ONTOLOGICAL, EPISTEMOLOGICAL, AND METHODOLOGICAL POSITIONS

James Ladyman

INTRODUCTION

This chapter summarises various important ontological, epistemological and methodological issues in the philosophy of science. Ontology is the theory of what exists and is the foremost concern of metaphysics, which is the study of the most fundamental questions about being and the nature of reality. Ontological issues in the philosophy of science may be specific to a particular special science, such as questions about the ontological status of biological species, or they may be more general, such as whether or not there are objective natural kinds. In the history of science ontological issues have often been of supreme importance; for example, whether or not atoms exist was a question that occupied many scientists in the nineteenth century. In what follows, some fundamental questions of ontology are discussed, some of which, such as those concerning laws of nature, are also addressed in analytic metaphysics. A number of these issues also relate to debates in the foundations of physics about the ontological implications of our best physical theories. Readers who wish to know more about the philosophy of space and time and the nature of matter and motion should consult the companion volume to this on the philosophy of physics.

Epistemology is the theory of knowledge and as such is concerned with such matters as the analysis of knowledge and its relationship to belief and truth, the theory of justification, and how to respond to the challenge of local scepticism, say about the past, or global scepticism, which suggests that there is no knowledge. The particular epistemological problems raised by science mostly concern inductive inference, since it is widely accepted that substantive knowledge of the world cannot be obtained by deduction alone. The most fundamental such problem is to explicate the relationship between theory and evidence. There are also epistemological issues that only arise when we reflect on the status of unobservable entities posited by our best scientific theories.

Finally, methodology here means the theory of the scientific method. Is there a single such method for all the sciences, and if so what is it? How much should we expect the theory of the scientific method to help with the progress of science? Is the scientific method fixed, or does it change over time?

Handbook of the Philosophy of Science: General Philosophy of Science - Focal Issues, pp. 303–376.
Volume editor: Theo Kuipers. General editors: Dov M. Gabbay, Paul Thagard and John Woods.

There is obviously a good deal of overlap between the areas discussed in this chapter. For example, accounts of scientific methodology may have implications for the epistemology of science, and vice versa, and the epistemological issues of whether we ought to be scientific realists has a lot of bearing on whether science can help us address ontological issues. There are vast literatures on all these issues and I have offered some advice about further reading that should help the reader find their way into the subject.

1 ONTOLOGICAL POSITIONS

(I) Natural Kinds

We divide the world up into individuals and then we class many individuals together in kinds. Hence, we distinguish between horses and donkeys, gold and silver, apples and pears and so on. One of the main concerns of science is to systematically classify natural phenomena and substances into kinds. A system for dividing things into kinds is called a taxonomy. The progress of chemistry, biology, physics and all the sciences is in part the history of a series of re-classifications, and refinements and reinterpretations of existing classifications. For example, in chemistry the kind acid has evolved from the rough notion of a liquid that would react with a base, such as a metal, and produce a salt and water, to the heavily theoretical idea of a chemical that can donate a hydrogen ion; in biology the living world is no longer divided into plants and animals, rather there is a complex taxonomy of phyla, kingdoms, groups, families, genii, and species; and in physics there have been several taxonomic revolutions from the Newtonian idea of a world of corpuscles and forces, through the late nineteenth century heyday of fields, and into the taxonomy of the four fundamental forces and their associated quantum fields and particles. Nonetheless, there is usually a large degree of retention of taxonomic structure between successive scientific theories. So for example, contemporary chemistry still classifies as acids the most important acids known in the age of alchemy.

A fundamental ontological question is whether taxonomy is a matter or discovery or invention, or, in other words, whether there are any objective natural kinds. It is often taken for granted that there are in the case of the natural sciences, but there has always been controversy about whether kinds picked out by the human sciences are objective, or whether they reflect the values of particular societies to which the science and scientists in question belong. For example, there is a considerable amount of contemporary debate about whether psychiatric classifications are objective. The idea of objective natural kinds in biology was also cast in doubt by the development of the theory of evolution by natural selection. Species came to be seen as historically contingent and defined by ancestry relationships rather than by morphology. Even when it comes to physics, there are those who deny that the natural kind distinctions it makes are objective rather than pragmatic or socially determined (see the section on Truth below).

In the history of philosophy, questions about natural kinds are closely related to the issue of essentialism because it is often thought that all the members of a kind possess some properties necessarily, and that these properties are characteristic of the kind in question. The idea is that some of the properties of some particular object, are properties that it couldn't lack without being a different kind of thing. For example, a piece of copper could be a different shape, but it could not fail to be a good conductor of electricity. The modern debate about natural kinds begins with John Locke's critique of Aristotelian essentialism and the former's distinction between real and nominal essences. Take gold: the nominal essence is the collection of ideas associated with the word gold, such as those of metallic, yellow, malleable and so on. Locke argued that whatever was the real essence, and he thought it would be some kind of characteristic microstructure, could not be known with any degree of certainty. Contemporary science has restored the faith of many philosophers in natural kinds and essences because our understanding of the chemical elements in terms of the periodic table seems to give us knowledge of the microstructure responsible for their sensible properties? Gold is characterised precisely as that atom that has 79 protons in its nucleus (an atomic number of 79). It is not only elements but also compounds that are often thought to have essences revealed by science. For example, it is widely said among philosophers that the essence of water is H_2O. Clearly, if contemporary science gives us knowledge of the essences of natural kinds it does so by empirical rather than a priori means.

W. v. O. Quine argued that appeal to the notion of natural kinds can dissolve two famous paradoxes of confirmation:

(i) Carl Hempel's raven paradox: Intuitively a generalisation like 'all ravens are black' is confirmed by observation of its instances, in other words, by the observation of black ravens. Yet this generalisation is logically equivalent to 'all non-black things are non-ravens' which is confirmed by observation of, say, a green leaf. But how can observing a green leaf confirm 'all ravens are black'?

(ii) Nelson Goodman's grue problem: Suppose that observation of the instances of green emeralds confirms the generalisation 'all emeralds are green'; they are also instances of the generalisation 'all emeralds are grue', where 'grue' means 'green before 2030 and blue afterwards'. Why do we take the one generalisation to be confirmed and not the other?

It seems that there are some predicates, such as 'green', that we are prepared to use to make projections about unobserved objects and others, such as 'grue', that are not projectible. We make judgements of similarity in terms of projectible predicates. Goodman argued that projectible predicates are those that are entrenched in our epistemic community, and denied that similarity judgements among objects are objective.

Quine argued that appeal to natural kinds allows Goodman's problem to be avoided, because not all predicates are projectible and those that are will be those that are true of all and only the things of a kind. Similarly, the Ravens paradox is avoided if predicates like 'is non-black' are ruled out of consideration because they do not refer to natural properties. Clearly the notions of kind and similarity are closely related. The question of whether there are natural kinds is the question of whether there are objective similarities between things, which is the question of whether there are some properties that are natural and in virtue of which similarity and kind membership is definable. But Quine is troubled by this since he thinks that the notion of a kind and the related notion of similarity are of dubious scientific standing. Quine would like to be able to do away with any metaphysical commitments other than to sets of concrete individual things. However, natural properties seem to be intensional while sets are extensional. This is because sets that have all the same members are the same set, but all and only the same things may instantiate two nevertheless distinct properties. There seem to be sets corresponding to the kinds that there are (the set of all the things that belong to the kind), but not all sets correspond to natural kinds.

We use judgements of similarity to learn language not least because we must learn how to decide which similarities in sounds are relevant to meaning. We seem to have innate rankings of similarity. Psychological tests show that people will class a red circle as more similar to a pink ellipse than a blue triangle, even though both red and blue are primary colours but pink is not. The reason for this is the subject matter of evolutionary psychology and allied sciences according to Quine. In order to learn a language we have to map our judgements of similarity onto those of our neighbours, and Quine says, induction itself is essentially only animal expectation or habit formation (he is close to Hume in this and other respects), only now we have to match our judgements of similarity to the world in order to be successful in generalising about the behaviour of things, and we have evolved more and more sophisticated forms of pattern recognition and accuracy as a means of survival.

The notion of kind relates to those of disposition, counterfactual conditionals and causation. We assert counterfactual conditionals like 'if that had been put in water it would have dissolved' when the thing in question is of the same kind as things that did or will behave that way. Quine claims that the notion of kind is what links general and singular causal claims: the singular causal claim can be made because we recognise the events as being of the same kind as those that feature in general causal claims. However, Quine thinks of the notions of causation, disposition, counterfactual conditional and kinds as scientifically disreputable. He is against any form of intensionality, de re modality, natural properties or objective similarity relations, and he thinks that the notions of kinds and property must be accounted for without them. He argues that it is a mark of maturity of a branch of science that the notion of similarity or of a kind is eliminable in favour of the structural properties that give rise to them: there is no need to talk about solubility since we can just talk about the relevant properties of the atomic lattice; there is no need to talk about gold when we can just talk about atoms of atomic number 79. Hence, we progress from our crude 'animal' sense of similarity to a more sophisticated and fine-grained set of similarity relations refined by scientific experimentation. The fact that we can ultimately eliminate talk of kinds and dispositions means that such talk is exonerated. We need only invoke the basic properties of physics and chemistry. Quine then holds an extensional view of these properties according to which they are simply the sets of things that possess them. Whether that view is defensible is another question.

(II) Truth and Mind-independence

Science is widely regarded as our most reliable source of true beliefs about the world. Philosophers quickly disagree about the nature of truth however, and differing views about it inform positions in epistemology and methodology. There is even controversy about the right account of 'truth-bearers', the entities that are true or false. Some philosophers posit propositions as truth-bearers, where propositions are abstract entities that are expressed by sentences, whereas others assert that the only truth-bearers are particular utterances or written sentences. In what follows, the term 'proposition' will be used to mean simply anything that can be true or false.

It is often said that the difference between the truth of a scientific theory, and the 'truth' of a piece of music, is that if the latter is any kind of truth at all, it is truth about the subjective world of emotions, whereas scientific truth concerns matters of objective fact. Objectivity may be taken to be equivalent to mind-independence in the sense that, in general, the objective facts about the world are what they are independently of whatever people happen to believe or desire.

Of course, for lay people looking to science to tell them important truths about the world, there is often no practical difference between the beliefs that are counted as truths by the epistemic authorities at a given time, and the genuine truths. Some sceptically inclined philosophers have therefore suggested that truth is nothing more than a certain kind of legitimacy that is bestowed on beliefs by those with the power to do so. On this view, truth is a social construction in the sense that social processes determine which beliefs are true and which are false. This 'social constructivism' about scientific knowledge suggests that, for example, the Special Theory of Relativity is true because those in the scientific establishment who advocated it overcame the opposition from those who denied it. Hence, social constructivists think that the order of explanation in the history of science goes from social processes to theoretical and experimental facts and not the other way around. They identify what is true with what is believed to be true (by those with the epistemic power), much as Euthythro identified the pious with what is loved by the gods in the famous Platonic dialogue. Socrates disputes the latter identification as follows:

> But if the god-beloved and the pious were the same, my dear Euthythro, and the pious were loved because it was pious, then the god-beloved would be loved because it was god-beloved, and if the god-beloved was god-beloved because it was loved by the gods, the pious

> would also be pious because it was loved by the gods; but now you see that they are in opposite cases as being altogether different from each other; the one is of a nature to be loved because it is loved, the other is loved because it is of a nature to be loved. [1981, 16]

Many philosophers take the contrast drawn by Socrates between the pious and the god-beloved to be analogous to the contrast between the true and what is believed to be true by the scientific establishment. A proposition, such as that the Earth is much more than six thousand years old, is believed and legitimated by the scientific establishment, because it is true.

The view of truth that initially seems appropriate to capture the idea of scientific truth being about facts that are independent of whatever scientists believe is called the 'correspondence theory of truth'. According to it, for a proposition to be true is for it to correspond to the facts. Unfortunately, defenders of the correspondence theory of truth have run into all manner of difficulties, most of which have to do with the nature of the correspondence relation, and with the nature of the ontology that must be posited as the entities to which true propositions correspond. Some philosophers who start out naturally sympathetic to the idea of truth as correspondence become disillusioned with metaphysical theories developed to explain correspondence and which posit facts or states of affairs as existent entities over and above objects and properties.

In practice it seems that whenever we are asked what the evidence for some proposition is, we can do nothing but assert one or more further propositions. Despairing of the project of explaining truth as the relationship between propositions and reality, many philosophers have been inclined to locate truth entirely in the realm of propositions and to treat as a relation, namely coherence, among them. According to coherentism about truth, a proposition is true just in case it coheres with other propositions in a system, and false otherwise. Truth is not a relation of correspondence between beliefs and reality but an internal relation of coherence among a set of beliefs. Many philosophers who have argued for coherentism have been motivated by holist considerations. If for one reason or another it is held that individual beliefs cannot be directly compared to the world, then the only test available for the truth of a belief is whether it coheres with the rest of a system of beliefs. Another source of motivation for coherentism is the fact that whenever we seek to describe the structure of the facts of the world, we always rely upon the structure of our thoughts and sentences. It may be argued that our judgements never confront the world directly but instead further judgement, beliefs and statements. Kant argued that noumenal reality (Ding an sich or what William James later called 'trans-empirical reality') is not accessible by human intuition. If we can never describe the mind-independent world then correspondence between thought and noumenal reality is not possible, and we have to explicate truth in a way which does not make reference to it. The coherence theory of truth is also implicit in some of the writings of rationalists like Leibniz and Spinoza. Other coherentists about truth include those who think the world is purely mental or spiritual in nature (idealists) such as Hegel and Bradley, and some logical positivists such as Neurath and Hempel.

The definition of coherence above seems much too weak since there are surely many sets of beliefs that are internally coherent but are nonetheless not true. There are consistent fictions and fairy-stories. Coherence must mean more than non-contradiction. There are various responses to this. One of the most influential is to argue that consistency must be supplemented by another relation such as explanation. On this view, a set of beliefs is true just in case they are consistent and if they are mutually explanatory. On the other hand, perhaps even consistency is too much to ask of a set of beliefs, since in practice we probably all believe some inconsistent propositions, but it does not follow that all our beliefs are false.

In the face of abstract worries of this kind many philosophers and non-philosophers alike are inclined to try to ground discussion of philosophical questions in our practical lives. One good practical reason to believe what is true is that it is generally a more successful strategy in guiding action than believing what is false. It is often said that knowledge is power and if this is true it is in part because what is known is also true. This motivates the pragmatic theory of truth which crudely put states that what is true is what it is useful to believe, or alternatively, the truth is what works. For example, William James argued that truth is what is expedient in thought, just as the right or the good is what is expedient in behaviour. According to the pragmatists like James, the meaning of a concept is given by the practical or experimental (pragmatic) consequences of its application. Both he and Charles Peirce thought that any difference in meaning must make some possible difference in practice. This is fairly close to the logical positivists verification principle according to which a proposition that cannot be empirically verified is meaningless (unless it is a tautology).

