted sense or broaden the view to the scientific significance of other forms of representing the results of observation, this does not eliminate the problem of semantic theory-dependence.

Epistemic theory-dependence

Epistemic theory-dependence is a consequence of the interpenetrating nature of all statements, and the absence of a principled way to separate self-justifying observational statements from derivatively justified theoretical statements. The idea is that some theoretical background is always required in order to produce statements from experiences. Theoretical statements need to be used or assumed (knowingly or unknowingly) in order to interpret the results of an experiment so as to produce an observation statement. This seems to be the form of theory-dependence that Duhem had in mind in the passage quoted previously, because immediately preceding the passage he gives a statement of his view as follows: 'An experiment in physics is not simply the observation of a phenomenon; it is, besides, the theoretical interpretation of this phenomenon' (Duhem 1962 [1906], 144). Popper expresses a similar idea: '...we can utter no scientific statement that does not go far beyond what can be known with certainty ‘on the basis of immediate experience’. (This fact may be referred to as the ‘transcendence inherent in any description’.)' (2002 [1935], 76)

In general, when we articulate any statement in response to our experience, this can be at best an interpretation of our experience. We can't know beforehand how influenced we are by cultural, theoretical, emotional, or innate biases when we make interpretations of our experience. The interpretation of any observation as a relation between entities or physical quantities is dependent upon some kind of theoretical structure, whether an explicit and sophisticated physical theory or implicit 'naïve' background knowledge. As Popper continues from the above quotation:

> Every description uses universal names (or symbols, or ideas); every statement has the character of a theory, of a hypothesis. The statement, ‘Here is a glass of water’ cannot be verified by any observational experience. The reason is that the universals which appear in it cannot be correlated with any specific sense-experience. (An ‘immediate experience’ is only once ‘immediately given’; it is unique.) By the word ‘glass’, for example, we denote physical bodies which exhibit a certain law-like behaviour, and the same holds for the word ‘water’. Universals cannot be reduced to classes of experiences; they cannot be ‘constituted’. (Popper 2002 [1934], 76)

Therefore, even inferences to the existence of entities which are seemingly directly observed cannot be made unproblematically. To further illustrate this, Churchland (1985) gives examples of entities which were once putatively observed directly, but which we no longer believe to exist, such as witches and 'the starry sphere that turns about us daily'. Observational statements thereby are unable to function as an independent basis upon which the acceptance or otherwise of theoretical statements depends, for they are just as fallible as the theoretical framework to whose concepts they must conform, and indeed this fallibility is dependent upon the fallibility of theoretical framework itself.

The problem of holism

The problem of holism, or *confirmation holism* or *epistemic holism* as it is also called, is sometimes thought to be a consequence of the so-called Duhem-Quine thesis. However, as Donald Gillies points out (1993, 98-116), Duhem and Quine actually put forward quite different ideas. Duhem formulated the problem as 'An experiment in physics can never condemn an isolated hypothesis, but only a whole theoretical group' (Duhem 1962

[1906], 183). This is because the test of a theoretical statement requires the deduction from it of a prediction which can have logical relations to an observational statement produced from an experiment, and such predictions in general cannot be deduced from a single theoretical statement, or even a single theory, in isolation from other hypotheses, and in the absence of the context of some conceptual scheme, which may be more or less explicitly stated. Therefore, in the event of an observational statement being produced which contradicts the prediction, we are not logically compelled to reject the original theoretical statement.

Duhem gives the example of O. Wiener's test of F. E. Neumann's hypothesis that the vibration in a polarised ray of light is parallel to the plane of polarisation. Duhem notes of this experiment that, using Neumann's hypothesis, it was deduced that 'if we cause a light beam reflected at 45 degrees from a plate of glass to interfere with the incident beam polarized perpendicularly to the plane of incidence, there ought to appear alternately dark and light interference bands parallel to the reflecting surface;' (Duhem 1962 [1906], 183) and that this effect did not occur. This was taken to refute Neumann's hypothesis. However, Duhem points out that the experimental demonstration does not logically compel us to reject the hypothesis, since other theoretical statements were required in order to deduce a prediction using the hypothesis. For example, 'that light consists in simple periodic vibrations, that these vibrations are normal to the light ray, that at each point the mean kinetic energy of the vibratory motion is a measure of the intensity of light, that the more or less complete attack of the gelatine coating on a photographic plate indicates the various degrees of this intensity,' (Duhem 1962 [1906], 183).

Extending Duhem

Quine famously extends Duhem's thesis with the idea that the production of an observation statement which contradicts a prediction made using a theoretical statement can be accommodated not only with a modification in a relatively small group of theories, but in all of our empirical knowledge: 'our statements about the external world face the tribunal of sense experience not individually but only as a corporate body,' (Quine 1951, 38). He expresses the extreme holistic statement more forcefully as 'The unit of empirical significance is the whole of science,' (39). The implication is that in the face of a recalcitrant observation statement, we are not only in a position where it is logically indeterminate which statement or statements within a relatively small group of theories are false, but it is indeterminate which of any statement in our knowledge might be in need of revision. It should be noted that this conclusion of Quine's only follows if we take it that all empirical knowledge is required for any prediction, since the force of Duhem's statement of the thesis is dependent upon the fact of elementary logic that, if a prediction is deduced from a group of statements which is inconsistent with an observational statement, then at least one of that group is responsible for the inconsistency. This is what allows the logical indeterminacy of the location of the error responsible for the inconsistency. If the indeterminacy were to range over the whole of our empirical knowledge, this would be because the whole of our empirical knowledge would have been required in order to deduce the prediction.

I suspect that Quine became carried away here in his iconoclastic excitement. It is implausible that all of our empirical knowledge is required in order to deduce every scientific prediction. In the derivation of a prediction using Neumann's hypothesis, Wiener did not use, for example, the hypothesis that the Galapagos finches were all descended from a common ancestor (cf. Gillies 1993, 111). This is understandable, because the finch hypothesis is irrelevant to his experiment, and an inconsistency in the prediction derived from the hypotheses that he did use with the observation statement that he produced would have no bearing on the finch hypothesis, since modification of this hypothesis would not eliminate the inconsistency that he had discovered. Therefore, in the event of an observation statement which is inconsistent with a prediction, the location of the error indicated by this inconsistency would not be indeterminate over the whole of our knowledge, but only that part which was involved in the deduction of the prediction and the production of the observation statement, which is that part within which modification could eliminate the inconsistency.

Quine also extended Duhem's thesis in another way. In addition to his thesis resting on the idea that all empirical knowledge is required for the derivation of any particular prediction, he sought to apply the thesis to all empirical knowledge, whereas Duhem expressly states that the range of application of his thesis is only to physical predictions. We can put their respective theses this way: Duhem's view is that, for each physical prediction, a group of physical theories (not necessarily the whole of physics) is required for its deduction. Whereas Quine's view is (or implies) that, for each scientific prediction, all of our empirical knowledge is required for its deduction. I find, for the reason discussed above, that Duhem's thesis is more plausible and sophisticated than Quine's, but I see no reason for restricting the thesis to apply only to physical predictions, because, as will be further argued for below, it is unlikely that any kind of scientific prediction can be deduced from a theoretical statement or theory in isolation.

Quine's extension of the range of application of the thesis (which I have expressed as the clause 'For each scientific prediction' in my formulation of his view) appears to be legitimate. The modified thesis of holism which I wish to consider can therefore be expressed as: For each scientific prediction, a subsection of our empirical knowledge, not limited to one theoretical statement or one theory, is required for its deduction. The problem of holism that I wish to consider does not, therefore, extend the logical indeterminacy of the location of an error to the whole of science, but neither does it allow pinpointing the source of the error to one statement or theory. Rather, I will be concerned with situations in which relatively small groups of theoretical statements are used in order to deduce a prediction. In addition to the deduction of a prediction, in order for there to be a possibility of an inconsistency between the prediction and an observation statement, an observation statement must be produced. A further clarification that must be made to the problem of holism set out so far therefore needs to be made. The group of theoretical statements from which a prediction which is inconsistent with an observational statement is deduced may not be responsible for the inconsistency. Rather, the responsibility may lie in the method by which the observational statement was produced. The wrong materials may have been used in the instrument or it was poorly constructed; the operationalisation of various concepts in the theory may have been flawed (more on this below); implicit background assumptions may have been involved which are in need of modification or rejection; there may be some unknown and heretofore unimagined factors affecting the procedure, and so on.