Peirce argued that truth is that set of beliefs which followers of the scientific methods of inquiry will converge upon in the long run (this is often referred to as 'the ideal endpoint of inquiry'). He thought this based on his psychology according to which beliefs are dispositions to behave, and doubts are the negative effect on such dispositions which arise from unruly experiences which subvert our theories. Doubt prompts inquiry because it induces as unpleasant state in us which we try to overcome by seeking stable beliefs. Truth is just the maximally stable state of belief. Note that Peirce thought that truth entailed correspondence with reality but he did not think it consisted in correspondence with reality. James' view was a hybrid of coherentism and pragmatism since he held that our set of beliefs is gradually adjusted to accommodate awkward experiences, and that the limit of this process is truth.

Other philosophers have advocated forms of minimalism about truth, an example of which is the redundancy theory according to which the truth predicate is redundant and so "*p*' is true.' means exactly the same thing as '*p*'. There is also a view of truth known as the identity theory of truth according to which truth-bearers and truth-makers are identical. Finally mention must be made of the famous T-schema of Alfred Tarski according to which, for any proposition '*p*', '*p*' is true if and only if *p*. For example, 'snow is white' is true if and only if snow is white. Some have argued that this exhausts what can correctly be said about truth.

(III) Properties and Universals

The problem of universals is an ancient but perennial philosophical problem about the ontological status of properties. Particulars are entities that only exist in a single instance. They include individual things, but also events. Universals on the other hand are said to be multiply instantiable; they can be multiply realised in space and time. If there are any universals they include properties and relations. Prima facie it seems obvious that the world consists of individual things, and that they have properties and there are relations among them. It is also seems obvious that several things can have the same property. The problem of universals is about whether or not properties themselves are real and hence whether talk of the same property being had by different particulars should be taken literally. Hence, the problem of universals is about the 'One over Many'.

woman, man, woman

How many words are there here? There are either three or two depending on whether we count the types or tokens. Similarly if we had two red apples and two green apples, then we would have either two colours or four instances of colour. The sets of the red and green apples each seem to have a natural unity. The problem of universals has to do with how a many can count as a one. The problem of universals is closely related to the problem of explaining what distinguishes a natural class from an artificial class.

The theory of universals solves the problem of the one over many because universals are capable of being instantiated by more than one particular. This seems to suggest an answer to the problem of similarity judgements; such judgements are correct when the individuals instantiate the same universals (and if those universals include all their essential properties they are members of the same natural kind). As well as objective resemblance, universals have been posited to account for two other important phenomena, namely predication and abstract reference and the meaning of general terms. Predication is exemplified by the following sentence: 'Socrates is a man'. The subject of the sentence is referred to by a singular term, 'Socrates', that denotes a particular thing in the world. It might be supposed that the predicate 'is a man' must also denote something in the world, namely the universal 'Man'. Predication is then analysed by saying that the particular instantiates or participates in the universal. Universals seem to be needed for the predication of relations too, such as Socrates is older than Plato. Abstract reference is when we refer directly to a property or relation, as when we say, for example, red is a colour. Here the meaning of the general terms 'red' and 'colour' is that they refer to the universals Red and Colour.

According to the Platonic theory of forms, universals transcend the reality of concrete particulars in space and time, and indeed there may be universals that are not instantiated in the actual world. On the other hand, the Aristotelian theory of forms is incompatible with the existence of uninstantiated universals, and on that view all universals are found in space and time (they are imminent). Hence the Aristotelian view is compatible with naturalism (the view that everything that exists is revealed to us, if at all, by scientific enquiry), and physicalism (the view that everything that exists is physical). Platonists may suppose that universals can be known a priori whereas Aristotelians insist that universals may only be known a posteriori.

One puzzle about universals is whether or not the universal that corresponds to a property itself has that property. For example, is the universal Man a man? If the answer is yes, and we needed universals to explain what all the men have in common, then it seems we must posit a new universal to explain what the universal Man and all the men have in common (this is called 'the third man argument'). On the other hand, if the answer is no, then it is mysterious how the manliness of men can be explained by their instantiating a universal that is not itself manly. Another puzzle about universals concerns the relation of exemplification that particulars bear to universals which seems to lead to a regress since exemplification must itself be a universal.

There is a good deal of debate among those who do believe in universals about such matters as whether they are abundant — there is a universal for more or less every predicate, or sparse — there are only universals for predicates that name natural properties. Consider, for example, disjunctive predicates, such as 'is red or square': reasons for denying the existence of disjunctive universals include that there is no common feature or common causal power of objects that satisfy such predicates. Similar considerations count against negative universals. Some argue for the special place of physics in saying what universals there are and deny that predicates that refer to non-physical properties name universals.

A state of affairs (or fact) obtains when a particular instantiates a universal. Clearly, there is more to a particular instantiating a universal than the existence of the particular and the universal. So it seems that there is more to a state of affairs than its constituents: as the latter can be 'summed' in different ways they are not merely parts to the state of affairs as whole. Must we then admit that states of affairs exist over and above particulars and universals?

David Lewis influentially defended the idea of universals in the context of his realism about concrete possible worlds, which are distinct since spatio-temporally isolated from each other. According to him, 'actual' is an indexical expression like 'here'; the former picks out the world the speaker is in just as the latter picks out the place where the speaker is. (Even philosophers who do not like Lewis believe in concrete possible worlds find it useful to employ them as models in modal reasoning.) Lewis holds that any mereological sum of any of the objects populating a world is itself an object, and any class of objects is an object. It is then possible to identify properties with classes of possibilia (the property F is the class of all the actual and possible objects that have the property), and propositions with sets of possible worlds (a proposition p is the set of possible worlds at which

p is true). Lewis conceives of universals as repeatable entities, wholly present wherever they are instantiated; hence for Lewis properties and universals differ. Lewis argues that universals may do useful work in various areas of metaphysics, epistemology and the philosophy of mind and science. Unlike properties, universals are present in every world, even those in which they are uninstantiated. There are relatively few universals but very many properties. Take any set of objects: these immediately give rise to as many properties as there are members of the power set of that set (the set of all its subsets). Any two things, actual or possible, share an infinite number of properties, because there are an infinite number of classes of which they are both members. Universals on the other hand are supposed to mark objective resemblances among things, and group things together partly according to whether they share important causal powers. Ockham's razor rules out any further universals as idle and the genuine universals are determined by our best scientific theories, and objective resemblances among particulars are primitive and unanalysable.

Another solution to the problem of universals posits entities that are neither particulars nor universals but property-instances or 'tropes'. Those who deny the existence of universals are called nominalists. They face the problem of accounting for the phenomena described above.

(IV) Identity and Individuality

We are concerned here with numerical identity — being one and the same thing — as opposed to qualitative identity which means being the same with respect to all qualities. There are two fundamental aspects to the problem of elucidating identity, namely identity at a time, or synchronic identity, and identity over time, or diachronic identity. Issues about identity are closely connected to issues about individuation. Problems of individuation concern what it is, if anything, in virtue of which some particular object is the object it is and not any other.

Aristotle defended the theory (hylomorphism) that individuals are the combination of matter and form (properties). Later philosophers argued that individuals are nothing more than a bundle of properties. If the bundle view is to be defensible then it would seem some version of the principle of the identity of indiscernibles (PII) must be true. PII states that there cannot be things with all the same properties, or equivalently, that qualitative identity implies numerical identity. The converse, the indiscernibility of identicals is uncontroversial, since clearly the if a and b have different properties then they are different objects. (Confusingly, both these principles are sometimes called Leibniz' Law.) We can only state these principles by using second-order logic, which quantifies over properties as well as over objects:

PII: $\forall x \forall y[(\forall P(P(x) \leftrightarrow P(y)) \rightarrow (x = y)]$
Converse: $\forall x \forall y[(x = y) \rightarrow (\forall P(P(x) \leftrightarrow P(y))]$

PII has been the subject of much controversy, and even if it is true there is the further question of whether it should be regarded as a necessary or a contingent truth. It is absurd to state PII as the claim that 'two things having all their properties in common are identical', since if they really are identical we don't have two things at all but one thing with two names. However, we can state PII in another logically equivalent form (as above): it is not possible that there be two things that share all their properties

$$(\neg\Diamond(\exists x\exists y[\forall P(P(x) \leftrightarrow P(y))\&\neg(x = y)]).$$

Obviously, if two objects, a and b, are really distinct then a has the property of being identical with a, and the property of being different from b, and b lacks these properties. However, such properties amount to nothing more than that a and b are distinct. (The property of primitive thisness or self-identity is also called 'haecceity'). If identity and difference count as properties PII becomes totally trivial. However, the question of whether PII is true when properties are restricted to be qualitative ones is still interesting and important. Qualitative properties are those properties which can be instantiated by more than one object and do not involve being related to a particular object, for example, being red, being on a brown table, and so on. It is also worth considering whether PII is only true if extrinsic properties are considered. Roughly, an intrinsic property is one which an object may possess even if it is the only thing that exists, for example, mass, charge, height, etc. An extrinsic or relational property is one which is not intrinsic, for example, being in the Northern Hemisphere. Qualitative properties include both intrinsic and extrinsic ones. Recall there is a further distinction between accidental and essential properties; the former are properties that some object can have or not and still be the same object, like being red or being on the Earth, the latter are properties that an object has in virtue of it being what it is and which it must have, like perhaps being charged for an electron or being H_2O for water (see Natural Kinds).

So now we are considering a stronger form of PII namely:

$\forall x\forall y\forall P[(P(x) \leftrightarrow P(y)) \rightarrow (x = y)]$, where P ranges over qualitative properties only. Obviously, the only way to discover that two different things exist is to find out that one has a quality not possessed by the other, or else that one has a relational characteristic that the other lacks. The epistemological question of how it can be known that a is different from b is the question of distinguishability. Verificationists about meaning like the logical positivists argue that if two things possessed all the same qualitative and relational properties, that they were different would be unverifiable and hence meaningless.

It is interesting to note that Leibniz believed that if an individual x is really distinct from an individual y then there is some intrinsic, non-relational property F that x lacks and y has or vice versa. Hence, he held that $\forall x\forall y\forall P[(P(x) \leftrightarrow P(y)) \rightarrow (x = y)]$, where P ranges over non-relational, qualitative properties. This principle does not apply even to classical particles since the ones of a given kind are all supposed to be identical in all their properties except spatio-temporal ones (which are usually taken to be relational). Some have argued that quantum particles can be numerically distinct and nonetheless share all their qualitative properties so that PII is contingently false. Others argue that quantum particles are not individuals. There is also controversy about whether or not spacetime points obey PII.

Consider now the extra problem of identity over time (or genidentity): What is it for a thing, by which we shall understand a concrete particular, to persist, that is, for it to be the same thing at different times? (This question is particularly pressing when we consider the identity of a person over time.) The problem of identity over time arises because things change over time. There are at least two types of change: change in parts and change in properties. The worry is that either of these two types of change construed as change of an individual entity seems to contradict the indiscernibility of identicals. For example, imagine a banana that turns from green to yellow — how can numerically one and the same object possess incompatible properties? Similarly, suppose a table has a leg on it replaced — how can numerically one and the same object have different parts? How many properties or parts can change before one object becomes another?

Theories of persistence divide into two main forms: endurance and perdurance theories. The former have it that one and the same object is wholly present at different times, while the latter have it that what we call the same object at different times are really different temporal parts of the whole object which is extended in time and hence never wholly present at a particular time. So on the perdurantist view identity over time is not really numerical identity at all, and ordinary concrete particulars are in fact aggregates made up of different temporal parts, time-slices or stages. (Perdurantists are divided over the question of whether temporal parts are infinitely divisible or not.)

(V) Matter and Motion

There were ancient philosophers who argued for materialism. This is the view that all that fundamentally exists is matter (there is only one substance), and that there is no immaterial soul beyond the body, the human mind being nothing more than the product of matter in motion. Materialism was advocated by the atomists Leucippus and Democritus. Atomism is the view that matter is ultimately composed of very small objects that are indivisible into further parts, and that all change in the world is attributable to changes in the position (motion) of elementary particles in the void. The only properties that atoms have are their size and shape, and their states of motion. Plato held that matter was essentially illusory and that the real world was the world of universals. On the other hand, for Aristotle the forms (although immaterial) depend on the existence of individual substances, and matter is a central component of Aristotle's theory of the nature of individual substances. Any such thing, for example, a marble statue, consists of matter in some form (see the section on Universals above). Aristotle distinguished between natural and unnatural motion, arguing that unnatural motion always involved some external cause. This gave him a problem in explaining the motion of an arrow in flight some time after it has left the bow. The pre-Socratic philosopher

Parmenides argued that all change was illusory. His follower Zeno tried to support this doctrine by giving a series of arguments to show that motion, being a kind of change, is impossible. Zeno's paradoxes of motion are so-called because they are arguments from seemingly plausible premises to the (unacceptable) conclusion that there is no motion.

Motion may be thought of in two ways, namely as absolute or as relative. Relative motion is only defined with respect to some object or frame or reference, and so the same object has many different states of relative motion at the same time depending in relation to what its relative motion is being considered. For example, when a ball is flying through the air its motion relative to the ground will not be the same as its motion relative to a bird which is flying along next to it. The revolution in physics initiated by Galileo is principally about this difference between absolute and relative motion. In Galileo and Newton's physics only relative (constant) motion is observable, which explains why we don't observe the effects of the Earth's motion around the Sun. In Newtonian mechanics absolute motion has no physical effects, but absolute acceleration, which includes rotation, does have physical effects.

In the seventeenth century, the idea of explaining all natural phenomena in terms of matter in motion became a goal of many of those known as the mechanical philosophers. Locke used the image of a clockwork machine to illustrate the goal of natural philosophy as he saw it: The hands seem to move in a co-ordinated way and the chimes ring out the hours, half hours and so on as appropriate; this corresponds to the appearances of things, the observable properties of, say, a piece of gold. However, the clock has inner workings and this mechanism produces the outer appearance of the clock; similarly the gold has an inner structure that gives rise to its appearance. The goal of natural philosophy is to understand the inner mechanisms responsible for what we observe. The point about a clockwork machine is that the parts all work together in harmony, not because they are co-ordinated by mysterious natural motions or final causes, but because each of them communicates its motion with the part adjacent to it by contact. Mechanists explain the behaviour of things in terms of motions of the particles that compose them, rather than in terms of essences and 'occult forces'. Mechanics, in the hands of Galileo, Descartes and Newton in particular, became a mathematically precise science of matter in motion, and what happens as a result of collisions between bits of matter. (All of them adopted a principle of inertia, which states that a body continues in its state of motion unless a force acts to change it, so only changes in motion require an explanation.)

Newton's theory of gravitation was problematic because Newton offered no explanation for how the force of gravity was transmitted between bodies separated in space. It seemed that gravity was an example of the kind of action at a distance, which mechanist philosophers were trying to avoid. Fields were introduced into physics to solve the problem of action at a distance posed by the force laws of classical mechanics, namely Newton's law of Gravitation and the law of electrostatic attraction and repulsion (Coulomb's law). The prototype for the field was the optical ether, which was thought to be material. Classical fields were replaced by quantum fields, and special relativity introduced the idea that mass and energy were somehow equivalent and that the amount of mass possessed by a body depends on its state of motion relative to the frame of reference in which the mass is being measured. There is much controversy about whether quantum fields and particles can be considered as material. Contemporary physics seems not to describe the world in terms of matter in motion, but rather in terms of spacetime, and fields of potentiality and probability.

(VI) Causation

Causation is apparently fundamental to the scientific understanding of the world. The idea of causation is closely linked with the concepts of laws of nature, dispositions, natural kinds and properties, necessity and possibility, and subjunctive conditionals; all these notions are modal (and, as mentioned above, for some philosophers like Quine therefore dubious).

Aristotle described four types of causation: efficient, material, formal and final. For example, what is the cause of a statue of Socrates? The efficient cause is the sculptor's actions, the material cause is the marble the statue is made of, the formal cause is the idea that the sculptor has of the finished image based on Socrates' appearance, and the final cause is the end for which the statue is made, perhaps to celebrate the intellectual virtues of philosophy. Aristotelian ideas were the subject of much criticism in the Scientific Revolution. The mechanical philosophers argued that science should not search for final causes (teleology), and some went further and argued that such explanations were vacuous; rather science should concentrate on finding the efficient or material causes of phenomena. (Evolutionary biology seems to reintroduce teleology but it is usually claimed that this is legitimate because the teleological talk of function and design is eliminable in favour of efficient cause.) However, there were those philosophers that were sceptical of the notion of causation as it was deployed by materialist and mechanist philosophers. Berkeley famously argued that matter could not be a cause of anything because he identified causation with activity, action and agency, and Malebranche argued that only God could be a true cause.

Hume's empiricist analysis of causation is the starting point for contemporary debates. He questions whether causation has anything to do with necessity. His own theory of causation is sometimes called the regularity theory of causation, according to which instances of the relation A causes B usually have each of the following features:

> Events of type A precede events of type B in time.
>
> Events of type A are constantly conjoined in our experience with events of type B.
>
> Events of type A are spatiotemporally contiguous with events of type B.

> Events of type A lead to the expectation that events of type B will follow.
>
> Hume says that we have no 'impression' of a necessary connection.

Hume's analysis maybe okay for generic causation where A and B type events occur lots of times and A is regularly followed by B, but it looks unable to handle single case causation where A and B type events only occur once in the whole history of the universe, so if A causes B this cannot be reduced to a regularity. One option is to deny that there really is any single case causation in the world. Another is to modify Hume's account. An influential account that clarifies the relationship between causation and necessary and sufficient conditions is due to John Mackie who argued that a cause is what he called an 'INUS' condition. Consider, a fire (A) caused by a match (B). A caused B does not imply, either that A is sufficient for B, because the match alone would not have caused a fire if there had been no combustibles, nor necessary for B, because something else could have lit the combustibles. A is an Insufficient but Necessary part of a set of conditions which are together Unnecessary but Sufficient for B (so A is said to be an 'INUS condition' for B).

However, when we consider a particular instance of an event of type A causing an event of type B, it is not enough that A and B happen and A is an INUS condition for B because of two problems:

(i) Epiphenomena

This is where A is a side effect of whatever the causal process is that causes B. A will always occur when B is being caused by the process and so A will be an INUS condition for B but it will not be a cause. For example, the sound of the heart beating is not a cause of the circulation but it is an INUS condition for it.

(ii) Pre-emption

This is where A would have caused B but some other cause of B happens first. For example, a match would have started a fire but another match was lit first and started it.

Notice that when A causes B we are often inclined to say 'if A hadn't happened B wouldn't have happened'(#). There is a close connection between causation and counterfactual or subjunctive conditions. Another way of expressing (#) is by saying 'A is necessary in the circumstances for B'. But how do we pick out the right counterfactuals, in other words, which circumstances do we hold fixed? For example, suppose a match sits next to a matchbox, we are inclined to say that had the match been struck it would have lit, but why don't we say that had the match been struck it would have been damp? One solution is to hold fixed the circumstances around the time when the match was struck. But this will allow us to say had the match been struck it would have been picked up. The solution is to only consider the circumstances prior to the time of the match being on the table or being struck.

(iii) C is an INUS condition for E iff E is an INUS condition for C: problem of the direction of causation (also problem of simultaneous causation)

David Lewis influentially took up an idea from Hume in his proposal of a counterfactual account of causation. The idea is that if (If C hadn't happened then E wouldn't have happened) is true then it is also often true that C caused E. Spelling out precisely how to turn this into a full analysis of causation turns out to be a complex matter that depends crucially on the analysis of counterfactuals, and on finding a way of dealing with the problems of epiphenomena and pre-emption mentioned above, and the problem of overdetermination, where E would still have happened if C hadn't because some other cause would have taken over.

Finally, there are now many theories of probabilistic causation, since many philosophers believe that there can be genuine causes that do not guarantee the occurrence of their effects. Unfortunately, the simplest account of probabilistic causation, according to which a probabilistic cause must make its effect(s) more likely than not is not true. There are many examples of probabilistic causes whose effects are relatively improbable. It is more plausible to say that probabilistic causes must raise the chances of their effects from what they otherwise would have been, but even this claim turns out to be false.

(VII) Laws of Nature

Discovering the laws of nature and using them for the prediction and explanation of observed phenomena is one, if not the most important job of science. However, it is not always easy to tell what the laws of a particular science are because there seems to be no rule about when to call something a 'law' rather than a 'principle'. Laws sometimes take the form of simple universal generalisations, such as all metals conduct electricity, but more often they have a mathematical form like Kepler's laws of planetary motion. Sometimes laws seem to express deep facts about the unobservable causes of phenomena, like the law that expresses the relationship between the energy and frequency of radiation, whereas other scientific laws seem almost homely by comparison, such as the law that if a gas is kept at a constant volume and its temperature is increased then its pressure will rise. Other ideas associated with laws include those of generalisation, regularity, pattern, stable relationship, symmetry and invariance.

Here are some important different kinds of laws:

(i) laws of motion or state evolution over time such as Newton's second law and the Schrödinger equation

(ii) laws of co-existence that constrain what states of some system are mutually compatible such as the ideal gas laws and Pauli's exclusion principle

(iii) conservation laws, such as the law of conservation of energy

(iv) phenomenological laws that describe the observable phenomena in a particular system, such as the law of the pendulum, versus fundamental laws that purport to explain the underlying unobservable entities and processes, such as the laws of electromagnetism

(v) deterministic laws are those such that given the values of all physical properties at a given time, there is only one possible state of the system at any other time. Probabilistic laws are those that only provide probabilities for the state of the system at other times, like the half-life laws of radioactive substances.

What is a law of nature? There three broad answers to this question. The first is Humeanism which says that a law of nature is a special kind of regularity among properties, events and/or objects in the natural world. The second is necessitarianism which says that a law of nature is a relation of necessity between properties, events and/or objects in the natural world. The third is the sceptical position that there are no laws of nature, or at least that there is no objective distinction between laws and mere regularities.

According to the naïve regularity theory of laws, there is no good reason to think there is a difference between laws and accidents (c.f. Hume on causation and induction). On this view, it is a law that all As are Bs iff all As are Bs. If it is correct there are not any regularities that are not laws, nor are there any laws that are not regularities.

A single case occurrence, such as a cat being on a mat at some time, is a trivial kind of regularity. So is it a law of nature that the cat is on the mat at that time? Further problems arise with disjunctions of regularities, and with regularities involving disjunctive predicates and predicates like grue (see (I)–(ii)). Furthermore, vacuous regularities, such as all unicorns love television, are always true (since they are analysed as ‘for anything, if it is a unicorn then it loves television’ which is true if there are no unicorns). Ought these to be regarded as laws of nature? There also seem to be regularities that are not laws (for example, all the presidents of the USA in the twentieth century were men). On the other hand, there are cases of scientific laws that do not seem to satisfy the regularity account. For example, Newton’s first law, which applies to bodies not acted upon by any external forces, is not actually instantiated since there are no such bodies, but it does not seem to be vacuous.

More complex problems besetting the regularity theory of laws include explaining the connection between laws, inference and explanation. Laws are supposed to be explanatory, and to support inductive inferences, but regularities do not seem to be explanatory, nor it is obvious why inductive inferences to the truth of regularities based on the truth of some of their instances are justified. Laws are also closely related to counterfactuals, so for example, it seems that if it is a law of nature that all metals expand when heated, it is true to say of a piece of metal that was not heated, that if it had been heated it would have expanded. But ordinary regularities do not seem to entail counterfactuals in the same way;

for example, that all the coins in my pocket are silver, does not entail that if some copper coin had been in my pocket it would have been silver.

So not all regularities are laws. The sophisticated regularity theorist therefore places restrictions on what regularities are to be counted as laws. These come in two varieties: epistemic restrictions are so-called because it is our cognitive attitudes that determine which regularities are laws. On such views, laws are regularities that play are certain role in our theories, or else they are just regularities to which we attach some significance or importance. The main problem for this account is that it seems plausible that laws can be unknown. Which of the unknown regularities are laws and which are not can only be a matter of what our attitude to them would be if we knew them. This is obviously problematic since such counterfactuals would seem to rely upon laws themselves. What about a world with no minds? There would be no laws either it would seem but surely the laws of nature could have ruled out the possibility of there being minds? Why do we have different attitudes to different regularities? Either this is arbitrary or grounded in some objective difference between them. If the former then this is no good, if the latter then the epistemic view collapses into the systemic view.

The second kind of modified regularity theory places 'systemic restrictions' on which regularities count as laws (this view is associated with Mill, Ramsey, and Lewis). Laws are the propositions we would use as axioms if we knew everything and organised it as simply as possible in a deductive system. On this view, laws are the result of a trade-off between simplicity and strength. Laws are the theorems and axioms of deductive systems that achieve the best combination of simplicity and strength.

Problems with this view include the following:

(i) Arguably, neither simplicity nor strength is an objective notion.

(ii) What achieves the best balance of strength and simplicity may not be agreed upon by all; for example, rationalists might weight simplicity more, whereas empiricists might weight strength more.

(iii) It is possible that the most systematic laws would involve grue or disjunctive predicates.

(iv) It is possible that there be equally systematic but different sets of laws. (Cf. Coherentism about truth.)

(v) The problem of inference and explanation is not obviously explained by the systemic view.

The necessitarian account of laws of nature says that they are relations among universals. On this view, laws of nature differ from mere universal generalisations. Laws of nature express necessary relations among universals (these relations are 2^{nd} order universals). Laws are singular statements about universals not universal generalisations about particulars. Laws imply universal truths but universal truths do not always imply laws.

F-ness $\rightarrow$ G-ness (X-ness is the property of being X, for example, F is being an electron and G is being negatively charged)

'$\rightarrow$' is to be read as 'brings with it' (Dretske), 'nomically necessitates' (Tooley), 'necessitates' (Armstrong).

This approach seems to offer an account of how laws support counterfactual statements, and to deal with the relation between laws, explanation and inference. However, necessitarianism faces a number of further questions:

(i) Are law statements themselves necessary?

(ii) The identification problem: what exactly is the necessitation relation between universals?

(iii) The inference problem: how can we make sense of the inference from 'F-ness $\rightarrow$ G-ness' and 'this is F' to 'this must also be G' if the laws of nature are themselves contingent?

Probabilistic laws raise further problems for all the views discussed above.

Nancy Cartwright argues that phenomenological laws may be true but that fundamental laws are not since their application to the world always involves modelling, idealisation and approximation. She argues that causal powers are more fundamental than laws. On the other hand, Bas van Fraassen argues that there are no laws of nature and that they are features only of the theoretical representation of the world and not the world itself.

Ceteris paribus laws are laws that hold 'all things being equal'. Giving an account of the ceteris paribus clause that does not make the truth of the law trivial, by saying that other things are equal just in case the law is true, turns out to be a difficult task. It is thought by some that the difference between the natural and the social sciences lies in the fact that the former and not the latter are able to find exact laws.

(VIII) Probability, Propensity and Dispositions

The formal theory of probability was invented relatively recently in the history of science and mathematics, but the idea of probabilistic reasoning is commonplace. Probability may be thought to have nothing to do with ontology, but rather to be the science of uncertainty, evidence and estimation, and hence to be part of epistemology and not metaphysics. There are accounts of probability that do indeed claim that probability is an entirely epistemic notion, but to do so is to adopt a position analogous to nominalism in the lively debate about whether there is such a thing as objective chance. Since the advent of quantum mechanics it has been widely thought that it is at least an open question whether the world has fundamentally probabilistic occurrences and causes in it. Probability in the world that does not arise from our ignorance is 'objective chance'.

Objective chance has been identified with:

(i) Finite relative frequencies

(ii) Infinite relative frequencies

(iii) Propensities (these are primitive single case probabilities)

There are problems with all of the above. The finite relative frequency of some occurrence may occasionally depart radically from its probability. For example, if a fair coin is tossed ten times it may well come up heads seven times, yet intuitively the probability of heads is only 50%. Infinite relative frequencies are problematic because the notion of a completed infinity is problematic and transcends the empirical world. It is an interesting question how epistemic and objective probabilities must be related.

Note that determinism is the doctrine that given the state of the world at one time, and the laws of nature, there is only one possible way the world could be at all other times. Indeterminism is the denial of determinism. Determinism is a modal claim about the world rather than a claim about what can be predicted. It is possible for there to be phenomena governed by deterministic laws that we are nonetheless unable to predict. This is the case where very divergent outcomes follow from very small differences in initial conditions, since then the smallest inaccuracy in measurements of the latter will make accurate prediction impossible (this sensitivity to initial conditions characterises chaotic systems).

Dispositions are properties, such as fragility and solubility, that may or not be actualised. Some philosophers hold that dispositions must be reducible to the structural properties of things, while others hold that dispositions may be primitive. Dispositional essentialists argue that the essential properties of physical kinds are dispositional.

(IX) Reductionism, Emergence and Supervenience

There is a great deal of debate in philosophy of science about the relationship between the sciences. How are the domains of physics, chemistry and biology related, and how are the laws, theories and explanations of these sciences related?

Fundamental intuitions of reductionism include:

1. The whole is not greater than the sum of the parts.

2. The behaviour of the whole is caused and explained by the behaviour of the parts.

3. There is a unity to the world and to science.

Reductionism is popular because in general: reduction seems to yield explanatory gain (some theories of explanation assimilate explanatory power to unification); reduction implies ontological unification and so is in keeping with the desire for parsimony in metaphysics and accordance with Occam's razor; and finally, reduction aids conceptual unification.

Here are some examples of different forms of reductionism:

(a) Philosophical/logical behaviourism about the mind that reduces thoughts and other mental states to relations among stimuli and behaviour. This is inspired by verificationism (the idea that all meaningful discourse concerns what can be verified in experience) conjoined with the claim that we can only verify propositions about the mind by observing behaviour.

(b) Logicism about mathematics that regards mathematical theorems as consequences of logical laws.

(c) Set-theoretic reductionism that reduces all mathematical objects to sets. For example, the natural numbers can be identified with a sequences of sets where each successive set contains all the sets that have gone before it.

(d) Semantic reductionism about theoretical terms that reduces sentences involving them into sentences only involving observational terms and logical constants. The Logical Positivists attempted to explicitly define theoretical terms in terms of observational language. For example, 'temperature' would be translated into statements about observable manifestations of it, and statements about mind-independent objects would be translated into statements about observations.

(e) Reductionism about the mind according to which types of mental states are identical to types of brain states.

(f) Reductionism about dispositions according to which the latter are reducible to categorical or structural properties.

(g) Reductionism about colours and sounds according to which they are identical with physical properties.

(h) Reductionism about natural kinds according to which macroscopic kinds, like water, are identical with their microstructural essences (water is identical with H_2O).