Ian Hacking (1988) adds further detail to this view with his idea of the self-vindication of laboratory sciences. Elaborating on the traditional view of scientific practice as being an interaction between theory and observation, he produces a detailed taxonomy of various areas of scientific methods which evolve together in the course of the development of a science, including systematic theory, background knowledge, topical or auxiliary hypotheses, the material instrumentation, the model of the instruments, the representations of data, data processing techniques, and more. The relevance of his work in this context is an extension of Duhem's thesis that one can change one's systematic theory, or one of many other instrumental or background theories, in the face of recalcitrant data. In fact, according to Hacking, each element in his classification is plastic, in the sense that each element is subject to change in order to produce a more harmonious whole. The great difficulty in achieving harmony among the disparate elements of such a complex system of representations, instruments, and practices produces stability in the results of a mature science, because it minimizes the possible modifications of the system available to us that produce such harmony. No single one of the elements can be logically identified as the source of an inconsistency, but the system serves as an epistemic unit within which we can be assured of the need for a modification

Relation of the two problems

The thesis of theory-dependence is that an observation statement cannot be produced from an isolated observation – theoretical statements or background knowledge are also required. The thesis underlying the problem of holism is that a prediction cannot be deduced from an isolated theoretical statement – other theoretical statements and background knowledge are also required. The two theses are thus seen to be very similar. However, there is a deeper connection.

The problematic epistemological consequence of the thesis underlying the problem of holism is that, in the event of an inconsistency arising between a prediction and an observation statement, the location of the error giving rise to it is logically indeterminate. The problem arose in the foregoing discussion of holism as a consequence of a prediction being derived from multiple theoretical statements, but it is also a problem implied by epistemic theory-dependence. This is because, given epistemic theory-dependence, in the event of an inconsistency arising, it is logically indeterminate as to whether the location of the error responsible for the inconsistency is in the theoretical statements used to derive the prediction, or in the theoretical statements which contributed to the production of the observational statement. The two theses are thus seen to be two facets of the same problem.

Both problems taken together imply the need to recognise the interdependence of several components of scientific practice, any of which might be modified in the event of an inconsistency: measuring techniques and empirical laws, standards and methods, theories, and the meaning of terms in observational statements. Theories, representations of observational results, models of instruments and of procedures, and the actual manipulatory and experimental operations, form an ecology of interrelated components forming our epistemic practices.

Chapter 2: Operationalism

Theory-dependence implies the absence of a foundation of self-justifying observational statements underlying our knowledge. Justification needs to involve logical relations between statements, but not be limited to them, otherwise it would be vacuous. We therefore need to show how there can be justification without a foundation, and how it avoids logical emptiness. In order to avoid this emptiness, there need to be points of contact between the abstract logical structure of a theory, and our experience of nature. Theoretically-guided operations offer a way of seeing how such a justification is possible without requiring the naïve empiricist view that there are statements that can be made infallibly purely on the basis of immediate experience.

Operationalism, as it was originally conceived, is the idea that some concepts within theoretical structures can be defined exclusively by means of operations or procedures, so that observational statements can be made on their basis without being infected by theory. Operationalism derives from the ideas of Percy Bridgman, especially his book *The Logic of Modern Physics* (1927). Bridgman was particularly influenced by Einstein's (1905) philosophical innovations regarding the definition of time in developing his Special Theory of Relativity. Einstein realised that in order to define the synchronicity of watches or timepieces distant from each other, and given the finite constancy of the average velocity of light and the principle of relativity whereby physical laws are not affected by differences in uniform translational movement of systems of coordinates, an operational definition of distant simultaneity was required. Einstein's insight here was that no 'physical meaning' (Einstein 1905) could be given to the concept of distant simultaneity without such an operational definition. Consequently, the uncritical extension of the concept of the simultaneity of two events very near to each other to the context of the simultaneity of events distant from each other without such an operational definition was seen to be an error. Impressed by this, Bridgman sought to build science on a foundation of operationally defined concepts, whereby no concept found useful in one domain should be uncritically extended into others: 'We must remain aware of these joints in our conceptual structure if we hope to render unnecessary the services of the unborn Einsteins' (Bridgman 1927).

To this end, Bridgman attempted to show how scientifically meaningful concepts must be given definite meanings by operational definition. He begins with an investigation of perhaps the simplest of scientific concepts, that of length:

> We evidently know what we mean by length if we can tell what the length of any and every object is, and for the physicist nothing more is required. To find the length of an object, we have to perform certain physical operations. The concept of length is therefore fixed when the operations by which length is measured are fixed: that is, the concept of length involves as much as and nothing more than the set of operations by which length is determined. In general, we mean by any concept nothing more than a set of operations; the concept is synonymous with the corresponding set of operations. (Bridgman 1927, 5)

The statement is clearly too strong as a general theory of meaning. However many operations might be used to define length, the concept might still be legitimately extended or used in domains where a physical operation is inapplicable. For instance, we can legitimately speak of the ratios of the lengths in geometry, which is entirely abstract, hence undefinable by any physical operation of length measurement. We cannot consider the geometrical concept of length to be entirely unrelated to concepts defined by physical operations, since this will have the consequence in general that concepts in abstract systems would be entirely unrelated to operations of measurement, and this is obviously false in the context of any physical science.

Another general objection to operationalism is that a single operation is a particular event. In order for a concept based upon an operation to have any relevance at all to scientific concepts it must be based on the idea that the result of the operation reflects a general empirical law, and an operation whose scientific significance depends upon one or more general empirical laws is not independent of theory. As Popper says,

> it can be shown that measurements presuppose theories. There is no measurement without a theory and no operation which can be satisfactorily described in non-theoretical terms. The attempts to do so are always circular; for example, the description of the

> measurement of length needs a (rudimentary) theory of heat and temperature-measurement; but these, in turn, involve measurements of length. (Popper 1962, 62)

Bridgman himself seems to have recognised the theory-dependence of measurement operations, but seems not to have fully appreciated how it might lead to circularity. He says, while continuing his foundational discussion of the measurement of length with a rigid rod:

> We must, for example, be sure that the temperature of the rod is the standard temperature at which its length is defined, or else we must make a correction for it: or we must correct for the gravitational distortion of the rod if we measure a vertical length; or we must be sure that the rod is not a magnet or is not subject to electrical forces. (Bridgman 1927, 10)

What a preponderance of theoretical assumptions goes into so simple an operation as the measurement of length with a rigid rod! Bridgman seems to believe that it is unproblematic that, in order to make corrections to a rod for measuring length, we have recourse to a measurement of temperature. But our measurements of temperature depend on the measurement of length. If an operation of measurement, in order to supply us with any useful results at all, needs to presuppose such theoretical assumptions then it seems that we are again faced with a circularity: we find ourselves still suspended in a web of statements, with no way to plant our feet upon a theory-free ground in order to escape. Operationalism was supposed to define concepts in a theory-free way, but now we find that operations themselves must be theoretically defined in order to be of any use. How are we to specify an operation, conceived as an interaction between entities in a physical space, without recourse to any theoretical assumptions?

Another important objection is raised by Gillies (1972), who objects to the idea of operations giving a complete meaning to a concept on the grounds that it makes measurement methods valid simply as a matter of convention, so that it would be senseless to ask whether some method is a good one, or whether one method is better than another, so long as the method gives consistent results. Once a definition of a concept is 'fixed' by a set of operations, there can be no grounds for modifying the definition by modifying which operations are used or by selecting novel operations as more reliable or useful than older ones.

Given these weaknesses, how can the idea of operationalising concepts serve in the present context of attempting to relieve the epistemological problems discussed above? I believe the answer is to be found from a perspective which adopts as a central evaluative criterion the coherence of practical procedures with each other and with theoretical structures. The insights that can be gained from a modified operational approach indicate how this can be the case.

In the first place, operationalism is clearly inadequate as a complete theory of meaning. There are many aspects of a concept's meaning which cannot be captured by such an account. The abstract meaning of a concept, and its broader meaning deriving from the context of its use, cannot be given entirely merely by an operation or set of operations. Furthermore, there are many scientifically useful concepts which cannot be directly or straightforwardly operationalised, such as the probability function for an elementary particle in quantum mechanics. This ought not to mean that their use is illegitimate. However, the core idea which I wish to take from operationalism is that a theoretical structure, though perhaps partly composed of concepts without a direct operationalisation, must also contain concepts which are operationalised. This is not to say that their meaning should be given completely by any single operation or set of operations, since we have seen that this is untenable for a scientific concept. The point is rather that a concept's meaning follows from its use, and that it can be used in theoretical and operational ways – in effect, the concept's definition is based upon the interdependence of theoretical structures and practical operations. The use of operationalisation of concepts is therefore a means of partially solving the problem posed by semantic theory-dependence

A simple illustration: perceptual comparison

In order to illustrate the form of my solution to the problem of semantic theory-dependence, I will now introduce a simple example of a scientific procedure, beginning with a rudimentary operation of perceptual comparison, and a correspondingly rudimentary conceptual scheme. I will start by discussing a rudimentary concept of heaviness.

The concept of heaviness might be said to initially derive the greater part of its meaning from a certain kind of perceptual comparison. We can tell, up to a certain limit of precision, and within a certain range, which of two objects is heavier. We don't need to put it into words in order to do this. However, by means of such an unarticulated perceptual comparison, the observation cannot be placed into logical relations with statements, since statements can only have logical relations to other statements, and observations are not statements.