Within science there have been reductionist programmes of great significance and some examples are listed below (the first three are intra-science, the rest are inter-science; (iv), (v) and (vi) are broad and programmatic/methodological)

(i) Galileo's law of freefall and Kepler's laws of planetary motion to Newtonian mechanics

(ii) optics to electromagnetism

(iii) thermodynamics to the kinetic theory of gases via statistical mechanics

(iv) laws in the social sciences to laws that only refer to individuals (methodological individualism), for example, laws about the behaviour of economic markets to rational choice theory

(v) social sciences to natural/physical sciences: socio-biology, evolutionary psychology, genetic reductionism

(vi) natural sciences to physics: geology to geophysics and geochemistry, neurophysiology - cell biology - molecular biology - molecular physics - quantum physics (the failure of vitalism/organicism, which posited a special status for living systems encouraged this kind of reductionism)

(vii) genetics to molecular biology

(viii) chemistry to quantum mechanics

There are various kinds of reductionism, notably semantic and having to do with meaning equivalence (a, b, c, d), and ontological (the rest of the above). In the case of the latter, translation is effected by means of 'bridge laws' which correlate terms in reduced theory's vocabulary to those in reducing theory's vocabulary. In the case of the former there must be strict identities between the terms in the reducing theory and the reduced theory.

There are various problems that may arise for reductionist programmes. One is that the bridge laws may turn out to be only partially true. Another is that the reduced theory is usually only approximately true and ends up being corrected rather than recovered exactly by the reducing theory. Reduction also usually relies heavily on idealisation. Finally the most celebrated problem is that of multiple/variable realisation; this is the fact that, for example, 'pain' seems to be realisable in animals with very different kinds of anatomies and physiologies, just as the same word processing programme can be realised by computers with very different internal workings. It is often said that multiple realisability means that mental events are only token identical with physical events, and not type identical with them, where the former means that each mental event is identical with some physical event but that each type of mental event need not be identical with the same type of physical event as the latter requires.

In the light of this, philosophers often think in terms of supervenience rather than reduction. A domain supervenes on another domain, if there can be no changes in the former without changes in the latter, but not necessarily vice versa. For example, arguably the mental state of a person cannot change without their brain state changing, but it is possible for their brain state to change in a way that does not affect their mental state. This is called local supervenience, whereas global supervenience is the claim that all the mental facts about the whole world supervene on the physical facts about the whole world. Dualism, for example, denies even global supervenience of the mental on the physical. On the other hand, emergentism is the doctrine that the whole is not reducible to the sum of the parts and that genuinely new properties and causal powers come into being when parts make up a whole.

The relationship between causation in physics and causation in the special sciences is much discussed in contemporary philosophy of mind. Causal exclusion reasoning proceeds along the following lines: Mental states must either be reducible to physical states, or cannot be the causes of actions, because, for any action A, since A is a physical event and as such, given the causal closure of the physical world, there is some set of physical causes that are sufficient for its occurrence (or at least to fix its objective chance).

Finally, there is the question of whether special science objects, for example, organisms, markets and people, can be identified with the mereological sums of physical objects. Some philosophers conclude that special science objects cannot be so identified and so do not therefore really exist, but realism about the special sciences seems at least if not more plausible than realism about physics. Other suppose that special science objects are individuated by their functional role and are only token identical with collections of physical objects. This is problematic in so far as such objects seem to be actually and not merely potentially multiply realised, for example, a given cat may be identified with numerous subsets of the maximal set of molecules that make it up, since for any set of the molecules that is a candidate for being token identical with the cat, removing a few molecules at random from this set will leave a new set that is also a candidate for being the cat.

(X) Space, Time and Spacetime

The Aristotelian theory of space grants a privileged position to the centre of the Earth, and this induces a privileged direction towards the centre of the Earth. Space is said to be absolute. The Galilean relativity principle entails that absolute position in space and absolute motion through it are physically undetectable. In Newton's theory of space absolute position is nonetheless an objective feature of the world and Newton also posited absolute time. Leibniz rejected the Newtonian ideas of absolute space and time and argued instead for the idea that space and time are nothing more than relations among phenomena. Leibniz appealed to the principle of sufficient reason and the PII to show that there was no such thing as absolute space. The former states that everything that occurs must have a sufficient cause; since position in absolute space and time make no observable difference to anything there could be no cause of why the universe begins in one position in space and time rather than another. PII is in conflict with absolute space and time since different positions for the whole universe in absolute space and time are qualitatively identical. These issues are now discussed in the context of general relativity.

There are several distinct, though often conflated issues in the metaphysics of time:

(i) Are all events, past, present and future, real?

(ii) Is there temporal passage or objective becoming?

(iii) Does tensed language have tenseless truth conditions?

‘Eternalism’ is the view that all times are real, whereas according to ‘presentism’ only the present is real (there is also the ’cumulative’ view that all past and present events are real). Those who believe in the passage of time or objective becoming often also believe that the process of becoming is that of events coming into existence and going out of existence, but this need not be so; to suppose there is becoming, one need only believe that there is some objective feature of the universe associated with the passage of time. Objective becoming could be like a light shining on events as they are briefly ‘present’, and is therefore compatible with eternalism. On the other hand, both presentism and cumulative presentism entail a positive answer to question (ii), since if events do come into existence, whether or not they then stay existent or pass out of existence, this is enough to constitute objective becoming. Presentism and becoming have also been associated with the idea that tensed language does not have tenseless truth conditions. However, this is not a necessary connection. So even though the standard opposition is between those who answer ‘no’ to (i), ‘yes’ to (ii), and ‘no’ to (iii) on the one hand (the defenders of McTaggart’s ‘*A*-series’), and those who answer ‘yes’ to (i), ‘no’ to (ii), and ‘yes’ to (iii) (the defenders of McTaggart’s ‘*B*-series’), a variety of more nuanced positions are possible.

There is a further celebrated question about time:

(iv) Does time have a privileged direction?

Clearly if (i) or (ii) are answered positively then that is enough to privilege a particular direction in time. However, eternalism and the denial of objective becoming are also compatible with time having a privileged direction, since there could be some feature of the block universe that has a gradient that always points in some particular temporal direction. For example, the entropy of isolated subsystems of the universe, or the universe itself, might always increase in some direction of time. Another well known possible source of temporal direction was proposed by Reichenbach who argued that temporal asymmetry is grounded in causal asymmetry: in general, the joint effects of a common cause are correlated but the joint causes of a common effect are uncorrelated.

However, it may be that no physical meaning can be attached to the idea of the direction of time in the whole universe, because no global time co-ordinate for the whole universe can be defined. This seems to be implied by special relativity. The status of time in special relativity differs from its status in Newtonian mechanics in that there is no objective global distinction between the dimensions of space and that of time. Spacetime can be split into space and time, but any such foliation is only valid relative to a particular inertial frame, which is associated with the Euclidean space and absolute time of the co-ordinate system of an observer. This seems to imply eternalism, since if there is no privileged foliation of spacetime, then there is no global present, and so the claim that future events are not real does not refer to a unique set of events. Furthermore, many have argued that, since special relativity implies the relativity of simultaneity, whether or not two events are simultaneous is a frame-dependent fact, and therefore there is no such thing as becoming.

On the other hand, special relativity is a partial physical theory that cannot describe the whole universe, even if there is good reason prefer it to its empirically equivalent rivals which some deny. The implications of general relativity for time are not clear. This is because the theory gives us field equations that are compatible with a variety of models having different global topological features, and different topological structures may have very different implications for the metaphysics of time. Clearly we must then turn to cosmological models of the actual universe, of which there are many compatible with the observational data. As yet there is no agreement about which of these is the true one. Highly controversial issues about quantum gravity bear on the question of whether spacetime will turn out to admit of a global foliation, and hence on whether absolute time is physically definable. Even if it does turn out to be definable, there remains the question of whether such a definition ought to be attributed any metaphysical importance.

Non-relativistic many-particle quantum mechanics does not directly bear on the philosophy of time since the status of time in the formalism is not novel in the same way as in relativity. However, it has often been argued that quantum physics is relevant to questions about the openness of the future, becoming, and the direction of time, because of the alleged process of collapse of the wavefunction. Since Heisenberg it has been popular to claim that the modulus squared of the quantum mechanical amplitudes that are attached to different eigenstates in a superposition represent the probabilities of genuinely chancy outcomes, and that when a measurement is made there is an irreversible transition from potentiality to actuality in which the information about the weights of the unactualised possible outcomes is lost forever. Hence, measurement can be seen as constituting irreversible processes of becoming that induce temporal asymmetry. However, quantum measurements need not be so understood. Furthermore, if there is no collapse, as in the Everett interpretation, then again there is no temporal asymmetry in quantum mechanics. The upshot seems to be that the status of the arrow of time in quantum mechanics is open.

There is also a vast literature about whether or not the second law of thermodynamics represents a deep temporal asymmetry in nature. The entropy of an isolated system always increases in time, and so this seems to be an example of the arrow of time being introduced into physics. If the whole universe is regarded as an isolated object, and if it obeys the second law, then it would seem that there is an objective arrow of time in cosmology. However, it is not clear what the status of the second law is with respect to fundamental physics. One possibility is that the second law holds only locally, and that there are other regions of spacetime where entropy is almost always at or very near its maximum. Even if thermodynamics seems to support the arrow of time, it is deeply puzzling how this can be compatible with an underlying physics that is time asymmetric. Conservative solutions to this problem ground the asymmetry of the second law in boundary conditions rather than in any revision of the fundamental dynamics. The most popular response is to claim that the law does indeed hold globally but that its so doing is a consequence of underlying time-reversal invariant laws acting on an initial state of the universe that has very low entropy. It is necessary to posit this because standard arguments in statistical mechanics that show that it is overwhelmingly likely that a typical state of an isolated system will evolve into a higher entropy state in the future, also show that it is overwhelmingly likely that the state in question evolved from a past state that had higher entropy too. A much more radical possibility is that the second law is a consequence of the fact there is a fundamental asymmetry in time built into the dynamical laws of fundamental physics. Given the outstanding measurement problem in quantum mechanics those who propose radical answers to problems in thermodynamics and cosmology often speculate about links between them and the right way of understanding collapse of the wavefunction. Roger Penrose, for example, suggests that gravity plays a role.

(XI) Events and Processes

Philosophers often think about the ontology in terms of what kinds of objects there are. So they ask whether there are only concrete objects or whether abstract objects also exist; they ask whether there are only the fundamental building blocks of the world (mereological atoms), or whether composite objects also exist, and they ask whether there are mental or spiritual objects, as well as physical ones. However, there are other influential accounts in metaphysics that hold that the world consists of entities that are partly temporal in nature, namely events or processes.

Donald Davidson influentially argued that the world consists of events, and that properties like colour and shape are properties of events not of objects (or at least that objects are arrangements or structures of events, rather than the other way round). For Davidson the relata of causal and lawlike relations are events rather than objects or facts. On the other hand, consider what physicist Lee Smolin says: "The universe is made of processes, not things" [2001, 49]. Smolin insists that a lesson of both relativity theory and quantum mechanics is that processes are prior to states. Classical physics seemed to imply the opposite because spacetime could be uniquely broken up into slices of space at a time (states). Relativity theory disrupts this account of spacetime and in quantum mechanics nothing is ever really still it seems, since particles are always subject to a minimum amount of spreading in space and everything is flux in quantum field theory within which even the vacuum is the scene of constant fluctuations.

2 EPISTEMOLOGICAL POSITIONS

(I) Rationalism

Philosophers described as rationalists include Plato, Descartes, Leibniz, and Spinoza. Rationalism is associated with two distinct but often conflated theses. The first is that some of our concepts (ideas) are innate ((vi) below); the second is that some of our knowledge of the world is a priori, that is justified independently of experience or empirical evidence ((v) below). Note that the a priori/a posteriori distinction is an epistemological one. Other related distinctions include the metaphysical distinction between what is necessary (could not have been otherwise), such as that 2+2=4, and contingent (could have been otherwise), such as that the largest mammals are blue whales; and the semantic distinction between the analytic (true or false in virtue of meaning), such as that all bachelors are unmarried, and the synthetic (true or false not merely in virtue of meaning), such as that Paris is the capital of France. Of course, these categories often overlap, for example, that bachelors are unmarried may well be analytic, necessary and a priori, and that Paris is the capital of France may well be synthetic, contingent and a posteriori. However, whether or not this overlap is partial or total is one of the central issues that divides rationalists and empiricists.

Some characteristic doctrines of rationalism (although not held by all rationalists) are as follows:

(i) Sensory knowledge is limited and we should be cautious about it and use reason correctly to overcome these limitations.

(ii) The universe is ordered and accessible to the rational mind.

(iii) Mathematics is general, and Euclidean geometry in particular, provides the model of well-founded and unified system of knowledge. The subject matter is intrinsically clear and knowledge of it is certain and based on reason.

(iv) Basic beliefs (or at least some) are known a priori by the use of pure reason / understanding.

(v) There is a faculty of rational intuition that delivers substantive a priori knowledge.

(vi) Concept innatism: some concepts are not derived from experience, for example those of event, cause, location, time, extension, and substance.

(vii) There are necessary connections in nature. The truth in science and philosophy must refer to what could not be otherwise.

There are various arguments for rationalism. Rationalists claim that certain concepts cannot be derived from experience because nothing that we perceive exemplifies them; for example, identity, equality, perfection, God, power, and cause. Descartes famously argued that even our concept of matter must be a priori. He considers a piece of wax that is heated and so changes its shape, its colour and its other sensible properties. He argues that since we continue to think of it as the same wax, we must be thinking of it as matter or pure extension in space, and that we have no direct experience of it as such and must therefore apply that concept to the world by the use of reason alone. Rationalists also argue that knowledge of the laws of logic (for instance, the law of identity states that everything is identical to itself) that describe which inferential connections among our beliefs are valid, and of mathematics, that apparently describe necessary truths about an abstract realm of mathematical objects, could only be known a priori. Some rationalists argue that metaphysics is knowledge of a priori necessary truths, for example, that every event has a cause or that an object cannot be in two places at the same time. They maintain that such truths, if truths they be, cannot be known by experience. Other domains of possible a priori knowledge include probability theory, decision theory (the theory of action) and mereology (the logic of part/whole relations).

Consider Euclidean geometry. There are primitive and undefined terms such as 'point' and 'line' and then there are a few axioms relating them, such as that any two points define a straight line. The former are alleged by rationalists to be innate (they are examples of Descartes 'clear and distinct ideas'), while the latter are supposed to be self-evident, in the sense that if one entertains the proposition in question one will thereby come to believe it. The rest of the theory is arrived by the use of proof, and the rationalist notions of clear and distinct perception (Descartes) and the 'natural light of reason' (Leibniz) are associated with the state of mind one is in when following a mathematical proof. Thinking about knowledge in terms of the paradigm of the axioms and theorems of Euclid leads naturally to a view about knowledge and justification called foundationalism. This view of knowledge goes back to Aristotle's *Posterior Analytics*, and is attractive because it offers a clear way out of the following famous sceptical problem known as Agrippa's Trilemma, or the Regress Argument: To be justified in believing something is to have a reason for believing it, but then one must have a justification of that reason, and so on ad infinitum. The idea is that this sceptical regress is halted with the intellect as the source of immediate and certain knowledge of foundational truths, upon which the rest of our knowledge is based.