The conditions required for an observational statement to be placed in a logical relation to a theoretical statement depends both upon the character of the theoretical statement and upon the procedure by which an observation statement is produced from an act of observation or experiment.

As an initial candidate for a theoretical statement which we wish to test by an observational procedure, I will discuss the statement 'Mangoes are heavier than strawberries,' which I will refer to as T in the following. We can test this statement by means of a perceptual comparison. As a first approximation, let's say that we can carry out this test if we know how to use the words. At this stage, 'heavier than' implies a certain kind of perceptual comparison, because the way the word is used implies a certain kind of perceptual comparison, e.g. holding one object in one hand and another object in the other hand. For the testing of T we also require categories, and means of identifying their members. Then the logical structure of T is one of a (transitive, non-symmetric, non-reflexive) relation between individual members of two non-overlapping categories.

We now have these procedures contributing to the testing and the physical meaning of T: a means of categorizing entities, and a certain kind of perceptual comparison. It should be noted here that the categorisation of entities is not an observation, but an interpretation of an observation, dependent upon some conceptual scheme. We test T as follows: We identify and procure individual entities belonging respectively to the categories 'Mango' and 'Strawberry'. By means of T, and the identification of the entities as members of the respective categories under consideration, we can make a prediction: 'The entity identified as a mango will be heavier than the entity identified as a strawberry'.

The relation 'heavier than' draws its empirical significance both from the operation of a certain kind of perceptual comparison, and from the logical role it plays in the conceptual scheme. The concept is operationalised because of the correspondence between a certain logical relation (in this case a transitive, non-symmetric, and non-reflexive one) and a certain kind of perceptual comparison. This use of the concept operationally and logically within the context of a particular conceptual scheme allows an observation statement to be produced from the result of the perceptual comparison, which can have a logical relation with the prediction. The possible observation statements are as follows: 'The entity identified as a mango is heavier than the entity identified as a strawberry' or 'The entity identified as a mango is not heavier than the entity identified as a strawberry'. The first is consistent with the prediction, while the second is not. The operationalisation of the concept, as a convention of correspondence between a logical relation and a perceptual comparison, gives us a system of statements and practices by which a logical structure can be related to certain observations, because the concepts involved in the statements are given their meaning from their use in both theoretical and practical contexts.

However, the production of the second possible observation statement from the perceptual comparison would still not logically compel us to reject the T, because there were other components involved both in making the prediction and in producing the observation statement from the observation: we required the identification of the entities as members of their respective categories, and also the correspondence between the relation 'heavier than', and a particular kind of perceptual comparison. Either of these components could also be modified in the face of an inconsistent observation statement, leaving T to stand. So operationalism as I have set it out does not save us from holism.

However, we have made a some kind of discovery: the derivation of our prediction from T and the method of producing an observational statement from the experiment together lead to an inconsistency. The problem of holism means that the location of the error responsible for the inconsistency is indeterminate, but it is not indeterminate over the whole of our knowledge, but must be among the small number of components used to derive the prediction and to produce the observation statement. Nevertheless, even at this extremely basic level, the complications in the testing of theoretical statements by observational means implied by holism is clear, and will be dealt with later. Epistemic theory-dependence appears here because the production of the observation statement from the operational procedure involved the categorisation of entities based on some conceptual scheme. While this does not appear to be particularly problematic, this is because of the simplicity of the scenario considered. We shall see, while considering more sophisticated instances of scientific practice, that epistemic theory-dependence still poses significant problems, to which I shall turn next.

The problem of nomic measurement

An acute instance of the challenge that the epistemic theory-dependence of observation poses to an epistemology of science is what Hasok Chang (2004) refers to as 'The Problem of Nomic Measurement'. Discussion of this problem can serve to introduce some of the concepts with which to address the more general problems of epistemic theory-dependence. The problem of nomic measurement occurs when we attempt to measure some quantity x as a function of some more observable variable y. The case that Chang discusses extensively is the problem of ascertaining temperature by means of a thermometer, where x is the temperature and y is the reading of the thermometer. In order to do this, we require an empirical law of the form x=f(y). The problem is that if we wish to test this empirical law we would need an independent test of how x varies with y, but none is available. Determining x from some other method would leave that method susceptible to the same problem.