Foundationalism says that there are *basic beliefs* which are justified independently of all other beliefs/non-inferentially justified. There are certain propositions that we seem to be justified in accepting but where that justification does not depend upon our acceptance of any other propositions, for example the aforementioned axiom that two points define a line. According to foundationalism, all justified beliefs are either basic or justified by being supported by basic beliefs, and justification is a 'one-way' relation. On this view, non-basic beliefs are deductively inferred from basic beliefs, and since deduction is truth-preserving, justification is assured. Basic beliefs are supposed by rationalists to be self-evident in the sense that if p is self-evident then if someone entertains it he or she will believe it. They are also required to be indubitable (not capable of being doubted) and incorrigible (not capable of being corrected by further experience). It is important to note however that empiricists may be foundationalists too: the proposition describing the immediate content of one's experience might be thought to be indubitable and incorrigible (although not self-evident).

By the seventeenth century rationalism was discredited in the eyes of many because of the failure of Aristotelian science, since the latter was widely regarded as overly reliant on reason at the expense of experience. Natural philosophers argued that certain knowledge of essences of things, or of substantive necessary truths about the world is not possible. When we consider examples of a priori knowledge, the propositions in question are often either questionable, or they seem to be true just in virtue of the meanings of the terms involved (analytic). Critics of rationalism argue that while there may be some a priori truths, there are no synthetic a priori truths. However, there is still some controversy among contemporary philosophers about whether thought experiments might offer a path to a priori knowledge in science.

(II) Empiricism

Classically empiricism is associated with Locke, Hobbes, Berkeley, and Hume. Empiricists tend to deny the existence of innate concepts and claim instead that the mind is a 'tabula rasa' at birth, and that all ideas are derived from experience. Experience either directly provides us with concepts via sensation, or indirectly via reflection and abstraction. Concepts and ideas are divided into the simple and complex, and the complex ones may not be derived from experience directly but rather composed of simple ideas. Empiricists also argue that there is no innate or a priori knowledge of the world. Rather all knowledge of reality is arrived at directly from particular experiences, or by extrapolating and generalising on the basis of experience.

Empiricists cannot consistently claim to know the truth of empiricism a priori, so they must argue for it on the basis of experience. The emerging natural philosophy led empiricists to make their model of knowledge not mathematics but experimental science. Francis Bacon was an important advocate of a new method of inquiry based on experiment; his vision of *New Atlantis* inspired the creation of *The Royal Society of London for the Improving of Natural Knowledge* (1660-). It is also possible to argue for empiricism from the implausibility and failure of rationalism. In the seventeenth century there were plenty of examples of embarrassing failures of science that was based on pure reason rather than experience and experimentation. The classic examples were Aristotle's theories of motion and cosmology that had been undermined by Galileo and Kepler. The idea of natural philosophers using their reason and intellect to apprehend the forms or essences of substances and processes in nature was discredited. Furthermore, empiricists can point out that there is no guarantee (at least for atheists) that a falsehood will not be self-evident, obvious, indubitable and clearly and distinctly perceived, for example, people might think it is self-evident that the earth is flat and doesn't move.

Hobbes thought he had squared the circle, so even the following of mathematical and logical proofs is subject to errors of reasoning.

Empiricists argue that pure reason cannot produce any useful or substantive knowledge of the world but only of the relations among our concepts. All a priori knowledge is of analytic truths, that is things that are true by definition and that tell us only about how we use words and concepts. This doctrine is often called Hume's Fork: all enquiry is about either, propositions about the 'relations of ideas' that are knowable a priori, for example, mathematics and logic, or, propositions about 'matters of fact and real existence' that are knowable only a posteriori, for example, physics and chemistry. Empiricists often add that all synthetic propositions are contingent and that since only analytic truths are necessary, the only necessity is verbal necessity.

In the nineteenth century an important empiricist movement called positivism came to prominence. The defining characteristic of positivism is that it is extremely in favour of science and opposed to metaphysics and theology. Positivists were also influenced by Hume in their disdain for ideas of necessitation or causation in nature, and their concern with ensuring the meaningfulness of language through an emphasis on verifiability or falsifiability. They also denied the existence of unobservable entities such as atoms.

Logical positivism was a movement of empiricist philosophers (associated with the Vienna Circle), in the twentieth century who used the new methods of mathematical logic to defend many of the traditional tenets of positivism. The logical positivists held that:

(i) Science is the only form of proper knowledge

(ii) All truths are either: (a) analytic, a priori and necessary, or, (b) synthetic, a posteriori and contingent

(iii) Logic is the science of elucidating the relationships among concepts.

(iv) The purpose of philosophy is to explicate the structure or logic of science.

(v) The verifiability criterion of meaning: A statement is held to be literally meaningful if and only if it is either analytic or empirically verifiable.

(vi) The Verification Principle: The meaning of a statement is its method of verification (except tautologies), that is the way in which it is shown to be true.

(vii) Metaphysical propositions are not verifiable and hence not meaningful.

In the light of (iv), the logical positivists held that epistemology just is the philosophy of science. Their projects included:

(a) the analysis of the meanings of theoretical terms in terms of observations or experiences — this is often referred to as operationalism

(b) the explication of the 'logic of confirmation', that is how evidence can confirm a hypothesis or theory

(c) show that a priori knowledge of mathematics and logic is compatible with the verification principle by showing that mathematics is reducible to logic and that logic is analytic.

(b) is discussed in the (III) of 3. Methodological Positions. (c) is beyond the scope of the present work. With respect to (a), an example of such a definition of a theoretical term V_T is:

$$\forall x(V_T x \leftrightarrow [Px \rightarrow Qx])$$

where P is some preparation of an apparatus (known as a test condition) and Q is some observable response of it (so P and Q are describable using only V_O terms). For example, suppose it is the explicit definition of temperature; any object x has a temperature of t iff it is the case that, if x is put in contact with a thermometer then it gives a reading of t. If theoretical terms could be so defined, then this would show that they are convenient devices that are in principle eliminable and need not be regarded as referring to anything in the world (this view is called 'semantic instrumentalism').

It was soon realised that explicit definition of theoretical terms is highly problematic. Perhaps the most serious difficulty is that, according to this definition, if we interpret the conditional in the square brackets as material implication, theoretical terms are trivially applicable when the test conditions do not obtain (because if the antecedent is false the material conditional is always true). In other words, everything that is never put into contact with a thermometer has temperature t.

The natural way to solve this problem is to allow subjunctive assertion into the explicit definitions. That is we define the temperature of object x in terms of what would happen if it were to be put into contact with a thermometer; temperature is understood as a dispositional property. Unfortunately this raises further problems. First, unactualised dispositions, such as the fragility of a glass that is never damaged, seem to be unobservable properties, and they give rise to statements whose truth conditions are problematic for empiricists, namely counterfactual conditionals such as 'if the glass had been dropped it would have broken' where the antecedent is asserted to be false. Dispositions are also modal, that is they pertain to possibility and necessity, and empiricists since Hume have disavowed objective modality. Like laws of nature and causation, dispositions are problematic for empiricists. Secondly, explicit definitions, dispositional or not, for terms like 'spacetime curvature', 'spin' and 'electron' have never been provided and there are no grounds for thinking that they could be.

When it comes to knowledge, many of the logical positivists initially adopted foundationalism about knowledge and justification but they take the foundations of knowledge to be immediate knowledge of our own sensory / perceptual states. The immediate objects of experience are called sense-data or the given, and so it was thought that the foundations of knowledge were to be given in terms of sense-data reports, which are also called protocol statements or basic proposition. These are first person singular, present tense, introspection reports and as such are supposed to be non-inferential, non-necessary, indubitable and incorrigible, and to refer solely to the content of a single experience.

One problem that the logical positivists faced was that of showing how knowledge of other minds and the public world could be built up from knowledge of private sense-data and analytic truths. Phenomenalism is the attempt to solve this problem by reducing all knowledge to knowledge of protocol statements and necessary truths: on this view physical objects are nothing but logical constructions out of actual and possible sense-experiences. Propositions asserting the existence of physical objects are analytically equivalent to ones asserting that subjects would have certain sequences of sensations in certain circumstances. A physical object is a permanent possibility of sensation (Mill).

Other problems concerned the status of the verification principle given that it appears to be neither empirically testable nor analytic, the fact that observation is theory-laden in the sense that all descriptions of observations involve interpretation and classification, and finally the problem of elucidating the logic of confirmation in the face of the problem of induction.

(III) Induction

In the broadest sense induction is any reasoning that is not deductively valid. In a narrower sense it is reasoning to a conclusion about unobserved cases on the basis of observed cases. There is also an even more narrow sense of induction that refers to inferences from finite sets of data to a universal generalisation; this is enumerative induction. The most general problem of induction is to explain when and how ampliative reasoning can be justified. The more specific problem is to explain how reasoning based on knowledge of unobserved cases can be a source of knowledge about unobserved cases.

Hume's problem of induction begins with the observation that all such reasoning is based on our knowledge of cause and effect. Given his analysis of causation, knowledge of cause and effect can only be knowledge that some regularity in has held in the past. Hence, induction is based on the assumption that the behaviour of things in the past is a reliable guide to their behaviour in the future, in other words it is based on the idea that nature is uniform in this respect. Hume then points out that the only reason we have for thinking that nature is uniform in the sense that the past is a good guide to the future is that in the past the past was a good guide to what was then the future. Hence, the justification of induction turns out to depend on circular reasoning and is therefore no justification at all.

There are a number of purported solutions and dissolutions of the problem of induction.

(a) Induction is rational by definition (analytic justification). The idea here is to argue that it is part of the ordinary meaning of the term 'rational' that inference from observed cases to unobserved cases can be rational.

(b) Hume is asking for a deductive defence of induction which is unreasonable. The claim is that just because induction cannot be deductively justified that does not mean induction is not justified.

(c) Induction is justified by the theory of probability. The idea is to construct an inductive logic by analogy with deductive logic. There have been some partial successes in this programme but it is generally agreed that they do not solve the problem of induction. The best that can be said to have been achieved is to show that if any form of non-ampliative rules of reasoning are to be employed then it is best to adopt standard induction. (This is sometimes called the pragmatic defence of induction and is associated with Reichenbach.)

(d) Induction is justified by a principle of induction or of the uniformity of nature. This principle could be claimed to be known a priori, since the claim that we know it a posteriori is denounced by Hume as circular.

(e) Hume's argument is too general. Since it does not appeal to anything specific about our inductive practices, it can only be premised on the fact that induction is not deduction.

(f) Induction is really (a species of) inference to the best explanation (see (IV)), which is justified.

(g) There really are necessary connections and we know that there are such. (It is often claimed that we know this by inference to the best explanation.)

(h) Induction can be inductively justified after all, because even deduction can only be given a circular (in other words, deductive) justifications.

(i) It may be agreed that induction is unjustified and an account of knowledge, in particular scientific knowledge, may be offered which dispenses with the need for inductive inference.

Note that these strategies may be combined.

(IV) Scientific Realism

Realism in the general sense has many faces, and this goes for scientific realism too. Critics differ as to which part of scientific realism it is to which they object, so there is a bewildering complexity of positions. In some contexts the significance of scientific realism is its commitment to the progressive and convergent nature of scientific inquiry and the privileging of the cognitive outcomes of that inquiry. This is a point of contention with some critics of science, but in recent times the debates about scientific realism in analytic philosophy of science have not questioned the success nor indeed the progress of science. Hence, for example, Bas van Fraassen, Larry Laudan, and Arthur Fine will all agree about the rationality of the scientific enterprise and its cumulative production of instrumental knowledge, even though none of them would be happy to be called a scientific realist. Scientific realists are united (and divided from sceptics) in their belief that scientific theories embody real knowledge about the world that goes beyond the observable, and further that the unobservable entities to which scientific theories refer really exist. If scientific theories are sets of statements, including laws, about observable phenomena and unobservable entities, processes and structures, then scientific realism claims that these are approximately true. Hence, according to scientific realism, scientific claims about electrons, quarks, spacetime curvature, and the energy of the vacuum are more or less true, and there really are such things to which these claims refer.

Notice that it has been implicitly supposed above that the language of science is to be taken literally pace the verificationist tradition that attempted to reconstrue theoretical talk as code for complicated sets of conditionals connecting observables. That such a project fails is taken for granted by all the main protagonists in contemporary philosophy of science. However, instead of reconstruing theoretical terms some antirealists reconstrue truth for claims about theoretical entities. So, we have not yet adequately characterised scientific realism. A social constructivist, for example, could assent to all the above, since they need not deny that theoretical terms refer nor that theories are true, but they may insist that truth is internal to our epistemic norms and practices, and that the entities to which we refer are socially constructed. There is no restriction to noncognitivist conceptions of truth in what has been said so far. This raises the question of to what extent a stand on such philosophical issues in defending scientific realism. For some, scientific realism simply amounts to the commitments at the end of the foregoing paragraph. However, usually scientific realists go further and commit themselves to the following claims

(i) the entities or kinds of entities talked about and/or described by theoreticalscientific discourse exist

(ii) their existence is independent of our knowledge and minds

These are the metaphysical requirements.

(iii) the statements of theoretical scientific discourse are irreducible/ineliminable and are genuinely assertoric expressions

(iv) the truth conditions for the statements of theoretical scientific discourse are objective and determine the truth or falsity of those statements depending on how things stand in the world.

These semantic requirements are often cashed out in terms of a correspondence theory of truth, as opposed to a pragmatic or a coherence theory of truth.

(v) truths about theoretical and unobservable entities are knowable and we do in fact know some of them, and hence the terms of theoretical scientific discourse successfully refer to things in the world.

This is the epistemic requirement.

For example, if we are considering electron theory then scientific realism says that:

(i) electrons exist

(ii) mind-independently

(iii) statements about electrons are really about subatomic entities with negative charge, spin 1/2, a certain mass, and so on

(iv) these statements are true or false depending on how the world is

(v) we should believe electron theory and much of it counts as knowledge

So standard scientific realism involves three kinds of philosophical commitment: a metaphysical commitment to the existence of a mind-independent world of observable and unobservable objects; a semantic commitment to the literal interpretation of scientific theories and a correspondence theory of truth; and finally an epistemological commitment to the claim that we can know that our best current theories are approximately true, and that they successfully refer to (most of) the unobservable entities they postulate which do indeed exist. To be an antirealist it is only necessary to reject one of these commitments, and antirealists may have very different motives, so there are a variety of antirealist positions which we ought now to be able to distinguish: Sceptics deny (i), reductive empiricists deny (iii), social constructivists deny (ii), while constructive empiricists like Bas van Fraassen deny only (v), but also don't believe or remain agnostic about (i).

(V) The Duhem-Quine Problem and Underdetermination

It is natural to suppose that scientific theories or hypotheses are tested by predictions being deduced from them. Then the appropriate experiment is performed and if the prediction agrees with what is observed then the theory or hypothesis is confirmed and if not it is falsified. However, in practice it is not possible to deduce statements about what will be observed from a single hypothesis. Rather, hypotheses have to be conjoined with background assumptions about the initial conditions of the system(s) in question, the reliability of the measurement procedures, and other relevant facts. Hence Pierre Duhem argued that experiments cannot confirm or falsify individual laws or hypotheses but only a whole collection of them. Consider the experimental test of Newton's law of gravitation by the observation of the path of a planet. The law of gravitation alone will not issue any prediction without values being given to variables representing the mass of the planet, the mass of the other planets in the solar system and the Sun and their relative positions and velocities, the initial position and velocity of the planet relative to the other planets and the Sun, and the gravitational constant. Newton's other laws of motion will also be needed. Once we have a prediction then we can observe whether it is confirmed or falsified, but suppose that it is the latter that occurs; which of the laws and assumptions we have made should we regard as being falsified? Perhaps none of them have been falsified because a mistake was made in the observation. Hence, Duhem argued that science must be treated as whole when it comes to testing it and considering the evidence for it, because no part of science on its own has determinate empirical content. This is often referred to as 'confirmational holism'.

Quine went further and argued that in principle even mathematics and logic, the laws of which must be used in deriving predictions from scientific theories, must be included in the whole that is confirmed or falsified by the experimental data. Quine argued that it would be reasonable to reject a law of logic, or change the meaning of our terms, if it was more convenient than rejecting a particular theory. Quine therefore rejected the distinction between analytic and synthetic truths that Hume, Kant and the logical positivists believed to be fundamental to epistemology. A trivial example of such a change in the meaning of a term is that of the change in meaning of 'atom' which once meant something indivisible and now refers to a particular type of collection of smaller particles. When physicists discovered that atoms were divisible, they redefined 'atom' rather than abandoning the term altogether.

The Duhem-Quine problem is that no part of science seems to be testable individually, and that therefore it is never possible to say that the a particular hypothesis or law is confirmed or falsified. In practice of course scientists do locate confirmation and falsification at the level of individual hypotheses. Duhem thinks that they use good sense to do so, but that this faculty and the basis on which such judgements are made cannot be fully characterised. Quine is a pragmatist and accepts that scientific knowledge is ultimately conventional. The Duhem-Quine problem is closely related to another problem that particularly undermines scientific realism, namely the underdetermination problem. There are two generic forms of underdetermination, namely weak and strong.

(i) Weak underdetermination:

1. Some theory, T is supposed to be known, and all the evidence is consistent with T.

2. There is another theory $T\#$ which is also consistent with all the available evidence for T. T and $T\#$ are weakly empirically equivalent in the sense that they are both compatible with the evidence we have gathered so far.

3. If all the available evidence for T is consistent with some other hypothesis $T\#$, then there is no reason to believe T to be true and not $T\#$.

Therefore, there is no reason to believe T to be true and not $T\#$.

This kind of underdetermination problem is faced by scientists every day, where T and $T\#$ are rival theories but agree with respect to the classes of phenomena that have so far been observed. What scientists try and do to address it is to find some phenomenon about which the theories give different predictions, so that some new experimental test can be performed to chose between them. For example, T and $T\#$ might be rival versions of the standard model of particle physics which agree about the phenomena that are within the scope of current particle accelerators but disagree in their predictions as to what will happen at even greater energies. The weak underdetermination argument is a form of the problem of induction: T is any empirical law, such as all metals expand when heated, and $T\#$ states that everything observed so far is consistent with T but that the next observation will be different. This form of underdetermination does not undermine scientific realism in particular since it does not entail or rely upon any epistemic differentiation between statements about observables and statements about unobservables.

(ii) Strong underdetermination:

To generate a strong underdetermination problem for scientific theories, we start with a theory H, and generate another theory G, such that H and G have the same empirical consequences, not just for what we have observed so far, but also for any possible observations we could make. If there are always such strongly empirically equivalent alternatives to any given theory, then this might be a serious problem for scientific realism. The relative credibility of two such theories cannot be decided by any observations even in the future and therefore, it is argued, theory choice between them would be underdetermined by all possible evidence. If all the evidence we could possibly gather would not be enough to discriminate between a multiplicity of different theories, then we could not have any rational grounds for believing in the theoretical entities and approximate truth of any particular theory. Hence, scientific realism would be undermined.

The strong form of the undetermination argument for scientific theories is as follows:

1. For every theory there exist an infinite number of strongly empirically equivalent but incompatible rival theories.

2. If two theories are strongly empirically equivalent then they are evidentially equivalent.

3. No evidence can ever support a unique theory more than its strongly empirically equivalent rivals, and theory-choice is therefore radically underdetermined.

Some who accept this argument adopt conventionalism according to which the choice among empirically equivalent rivals is a pragmatic one that involves freely chosen conventions based on simplicity and convenience. However, there are various ways of arguing that strong empirical equivalence is incoherent, or at least ill-defined:

(a) The idea of empirical equivalence requires it to be possible to clearly circumscribe the observable consequences of a theory. However, there is no non-arbitrary distinction between the observable and unobservable.

(b) The observable/unobservable distinction changes over time and so what the empirical consequences of a theory are is relative to a particular point in time.

(c) Theories only have empirical consequences relative to auxiliary assumptions and background conditions. So the idea of the empirical consequences of the theory itself is incoherent.

Furthermore, it may be argued that there is no reason to believe that there will always, or often, exist strongly empirically equivalent rivals to any given theory, either because cases of strong empirical equivalence are too rare, or because the only strongly empirically equivalent rivals available are not genuine theories (against this it is often claimed that quantum physics gives genuine examples of empirical equivalence). Whether or not any of these objections to (1) works, many scientific realists argue that (2) is false. They argue that two theories may predict all the same phenomena, but have different degrees of evidential support. In other words, they think that there are non-empirical features (superempirical virtues) of theories such as simplicity, non-ad hocness, novel predictive power, elegance, and explanatory power, that give us a reason to chose one among the empirically equivalent rivals. Some philosophers agree that superempirical virtues break underdetermination at the level of theory choice, but argue, following van Fraassen, that their value is merely pragmatic, insofar as they encourage us to chose a particular theory with which to work, without giving us any reason to regard it as true. This may motivate the conclusion that science can never give us knowledge of the unobservable world, and that our best scientific theories are empirically adequate rather than true. Strong empirical equivalence shows that theories have extra structure over and above that which describes observable events, so clearly belief in empirical adequacy is logically weaker than belief in truth simpliciter. Note however, that even if the choice among competing ways of embedding empirically equivalent substructures in fundamental theory is a pragmatic one, ultimately different formulations may lead naturally to the discovery of new laws. For example, Newton's force law suggested the mathematical form for Coulomb's law.

The problem that critics of scientific realism, who are not also inductive sceptics, face is how to overcome the weak underdetermination argument. It may be argued that the same superempirical considerations that entitle us to regard a well-tested theory as describing future observations as well as past ones, also entitle us to choose a particular theory among strongly empirically equivalent ones. The particular strong underdetermination problem for scientific realism is that all the facts about observable states of affairs will underdetermine theory-choice between T_0, a full realistically construed theory, and T_1, the claim that T_0 is empirically adequate. However, all the evidence we have available now will underdetermine the choice between T_1 and T_2, the claim that T_0 is empirically adequate before the year 2010 (the problem of induction). Furthermore all the facts about all actually observed states of affairs at all times will underdetermine the choice between T_1 and T_3, the claim that T_0 describes all actually observed events. So, even the judgement that T_0 is empirically adequate is underdetermined by the available evidence, and hence, the advocate of the underdetermination argument against scientific realism must be an inductive sceptic in the absence of a positive solution to the underdetermination problem.

(VI) Inference to the Best Explanation

Inference to the best explanation (IBE) is a (putative) rule of inference according to which, where we have a range of competing hypotheses all of which are empirically adequate to the phenomena in some domain, we should infer the truth of the hypothesis which gives us the best explanation of those phenomena. It is often claimed that IBE gives us justified beliefs and knowledge. It certainly seems that in everyday life when faced with a range of hypotheses that all account for some phenomenon, we usually adopt the one which best explains it. Here is an example from van Fraassen: Suppose you hear scratching in the wall of your house, the patter of little feet at midnight, and cheese keeps disappearing. You would doubtless infer that a mouse has taken up residence [1980, 19]. This inference has the structure, if p then q, q therefore p, in other words, you know that if there is a mouse then there will be droppings, noises and other observable evidence, and you observe the evidence and so infer the existence of a mouse. However, consider the following: if something is a square, then it has four sides, a rectangle has four sides, therefore it is a square; this is deductively invalid because it is possible for the conclusion to be false when the premises are both true, for example, if two of the sides of the rectangle are of different lengths (this is the fallacy called 'affirming the consequent'). Similarly, there is no contradiction in supposing that there is no mouse in the house despite the evidence, so that instance of inference to the best explanation is also deductively invalid.

IBE is usually ampliative and invalid so the problem is to explain what distinguishes justifiable and knowledge-producing instances of IBE, from other invalid inferences that are clearly just bad reasoning. Here are some features that instances of IBE might be required to have:

1. Otherwise surprising phenomena are to be expected if the hypothesis is true.
2. Predictions of empirical consequences must be inferred from the hypothesis and tested and confirmed.
3. Simple and natural hypotheses are to be favoured.

4. Hypotheses which cohere with metaphysical views are to be favoured.

5. Unifying power and wideness of scope of hypotheses are to be favoured.

6. Hypotheses that cohere with other scientific theories are to be favoured.

Inference to the best explanation is used to defend scientific realism in two ways: at the local level, the idea is that if we are to follow the same patterns of inference in philosophy, science and ordinary life then we will be scientific realists since, for example, our best explanation of many phenomena involves the theoretical unobservable entities postulated by science.

IBE is invoked by scientific realists to break the underdetermination of theory by evidence. Recall the second premise of the underdetermination argument: If two theories are empirically equivalent then they are evidentially equivalent. If two theories are empirically equivalent but one of them offers a better explanation of the phenomena, then advocates of IBE will argue that we can infer the truth of the more explanatory theory. Hence advocates of IBE think that the explanatory power of a theory is evidence for its truth and hence that the second premise of the underdetermination argument is false. But van Fraassen argues that explanatory power is a merely pragmatic virtue of theories and does not give us evidence for their truth, and that IBE at the everyday level can always be recast as inference to the empirical adequacy of the best explanation. He also argues that the realist demand for explanation of every regularity leads to infinite regress.

The defence of scientific realism by appeal to IBE at the global level is based on the claim that scientific realism is the best explanation of the overall success of scientific theorising — this is known as 'the no-miracles argument' because the idea is that the success of science would be miraculous on anything but a scientific realist view. In particular, realists (following Richard Boyd) argue that we need to explain the overall instrumental success of scientific methods across the history of science. All parties in the scientific realism debate agree that:

(i) Patterns in data are projectable from the observed to the unobserved using scientific knowledge, which is to say that induction based on scientific theories is reliable.

(ii) The degree of confirmation of a scientific theory is heavily theory-dependent, in the sense that background theories inform judgements about the extent to which different theories are supported by the available evidence.

(iii) Scientific methods are instrumentally reliable, in other words, they are reliable ways of achieving practical goals like prediction and the construction of technological devices.

Scientific realists argue that these features of science would be utterly mysterious if the theories involved were not true or approximately true.

Another feature of scientific practice that realists have long argued cannot be explained by antirealists is the persistent and often successful search for unified theories of diverse phenomena. The well known 'conjunction objection' against antirealism is as follows: Consider two scientific theories, T and T', from different domains of science, say chemistry and physics. That T and T'' are both empirically adequate does not imply that their conjunction T & T' is empirically adequate, however, if T and T' are both true this does imply that T & T' is true. So, the argument goes, only realists are motivated to believe the new empirical consequences obtained by conjoining accepted theories. However, it is claimed that in the course of the history of science the practice of theory conjunction is widespread and a reliable part of scientific methodology. Therefore, if scientists are not irrational, since only realism can explain this feature of scientific practice, then realism must be true.

A fundamental criticism of the use of IBE at the global level was made by Larry Laudan and Arthur Fine, both of whom pointed out that since it is IBE involving unobservables that is in question in the realism debate, it is circular to appeal to the explanatory power of scientific realism at the meta-level to account for the overall success of science because realism is itself a hypothesis involving unobservables. Hence, it is argued that the global defence is question begging. There is a similarity here with the inductive vindication of induction. Richard Braithwaite, and Carnap, defended the view that the inductive defence of induction — induction has worked up to now so it will work in the future — was circular but not viciously so, because it is rule circular not premise circular. In the case of IBE such a view has been defended by David Papineau and Stathis Psillos. The idea is that premise circularity of an argument is vicious because the conclusion is taken as one of the premises; on the other hand rule circularity is when the conclusion of an argument states that a particular rule is reliable, but that conclusion only follows from the premises when that very rule is used. Now notice that the global defence of realism is rule but not premise circular. The conclusion that the use of IBE in science is reliable is not a premise of this defence of realism, but the use of IBE is required to reach this conclusion from the premises that IBE is part of scientific methodology and that scientific methodology is instrumentally reliable. It is conceded that, although it is not viciously circular, this style of argument will not persuade someone who totally rejects IBE. However, what the argument is meant to show is that someone who does make abductive inferences can show the reliability of their own methods. So, it seems that IBE is on a par with inductive reasoning; it cannot be defended by a noncircular argument, but recall that even deduction cannot be defended by a non-circular argument either. Hence, the realist may claim that although they cannot force the non-realist to accept IBE, they can show that its use is consistent and then argue that it forms part of a comprehensive and adequate philosophy of science.

However, an antirealist could agree with the descriptive claim that often our inductive inferences are guided by explanatory considerations, and accept that to be so guided is not prohibited by the canons of rationality. However, it may be argued that nobody is ever rationally compelled to believe something because it is the best explanation of the phenomena. Furthermore, arguably IBE is only pragmatically motivated in general: as it turns out, being guided by explanatory considerations has led us to arrive at empirically adequate theories, and that gives us some reason to search for explanations in the future, but we should not admit explanatory considerations as reasons for belief if we are good empiricists, and we should certainly never regard IBE as rationally compelling. It may be objected that it is capricious to use inference to the best explanation widely, but to always abstain from inferring the truth of the conclusion in the case of unobservable entities. However, there is a salient difference between inferring the existence of an unobserved observable and inferring the existence of an unobservable, namely that the former case is usually the inferring of the existence of an unobserved token of an observed type that is at issue. (In the next section it is shown that the history of science gives us further reasons to be wary of committing ourselves to the existence of the unobservables postulated to explain observable phenomena.)

Van Fraassen offers several arguments against the idea that IBE is a compelling rule of inference:

(i) The Argument from Indifference

The argument from indifference is roughly that, since there are many ontologically incompatible yet empirically equivalent theories, we have no reason to choose among them and indentify one of them as true. This argument appeals to the existence of empirical equivalents to any theory that we have. In the discussion of the underdetermination problem above it was concluded that the antirealist may also be threatened by the existence of empirical equivalents since any finite set of theories that we consider is just as highly unlikely to contain an empirically adequate theory. However, this does not help defend IBE.

(ii) The Argument from the Best of a Bad Lot

This argument is that some 'principle of privilege' is required if we are to think that the collection of hypotheses that we have under consideration will include the true theory. The best explanatory hypothesis we have may just be the best of a bad lot, all of which are false. In other words this argument challenges the proponent of IBE to show how we can know that none of the other possible explanations we have not considered is as good as the best that we have. Unless we know that we have included the best explanation in our set of rival hypotheses, even if it were the case that the best explanation is true, this would not make IBE an acceptable rule of inference.