Comparability

Part of the solution that Chang proposes for the problem of nomic measurement, and part of the answer to the general problems with operationalism already discussed, is invoking comparability: an experimental operation should give the same result in the same circumstance. Also, two instruments, if they are supposed to be reliable indicators of some single quantity, should agree with each other in the same circumstance. This is the rationale for the general practice of increasing our confidence in the results of experiments by repeating them. The same result occurring in different instances of the experiment reinforces the idea that the physical factors one has isolated in constructing the experimental situation are those responsible for the result, and that the operation itself is repeatable and reliable. Science is built upon locating and systematising the invariances in nature. Regularities in the dynamic relations between quantities or entities which we can perceive are required for this. Without such regularities, all change would be random and there would be no way to understand the world. Any project to understand the world must therefore be based upon the assumption that such regularities exist. Comparability in the results of a process or in results of diverse processes is therefore a partial solution to the problem of epistemic theory-dependence, because it reduces the number of theoretical assumptions required in order for us to be justified in thinking that the processes which we are using as components of our epistemic practices reflect a natural regularity.

In the case of thermometry, comparability was partly achieved by calibrating different instruments at fixed points, for example, the freezing and boiling points of water. But this still left the possibility that instruments would disagree in their readings between and beyond the fixed points. The usual practice of dividing the interval on the thermometer into equal units between the fixed points, and to extend this scale beyond them, cannot be considered scientifically meaningful in the absence of a numerical temperature scale. Again, there is no way to construct such a numerical concept without already having a reliable thermometer (a thermometer being a device which assigns numerical values to temperatures – instruments calibrated at fixed points without a numerical temperature scale are more properly called thermoscopes (Chang 2004)). It was found in fact that thermometers using different thermometric fluids gave different readings in between the fixed points (examples compared by eighteenth century Genevan scientist Jean-André De Luc include water, distilled spirit, brandy, and old Languedoc wine). The readings of alcohol-based thermometers deviated from mercury and air thermometers, especially in higher temperature ranges. Mercury thermometers were also found to deviate from each other, as there seems to have been no way of constructing multiple instruments in precisely the same way, including using precisely the same chemical composition of the glass and blowing the bulb into exactly the same shape and thickness.

A significant advance in finding a standard way of measuring temperature was made by French physicist, Victor Regnault. He found that constant-volume air thermometers could be constructed so as to agree with each other when placed in the same heated baths of oil. These are thermometers based on the increase of the pressure of air when confined in a space of constant volume. The readings of the thermometers never diverged more than 0.3° between 0° and 340°, and never diverged by more than 0.1% of the measured magnitude. These air thermometers were also found to be in close agreement with thermometers using other gases such as hydrogen and carbon dioxide, but differed from sulphuric acid gas thermometers. However, the strong comparability between different thermometers using different gases gave good grounds for believing that these were instruments whose operations reflected a significant regularity in nature.

The same point can be illustrated with reference to time. Time needs to be measured according to the countable periodicity of some phenomenon. In order to be a precise measurement, the phenomenon has to be regular. But this means that the phenomenon should complete a cycle in a fixed amount of time, yet we can't give any meaning to this without an independent method for measuring time. Again, we must resort to comparability in methods of measurement in order to achieve a standard, testing diverse methods against each other and placing greater confidence in those methods which display a high degree of precision in their convergence.

Chapter 3: Coherentism

Reliability

As the foregoing demonstrates, it is seldom a straightforward matter to construct experimental procedures which give stable and reliable results. It is even more difficult to construct a network of procedures whose results cohere with each other. Yet it is this coherence which gives grounds for confidence in the results. It should be clear by now that there is little basis in the history of science for the claim that interpretations of sensory experiences are epistemically privileged in the way traditionally envisioned by empiricists – as forming a self-justificatory foundation for scientific knowledge. Rather, of higher epistemic value are procedures which give reliable results which cohere with the results of other reliable procedures. All empirical procedures involve perception of course, and procedures involving perception can be generally reliable and scientifically useful. But they can be evaluated as such not by virtue of forming a self-justificatory foundation for scientific knowledge, but by virtue of the very criteria of coherence with other such procedures and with theoretical knowledge which I have been discussing.