Realists tend to bite this bullet and argue that scientists do have privilege which issues from background knowledge. Theory choice is informed by background theories which narrow the range of hypotheses under consideration, and then explanatory considerations help select the best hypothesis. Furthermore, they argue that both the realist and the constructive empiricist need privilege, because the constructive empiricist needs to assume that the empirically adequate theory is among the ones considered in order to have warranted belief in the empirical adequacy of the chosen theory. Hence the dispute can only be about the extent of that privilege.

(iii) The argument from Bayesianism

The idea here is that any rule for the updating of belief that goes beyond the rules of Bayesian conditionalisation (see VII of 3) will lead to probabilistic incoherence.

(VII) Arguments from Theory Change

Unlike the underdetermination problem which may seem to be generated a priori, the arguments against scientific realism from theory change are empirically based and their premises are derived from data obtained by examining the practice and history of science. Furthermore, ontological discontinuity across radical changes in theories seems to give us grounds not merely for doubt, but for the positive belief that many central theoretical terms of our best contemporary science will be regarded as non-referring by future science. Hence, the strongest argument from theory change has as its conclusion that scientific realism is not true because it is not even empirically adequate. The argument in question is the 'pessimistic meta-induction', and was anticipated by the ancient Greek sceptics, but in its contemporary form it is due to Larry Laudan. It has the following structure:

(i) There have been many empirically successful theories in the history of science which have subsequently been rejected and whose central theoretical terms do not refer according to our best current theories.

(ii) Our best current theories are no different in kind from those discarded theories and so we have no reason to think they will not ultimately be replaced as well.

So, by induction we have positive reason to expect that our best current theories will be replaced by new theories according to which some of the central theoretical terms of our best current theories do not refer, and hence, we should not believe in the approximate truth or the successful reference of the theoretical terms of our best current theories.

The most common realist response to this argument is to restrict realism to theories with some further properties (usually, maturity, and novel predictive success) so as to cut down the inductive base employed in (i). However, assuming that such an account can be given there are still a couple of cases of mature theories which enjoyed novel predictive success by anyone's standards, namely the ether theory of light and the caloric theory of heat. If their central theoretical terms do not refer, the realist's claim that approximate truth explains empirical success will no longer be enough to establish realism, because we will need some other explanation for success of the caloric and ether theories. If this will do for these theories then it ought to do for others where we happened to have retained the central theoretical terms, and then we do not need the realist's preferred explanation that such theories are true and successfully refer to unobservable entities. To be clear:

(a) Successful reference of its central theoretical terms is a necessary condition for the approximate truth of a theory.

(b) There are examples of theories that were mature and had novel predictive success but which are not approximately true.

(c) Approximate truth and successful reference of central theoretical terms is not a necessary condition for the novel-predictive success of scientific theories

So, the no-miracles argument is undermined since, if approximate truth and successful reference are not available to be part of the explanation of some theories' novel predictive success, there is no reason to think that the novel predictive success of other theories has to be explained by realism.

Hence, we do not need to form an inductive argument based on Laudan's list to undermine the no-miracles argument for realism. Laudan's paper was also intended to show that the successful reference of its theoretical terms is not a necessary condition for the novel predictive success of a theory, and that there are counter-examples to the no-miracles argument.

There are two basic (not necessarily exclusive) responses to this:

(I) Develop an account of reference according to which the abandoned theoretical terms are regarded as referring after all.

(II) Restrict realism to those parts of theories which play an essential role in the derivation of subsequently observed (novel) predictions, and then argue that the terms of past theories which are now regarded as non-referring were non-essential so there is no reason to deny that the essential terms in current theories will be retained.

Realists have used causal theories of reference to account for continuity of reference for terms like 'atom' or 'electron', when the theories about atoms and electrons undergo significant changes. The difference with the terms 'ether' and 'caloric' is that they are no longer used in modern science. In the nineteenth century the ether was usually envisaged as some sort of material solid that permeated all of space. It was thought that light waves had to be waves in some sort of medium and the ether was posited to fulfil this role. Yet if there really is such a medium then we ought to be able to detect the effect of the Earth's motion through it, because light waves emitted perpendicular to the motion of a light source through the ether ought to travel a longer path than light waves emitted in the same direction as the motion of the source through the ether. Of course, various experiments, the most famous being that of Michaelson and Morley, failed to find such an effect. By then Maxwell had developed his theory of the electromagnetic field, which came to be regarded as not a material substance at all. As a result the term 'ether' was eventually abandoned completely.

However, the causal theory of reference may be used to defend the claim that the term 'ether' referred after all, but to the electromagnetic field rather than to a material medium. If the reference of theoretical terms is to whatever causes the phenomena responsible for the terms' introduction, then since optical phenomena are now believed to be caused by the oscillations in the electromagnetic field, than the latter is what is referred to by the term 'ether'. Similarly, since heat is now believed to be caused by molecular motions, then the term 'caloric' can be thought to have referred all along to these rather than to a material substance. The danger with this is that it threatens to make the reference of theoretical terms a trivial matter, since as long as some phenomena prompt the introduction of a term it will automatically successfully refer to whatever is the relevant cause (or causes). Furthermore, this theory radically disconnects what a theorist is talking about from what they think they are talking about. For example, Aristotle or Newton could be said to be referring to geodesic motion in a curved spacetime when, respectively, they talked about the natural motion of material objects, and the fall of a body under the effect of the gravitational force.

The essence of the second strategy is to argue that the parts of theories that have been abandoned were not really involved in the production of novel predictive success. Philip Kitcher says that: "No sensible realist should ever want to assert that the idle parts of an individual practice, past or present, are justified by the success of the whole" [1993, 142]. Kitcher suggests a model of reference according to which some tokens of theoretical terms refer and others do not, but his theory allows that the theoretical descriptions of the theoretical kinds in question may have been almost entirely mistaken, and seems to defend successful reference only for those uses of terms that avoid ontological detail at the expense of reference to something playing a causal role in producing some observable phenomena.

Similarly, Stathis Psillos argues that history does not undermine a cautious scientific realism that differentiates between the evidential support that accrues to different parts of theories, and only advocates belief in those parts that are essentially involved in the production of novel predictive success. This cautious, rather than an all or nothing, realism would not have recommended belief in the parts of the theories to which Laudan draws attention, because if we separate the components of a theory that generated its success from those that did not we find that the theoretical commitments that were subsequently abandoned are the idle ones. On the other hand, argues Psillos: "the theoretical laws and mechanisms that generated the successes of past theories have been retained in our current scientific image" [1999, 108]. Such an argument needs to be accompanied by specific analyses of particular theories which both identify the essential contributors to the success of the theory in question, and show that these were retained in subsequent developments.

Psillos takes up Kitcher's suggestion of (II) and combines it with (I). Laudan claims that if current successful theories are approximately true, then the caloric and ether theories cannot be because their central theoretical terms don't refer (by premise (ii) above). Strategy (I) accepts premise (ii) but Psillos allows that sometimes an overall approximately true theory may fail to refer. He then undercuts Laudan's argument by arguing that abandoned theoretical terms that do not refer, like 'caloric', were involved in parts of theories not supported by the evidence at the time, because the empirical success of caloric theories was independent of any hypotheses about the nature of caloric. Abandoned terms that were used in parts of theories supported by the evidence at the time do refer after all; 'ether' refers to the electromagnetic field. The problem with strategy (II) is that its applications tend to be ad hoc and dependent on hindsight. Furthermore, by disconnecting empirical success from the detailed ontological commitments in terms of which theories were described, it seems to undermine rather than support realism.

As we have seen, in the debate about scientific realism, the no-miracles argument is in tension with the arguments from theory-change. In an attempt to break this impasse, and have "the best of both worlds", John Worrall [1989] introduced structural realism (although he attributes its original formulation to Poincaré). Using the case of the transition in nineteenth century optics from Fresnel's elastic solid ether theory to Maxwell's theory of the electromagnetic field, Worrall argues that:

> There was an important element of continuity in the shift from Fresnel to Maxwell — and this was much more than a simple question of carrying over the successful empirical content into the new theory. At the same time it was rather less than a carrying over of the full theoretical content or full theoretical mechanisms (even in "approximate" form) ... There was continuity or accumulation in the shift, but the continuity is one of form or structure, not of content. [1989, 117]

According to Worrall, we should not accept full blown scientific realism, which asserts that the nature of things is correctly described by the metaphysical and physical content of our best theories. Rather we should adopt the structural realist emphasis on the mathematical or structural content of our theories. Since there is (says Worrall) retention of structure across theory change, structural realism both (a) avoids the force of the pessimistic meta-induction (by not committing us to belief in the theory's description of the furniture of the world) and (b) does not make the success of science (especially the novel predictions of mature physical theories) seem miraculous (by committing us to the claim that the theory's structure, over and above its empirical content, describes the world). A different form of structural realism is also defended by Steven French and James Ladyman in the context of interpreting contemporary physics.

(VIII) Contemporary Empiricism

The constructive empiricism of van Fraassen has provoked renewed debate about scientific realism. Van Fraassen accepts the semantic and metaphysical components of scientific realism, but, he denies the epistemic component. So he thinks that scientific theories about unobservables should be taken literally, and are true or false in the correspondence sense, depending on whether the entities they describe are part of the mind-independent world. However, he argues that acceptance of the best theories in modern science does not require belief in the entities postulated by them, and that the nature and success of modern science relative to its aims can be understood without invoking the existence of such entities.

Van Fraassen defines scientific realism as follows:

> Science aims to give us, in its theories, a literally true story of what the world is like; and acceptance of a scientific theory involves the belief that it is true. [1980, 8]

Constructive empiricism is the view that:

> Science aims to give us theories which are empirically adequate; and acceptance of a theory involves as belief only that it is empirically adequate. [Ibid., 12]

To say that a theory is empirically adequate is to say: "What it says about the observable things and events in this world, is true (ibid.)". In other words:

> the belief involved in accepting a scientific theory is only that it 'saves the phenomena', that is that it correctly describes what is observable. [Ibid., 4]

Note that this means that it saves *all* the *actual* phenomena, past present and future, not just those that have been observed so far, so even to accept a theory as empirically adequate is to believe something more than is logically implied by the data [ibid., 12, 72]. Moreover, for van Fraassen, a phenomenon is simply an *observable* event and not necessarily an observed one. So a tree falling over in a forest is a phenomenon whether or not someone actually witnesses it.

The scientific realist and the constructive empiricist disagree about the purpose of the scientific enterprise: the former thinks that it aims at truth with respect to the unobservable processes and entities that *explain* the observable phenomena; the latter thinks that the aim is merely to tell the truth about what is observable, and rejects the demand for explanation of all regularities in what we observe. Van Fraassen says that explanatory power is not a "rock bottom virtue" of scientific theories whereas consistency with the phenomena is [ibid., 94]. Hence, for the constructive empiricist, empirical adequacy is the internal criterion of success for scientific activity.

Note that

(a) Both doctrines are defined in terms of the aims of science, so constructive empiricism is fundamentally a view about the aims of science and the nature of 'acceptance' of scientific theories, rather than a view about whether electrons and the like exist. Strictly speaking it is possible to be a constructive empiricist and a scientific agnostic, or a scientific realist and scientific agnostic. That said, it is part of van Fraassen's aim to show that abstaining from belief in unobservables is perfectly rational and scientific.

(b) Scientific realism has two components: (i) theories which putatively refer to unobservable entities are to be taken literally as assertoric and truth-apt claims about the world (in particular, as including existence claims about unobservable entities); and (ii) acceptance of these theories (or at least the best of them) commits one to belief in their truth or approximate truth in the correspondence sense (in particular, to belief that tokens of the types postulated by the theories in fact exist). Van Fraassen is happy to accept (i). It is (ii) that he does not endorse. Instead, he argues that acceptance of the best theories in modern science does not require belief in the entities postulated by them, and that the nature and success of modern science relative to its aims can be understood without invoking the existence of such entities.

(c) Empirical adequacy for scientific theories is characterised by van Fraassen in terms of the so-called 'semantic' or 'model-theoretic' conception of scientific theories, the view that theories are fundamentally extra-linguistic entities (models or structures), as opposed to the syntactic account of theories, which treats them as the deductive closure of a set of formulas in first order logic. The semantic view treats the relationship between theories and the world in terms of isomorphism. On this view, loosely speaking, a theory is empirically adequate if it "has at least one model which all the actual phenomena fit inside" [1980, 12].

Initial criticism of van Fraassen's case for constructive empiricism concentrated on three issues:

(i) The line between the observable and the unobservable is vague and the two domains are continuous with one another; moreover the line between the observable and the unobservable changes with time and is an artefact of accidents of human physiology and technology. This is supposed to imply that constructive empiricism grants ontological significance to an arbitrary distinction.

(ii) Van Fraassen eschews the positivist project which attempted to give an a priori demarcation of predicates that refer to observables from those that refer to unobservables, and accepts instead that: (a) all language is theory-laden to some extent; and (b) even the observable world is described using terms that putatively refer to unobservables. Critics argue that this makes van Fraassen's position incoherent.

(iii) The underdetermination of theory by evidence is the only positive argument that van Fraassen has for adopting constructive empiricism instead of scientific realism; but all the data we presently have underdetermine which theory is empirically adequate just as they underdetermine which theory is true (this is the problem of induction), and so constructive empiricism is just as vulnerable to scepticism as scientific realism. This is taken to imply that van Fraassen's advocacy of constructive empiricism is the expression of an arbitrarily selective scepticism.

(i) is rebutted firstly by pointing out that vague predicates abound in natural language but clear extreme cases suffice to render their use acceptable, and secondly by arguing that epistemology ought to be indexical and anthropocentric, and that the distinction between the observable and the unobservable is not to be taken as having direct ontological significance, but rather epistemological significance. Says van Fraassen: "even if observability has nothing to do with existence (is, indeed, too anthropocentric for that), it may still have much to do with the proper epistemic attitude to science" [1980, 19].

For van Fraassen, 'observable' is to be understood as 'observable-to-us": "X is observable if there are circumstances which are such that, if X is present to us under those circumstances, then we observe it" [1980, 16]. What we can and cannot observe is a consequence of the fact that

> the human organism is, from the point of view of physics, a certain kind of measuring apparatus. As such it has certain inherent limitations — which will be described in detail in the final physics and biology. It is these limitations to which the ''able' in 'observable' refers — our limitations, *qua* human beings. [1980, 17]

So we know that, for example, the moons of Jupiter are observable because our current best theories say that were astronauts to get close enough, then they *would* observe them; by contrast the best theories of particle physics certainly do not tell us that we are directly observing the particles in a cloud chamber. Analogous with the latter case is the observation of the vapour trail of a jet in the sky, which does not count as observing the jet itself, but rather as detecting it. Now if subatomic particles exist as our theories say that they do, then we detect them by means of observing their tracks in cloud chambers, but, since we can never experience them directly (as we can jets), there is always the possibility of an empirically equivalent but incompatible rival theory which denies that such particles exist. This fact may give the observable/unobservable distinction epistemic significance. Note, that van Fraassen adopts a direct realism about perception for macroscopic objects: "we can and do see the truth about many things: ourselves, others, trees and animals, clouds and rivers — in the immediacy of experience" [1989, 178].