Another point which I have emphasised above is that perceptions by themselves are not in a form which permits them to be related to an abstract or symbolic structure such as a scientific theory. They must be interpreted and expressed symbolically in order to be placed in such a relation. These processes of interpretation and symbolic expression are not trivial, and should not be underestimated. Bogen & Woodward's (1988) distinction between data and phenomena is instructive in this case. In Bogen & Woodward's terminology, 'data' are the actual representations of observations made during an experiment. The data may be, and probably will be, different each time the procedure is carried out, due to ubiquitous uncertainty in measurement. 'Phenomena' are the stable results of such procedures, once controlled for systematic error and once the data are analysed, etc.The terminology can seem strange, since we naturally think of ourselves as observing phenomena directly, and that data are the representations of the results of such observations. But the term 'phenomena' in the sense employed here is supposed to represent what I have called 'observational statements' above, and consequently should not be confused with sense-experiences.

Bogen & Woodward's emphasis is on drawing stable and reliable results from data which are often noisy and not conducive to being given full explanations, due to unavoidable experimental uncertainty. For example, measurements of the melting point of lead rarely if ever give the same value, but cluster within a certain distance from some particular value. If there is no systematic reason for their departure from a single value, and the reasons for their departure from that value are random and independent, then data-analysis gives good reason for accepting the mean value of the data as the melting point of lead. This is in order to produce a reliable observational statement which can have logical relations with predictions from theory. The repeatability of a procedure is important. None of the individual observations are exactly repeated, thereby resisting description by a single statement. A procedure involving the analysis of the data derived from numerous observations gives greater stability and reliability. The relevance of this for my purposes is that the entire procedure, including data analysis and control for systematic error, constitutes a procedure which produces an observational statement which is stable enough to have scientifically relevant relations with statements derived from theory.

Inconsistency

Holism means that determining whether a theoretical statement is true requires determining the truth of other theoretical statements. In the absence of a self-justifying foundation for empirical knowledge, the evaluation of theoretical statements as true is always provisional. This is not to say that there are no statements beyond reasonable doubt, but only to say that being beyond reasonable doubt *given limited evidence* is never a guarantor of certainty. As I have mentioned previously however, scientific certainties can be achieved through the discovery of inconsistencies.

In the event of an inconsistency, we can be assured that something is wrong with one or more of the components involved in the procedure, whether this be the systematic theory, our background assumptions, a malfunctioning instrument, a poor experimental design, or an error in data analysis. There need be no law-like method of determining which among these requires modification.. Further tests may indicate which of these components is weakest (in terms of coherence with other parts of our knowledge or practice), and consequent modifications can be made to that component. Minimising inconsistency, while maximising the range and precision of consistency, provides a criterion by which we can judge whether the resulting modified system is superior to the previous one.

Furthermore, inconsistency between a prediction and an observational statement can admit of degrees when we have a scale of measurement. To make consistency scientifically interesting, we need to design experiments which limit the number of possible outcomes which are consistent with a prediction. Quantitative prediction and measurement are means by which we can restrict the possible outcomes consistent with the prediction. Since measurements are operations giving a quantitative output, they allow greater precision in the testing of theories. Limiting the theoretical statements necessary for the experimental production of an observational statement also limits the (anticipated) possible number of outcomes of the experiment. In this way, we can aim to design an experiment which has only one (or a small range of) possible outcomes which are consistent with the prediction. The problem of the indeterminate location of error can be ameliorated by taking these efforts to minimise the number of theoretical statements and concepts which we employ in any test of a theory, and to design experiments designed to restrict the number of possible outcomes which are consistent with the prediction. This gives a clear rationale within the context of science for Ockham's razor. Operationalising concepts as far as is possible is one way in which this can be done, since this minimises the number of theoretical assumptions that must be employed to produce an observational statement. Employing economy in theory is another way to do this, since it minimises the number of theoretical statements used to derive a prediction.

The account has a Popperian flavour because I, like Popper, wish to retain the idea that the strength of scientific results depends in large part upon such epistemically sound principles as those of elementary logic. However, the results of which we can have the greatest confidence are not theoretical statements, but inconsistencies. This differs from a naïve falsificationism because the falsifiable units are not theories or theoretical statements, but systems of theory and practice. The theory-dependence of observation statements is actually essential to this scientific project of discovering inconsistencies, because it is this fact which allows observation statements to have logical relations to predictions.

Coherence

An important development which emerges from the foregoing discussion is how using diverse methods to give the same result reinforces belief in the stability and reliability of the result. These diverse methods might be comparison between theoretical predictions and practical operations, as in the classic view of theory testing, or, as in the case of Regnault's thermometers discussed above, between two or more practical procedures. In both cases agreement between the results induces confidence in each of the things being compared, whether a practical operation or a theoretical structure. Progressive and more precise agreement of scientific results, spreading in scope and depth, justifies confidence in the methods used to gain these results. This agreement in the scientific results, whether practical or theoretical, constitutes coherence in scientific systems of explanation which can act as a criterion of acceptance for theoretical structures and practical procedures. In order to be empirical, this cohering scientific system must include practical procedures. This is how the problem of seeming failure to test scientific results based on holism is solved – we are not concerned here with need for the possibility to unequivocally falsify a specific theoretical statement. Rather, we can be assured for the need for modification within a specific system of interrelated theoretical statements and practical procedures, and we determine whether such modification constitutes an improvement over other possible modifications by seeing whether the resulting system coheres with greater range and precision than its rivals.

Conclusion

The analysis of the problems considered at the beginning of the thesis showed the necessity to adopt a perspective of the testing of scientific theories as an activity formed of multiple components interacting to produce each result. It was found that semantic theory dependence occurs because the meanings of words in an observation statement are influenced by words in theoretical statements, because observation statements need to contain the same concepts as the predictions derived from theoretical statements in order to be logically related to them. The solution offered to this problem was the operationalisation of the concepts involved so that the meanings of the term have a co-dependence on theory and the practical context of their use.

Epistemic theory-dependence was found to occur because theoretical statements need to be used or assumed in order to interpret the results of an experiment so as to produce an observation statement. Comparability is a way of reducing the number of such assumptions, by relating different processes to give the same result, thereby indicating that a significant natural regularity is being employed to give the result.

The problem of holism showed that the location of an error leading to an inconsistency is in general not logically determined, and so a procedure of testing cannot result in a logical compulsion to reject any single theoretical statement. The discovery of an inconsistency, however, can be considered a scientific result in which we can have complete confidence, providing us with the knowledge that we must make a modification somewhere among the components of our epistemic practices which were used in order to give rise to it. Ingenuity in the modifications undertaken in response to inconsistency, coupled with the aim of achieving an ever wider range and ever greater precision in the coherence of our epistemic methods serve to produce an epistemologically sound and progressive view of scientific practices, and illustrate the logical structure of justification which underlies the rationality of science.

References

Bogen, J. (2009) 'Theory and Observation in Science' *Stanford Encyclopedia of Philosophy*.

Bogen, J. & Woodward, J. (1988) 'Saving the Phenomena' *The Philosophy Review*, vol. 97, pp. 303-352.

Brewer, W. & Lambert, B. (1993) 'The Theory-ladenness of Observation: Evidence from Cognitive Psychology' *Proceeding of the Fifth Annual Conference of the Cognitive Science Society*, pp. 254-259.

- (2001) 'The Theory-ladenness of Observation and the Theory-ladenness of the Rest of the Scientific Process'.

Bridgman, P. (1927) *The Logic of Modern Physics.* New York: Macmillan

Bruner, J. and Postman, L., (1949) 'On the Perception of Incongruity: A Paradigm", 18: 206-23.

Carnap, R. (1928) *Der logische Aufbau der Welt.* Berlin: Bernary, transl. by R.A. George , Berkeley: University of California Press, 1967

Chang, H. (2004) *Inventing Temperature: Measurement and Scientific Progress.* Oxford: OUP

Churchland, P. M. (1985) 'The Ontological Status of Unobservables: In Praise of the Superempirical Virtues', in Churchland & Hooker (eds.) *Images of Science.*

Duhem, P. (1962 [1906]) *The Aim and Structure of Physical Theory.* New York: Athenaeum

Einstein, A. (1905) 'On the electrodynamics of moving bodies'. *Annalen der Physik*, 17: 891–921

Gillies, D. A. (1972) 'Operationalism' *Synthese*, Vol. 25, No. 1/2, , pp. 1-24.

Hacking, I. (1983) *Representing and Intervening.* Cambridge: CUP

- (1988) 'On the Stability of the Laboratory Sciences' *The Journal of Philosophy*, vol. 85, no. 10, pp. 507-514.

Hempel, C. (1952) 'Fundamentals of Concept Formation in Empirical Science' in , Volume 2, O. Neurath, R. Carnap, C. Morris (eds.), pp. 651–746.

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

Quine, W. (1951) 'Two dogmas of Empiricism'. *The Philosophical Review*, Vol. 60, No. 1, pp. 20-43

Steuwer, R.H., 'Artificial Disintegration and the Cambridge-Vienna Controversy,' in P. Achinstein and O. Hannaway (eds.), , pp. 239–307.

Van Helden, A. (1974) 'Saturn and his Anses' *Journal for the History of Astronomy*, 5: 105-121.