(ii) is rebutted by showing that there are at least some entities which if they exist are unobservable, for example, quarks, spin states of sub-atomic particles, and light.

(iii) is the most serious problem for van Fraassen. Note first that, contrary to what is often claimed, van Fraassen does not accept that inference to the best explanation is rationally compelling in the case of the observable world while denying it this status for the case of the unobservable world. Furthermore, van Fraassen does not appeal to any global arguments for antirealism such as the underdetermination argument or the pessimistic meta-induction. He rejects realism not because he thinks it irrational but because he rejects the "inflationary metaphysics" which he thinks must accompany it, i.e., an account of laws, causes, kinds and so on, and because he thinks constructive empiricism offers an alternative view that offers a better account of scientific practice without such extravagance [1980, 73]. Empiricists should repudiate beliefs that go beyond what we can (possibly) confront with experience; this restraint allows them to say "good bye to metaphysics" [1989; 1991, 480].

What then is empiricism and why should we believe it? Van Fraassen suggests that to be an empiricist is to believe that "experience is the sole source of information about the world" [1985, 253]. The problem with this doctrine is that it does not itself seem to be justifiable by experience. However, he has argued in recent work that empiricism cannot be reduced to the acceptance of such a slogan, and that empiricism is in fact a stance in Husserl's sense of an orientation or attitude towards the world.

Constructive empiricism is supposed to offer a positive alternative to scientific realism that dispenses with the need for metaphysics. It is a positive view of science which is intended to free us from the need to articulate accounts of laws, causes, and essential properties that take seriously the apparent modal commitments of such notions. This promised liberation from metaphysics is fundamental to van Fraassen's advocacy of a constructive empiricist view of science. Indeed, from his point of view, scepticism about objective modality is partly definitive of an empiricist outlook: "To be an empiricist is to withhold belief in anything that goes beyond the actual, observable phenomena, and to recognise no objective modality in nature" [1980, 202]. However, arguably, to be a constructive empiricist one must, in spite of what van Fraassen says here, recognise objective modality in nature. This is largely because constructive empiricism recommends, on epistemological grounds, belief in the empirical adequacy rather than the truth of theories, and hence requires that there be an objective modal distinction between the observable and the unobservable.

(IX) Pragmatism

Various philosophers have defended forms of pragmatism in philosophy of science, not least because it seems to help scientific realists avoid problems like the underdetermination problem. Brian Ellis says: "scientific realism can be combined with a pragmatic theory of truth: and given such a theory of truth, all of the criteria which we use for the evaluation of theories, including the so-called pragmatic ones, can be seen as being relevant to their truth or falsity" [1985, 41]. Other forms of pragmatism include that of Ian Hacking and Nancy Cartwright who both defend entity realism, which is the idea that the unobservable entities that are posited by science can be known about even if the fundamental theoretical claims of scientific theories are not true, because the entities can be manipulated and play a role in the practical life of experimentalists, engineers and technologists.

Arthur Fine defends what he calls the Natural Ontological Attitude (NOA). NOA is, he claims, a minimalist view that avoids the philosophical conceits of both scientific realism and antirealism, and simply incorporates the 'homely line' that we should regard the pronouncements of science as on a par with everyday talk about objects observed with the senses. Realists have argued in response that NOA is all the scientific realist needs, since it says that the unobservable objects postulated by scientific theories have the same status as tables and chairs, and, in particular, this suggests that we can refer to and know the truth about them. However, Fine argues against all the standard philosophical arguments for scientific realism with some vigour, and his position may therefore seem to be antirealist. For Fine, asking whether electrons really exist is like asking whether tables do, and in both cases he refuses to engage with the question. He seems to have adopted a philosophical quietism that is consistent even with solipsism or Berkelian scepticism. Fine is quite deliberate about this since he claims that precisely what distinguishes NOA from scientific realism is that the latter involves a metaphysics of the external world, and a theory of truth and so on, while the former does not bother with them. Thus, NOA seems to be a recommendation for the abandonment of certain philosophical questions. According to him what marks out the realist or antirealist is what they add on to the everyday talk of scientists. Hence, theories of truth, whether, for example, correspondence (realist) or coherence (antirealist), are equally unnecessary and unhelpful. Rather, it seems Fine proposes that we should simply stop talking about truth per se and accept the homely truths that scientists use just as we accept the truths of everyday life.

(X) Naturalism and Normativity

Naturalism is the view that philosophy is continuous with science. According to naturalists, traditional philosophical questions concerning knowledge ought to be investigated by cognitive science and evolutionary psychology, rather than by a priori reflection. Naturalists also think that metaphysical questions can only be answered by science rather than by thought experiments and other traditional philosophical methods.

Normativity concerns not what is the case but what ought to be the case. The main source of normative claims is ethics, however, logic, rationality and reason also seem to be concerned with what ought to be the case. For example, it seems that we ought to be believe what is true, and that we ought not to be believe what is false. Those opposed to naturalism argue that it will never be possible to explain normativity in scientific terms since science can only describe how the world is and claims about what ought to be the case can never be tested. Hume famously argued that it is not possible to derive an 'ought' from an 'is', and if this is right there would seem to be something to the idea that normativity lies outside the scope of naturalism. However, this leads some philosophers to adopt scepticism about normativity and regard what we ought to do as a mere matter of convention.

3 METHODOLOGICAL POSITIONS

(I) Inductivism

The most general characterisation of inductivism is that it is any position according to which a universal generalisation is positively supported by observation of its instances. Philosophers refer to the idea of evidence positively supporting a law or theory as confirmation. The idea of confirmation is fundamental to most but not all theories of the scientific method.

Naïve inductivism states that the basic means by which scientific knowledge is advanced is generalisation from experience. It is associated with Francis Bacon who criticised the natural philosophy of his day for being insufficiently empirical and experimental. Bacon advocated the influential idea that science in any domain must begin with numerous and wide-ranging observations that are undertaken without prejudice or preconception. Many scientific laws are of the form of universal generalisations (statements that generalise about the properties of all things of a certain kind). For example, 'all metals expand when heated' is a universal generalisation about metals. Induction in the broadest sense is just any form of reasoning which is not deductive, but in the narrower sense which Bacon uses it is the form of reasoning where we generalise from a collection of particular instances to a general conclusion. The simplest form of induction is enumerative induction, which is where we observe that some large number of instances of some phenomenon has some characteristic and then infer that the phenomenon always has that property. Bacon also discussed more involved methods involving the drawing up of tables to compare and contrast different instances of some phenomenon, so that it can be inferred what all such instances have in common.

According to naïve inductivism it is legitimate to infer a universal generalisation from a collection of observation statements when a large number of observations of Xs under a wide variety of conditions have been made, and when all Xs have been found to possess property Y, and when no instance has been found to contradict the universal generalisation 'all Xs possess property Y'. This is known as a Principle of Induction. Once a generalisation has been inductively inferred in accordance with this principle then it assumes the status of a law or theory and we can use deduction to deduce consequences of the law that will be predictions or explanations.

The obvious problem with this is how to make precise the idea of a large number of observations. One common response to this problem is to claim that, given that no amount of evidence of observed cases will ever logically entail a claim about unobserved cases (the problem of induction), then it is never the case that enough observations have been made to establish a hypothesis with certainty, and that therefore we ought instead to think it terms of probabilities, so that that the larger the number of observations that have been made then the higher the probability that the universal generalisation is true. However, it is easy to see that the move to probabilities does not solve the problem since a universal generalisation covers potentially infinitely many cases, so no matter how many instances are observed, if there are a finite number, it would seem that the probability of the universal generalisation will always be small.

In any case this is simplistic, even if it works for some parts of science. For example, it is arguably impossible to engage in observation without any preconceptions or presuppositions, since theory guides the decision as to what to pay attention to when observing, and also theories often suggest experiments that might be performed to test them. Inductivists may appeal to more sophisticated kinds of induction such as Mill's methods for eliminative induction, which attempts to find the right universal generalisation by eliminating the alternatives rather than be enumerative generalisation. Nonetheless, all forms of inductivism face the problem of induction. Naïve inductivism and more sophisticated forms of inductivism face other problems that arise when it is observed that often in the history of science, great advances have been made by people who have not followed inductive methods. In particular, sometimes scientifically valuable hypotheses have been the result of speculation rather than generalisation from experience.

(II) The Context of Discovery and the Context of Justification

There is a fundamental difference between accounts of the scientific method, namely that some are accounts of how to generate scientific theories, and also how to test scientific theories and how to respond to the results of testing them, while others do not attempt to describe how scientific theories should be generated. Clearly, Bacon's inductive logic is an example of the former, since it proscribes how to begin the investigation of some range of phenomena, and the production of generalisations and laws is supposed to be an automatic outcome of the mechanical operation of the method. Examples of the latter include falsificationism which is discussed below, but also some versions of inductivism.

It may be desirable that laws and theories be derived from experimental data (as Newton claimed that he did not speculate but rather deduced the laws of mechanics from the results of observations), but in most interesting cases this is just not possible. The generation of scientific theories is not in general a mechanical procedure, but a creative activity. Scientists have drawn upon many sources of inspiration, such as metaphysical beliefs, dreams, analogies and so on, when trying to formulate a theory. The kind of speculation and imagination which scientists need to employ cannot be formalised or reduced to a set of rules, but once a hypothesis is generated it must be subject to testing by experience, and this must be the final arbiter of any scientific dispute. If this is right then, when we are thinking about scientific methodology, perhaps we ought to make a distinction between the way theories are conceived and the subsequent process of testing them. The scientific method may be silent about where hypotheses should come from, but, faced with two rival hypotheses that equally account for the data, scientists ought to construct an experimental situation (crucial experiment) about which the hypotheses will predict different outcomes.

In general, the evidence in favour of a hypothesis is independent of who believes it, and whether an idea is a good one does not depend on who first thinks of it. So it seems plausible to argue that evaluation of the evidence for a hypothesis ought to take no account of how, why and by whom the hypothesis was conceived. This distinction between the causal origins of scientific theories and their degree of confirmation and scientific status is often thought to be important for the defence of the objectivity of scientific knowledge. Many philosophers of science, who otherwise disagree with each other about fundamental matters, believe that the task of philosophy of science is to logically analyse the testing of scientific theories by observation and experiment. How theories are developed is a matter for psychology not philosophy. Scientists do not need to make presuppositionless observations, nor does it matter if they use background theories to develop new theories.

On this view, there are two contexts in which the history of science may be investigated, namely the context of discovery and the context of justification. The distinction between the context of discovery and the context of justification separates the question of how scientific theories are developed from the question of how to test them against their rivals. The degree of confirmation of a theory is a relationship between it and the evidence and is independent how it was produced.

(III) The Paradoxes of Confirmation

Any theory of confirmation must avoid the following paradoxes (see 1.I Natural Kinds).

(i) The Ravens Paradox

'Nicod's criterion' states that laws are confirmed by observation of their instances. If we assume that logically equivalent propositions are equally confirmed by the same evidence, then the logical equivalence of 'all Ravens are black' and 'all non-black things are non-ravens' implies that observation of a green leaf confirms the law that all ravens are black.

(ii) Goodman's New Riddle of Induction

Consider a law of nature of the form 'all Fs are Gs'. Construct the predicate G^* as follows: a is G^* iff a is G before time t and H after time t, where 'a is H' entails 'it is not the case that a is G'. It seems that all the evidence gathered before time t must equally support the law 'all Fs are G^*s'. Hence, one question Goodman's riddle poses is 'what is the justification for taking generalisations with ordinary predicates to be confirmed by the instances we have so far observed, but not generalisations with predicates like G^*? It seems that appeal to the uniformity of nature alone does not solve the problem of induction because the world may be uniform in different ways; if it is uniform in that all Fs are G^*s then our ordinary inductive inferences will be unreliable.

(iii) The Tacking Paradox

The special consequence condition states that evidence which supports a theory also supports the logical consequences of the theory. This is plausible because it seems to be necessary to explain why evidence that confirms a theory also gives us a reason to believe that the predictions the theory makes are true. The converse consequence condition states that evidence which supports a theory, T, also supports any other theory which entails T. This seems plausible because there are many cases in the history of science where a high level theory entailed a low level law that was already supported by the evidence, and where that evidence was then taken to also support the high-level law. However, it follows from these two conditions that any piece of evidence for an arbitrary theory supports any hypothesis whatsoever. Consider e which supports theory T. Since T is entailed by $T\&G$ for any G, it follows from the converse consequence condition that e supports $T\&G$. But then since $T\&G$ entails G, it follows from the special consequence condition that e supports G.

Each of these paradoxes seems to rely on the assumption that the relation between a theory and the evidence which supports it is a logical one. Some think the paradoxes show that a purely logical theory of confirmation is impossible. Historical theories of confirmation make the history and origin of a theory relevant to its evidential status. For example, Goodman's own response to his problem was to argue that entrenched predicates, ones that have been used in successful inductions in the past, are more confirmed by new evidence than un-entrenched ones like G^*. Historical theories of confirmation collapse the distinction between the context of discovery and the context of justification, according to which the causal history of a theory is quite irrelevant to the extent to which it is supported by the evidence. Many philosophers worry that a historical theory of confirmation is inconsistent with the idea that the evidential basis of scientific knowledge is an objective matter, and hence to invite relativism and subjectivism. On the other hand, many other philosophers have given up on the ideal of an ahistorical theory of confirmation.

(IV) Explanation versus Prediction

Carl Hempel advocates the thesis of structural identity, according to which explanations and predictions have exactly the same structure: they are arguments where the premises state laws of nature and initial conditions. The only difference between them is that, in the case of an explanation we already know that the conclusion of the argument is true, whereas in the case of a prediction the conclusion is unknown. For example, Newtonian physics predicted the return of Halley's comet in December 1758, and the same argument explains its return. However, there are many cases where the observation of one phenomenon allows us to predict the observation of another phenomenon but where the former does not explain the latter. For example, the fall of the needle on a barometer allows us to predict that there will be a storm but doesn't explain it. Similarly, the length of a shadow allows us to predict the height of the building that cast it, and the period of oscillation of a pendulum we can predict its length, but in both these cases the latter explains the former and not the other way round. There also seem to be theories that provide adequate explanations but that cannot make precise predictions. For example, evolutionary theory explains why organisms have the morphology that they do, but it cannot make specific predictions because evolutionary change is subject to random variations in environmental conditions and the genotype of organisms. Furthermore, there are cases of probabilistic explanations where the probability conferred by the explanans on the explanandum is low, so we cannot predict that the explanandum is even likely to happen although we can explain why it did if it does.

According to hypothetico-deductivism, there is also a symmetry between predictions and explanations in respect of confirmation; because an explanation is simply a prediction where the phenomenon predicted has already been observed, the degree of confirmation conferred on a theory is the same for predictions and explanations. Hypothetico-deductivism is a purely logical theory of confirmation, and the origin of a theory, in particular, when it was proposed relative to when the evidence for it was gathered, is irrelevant to its epistemic status. On the other hand, predictivists think that only successful predictions of previously unknown phenomena count as evidence, and explanationists think that only explanations of previously known about phenomena count as evidence. Intermediate positions accord some confirmational power to both predictions and explanations but weight one more highly than the other. Many scientific realists argue that novel predictions of new and unsuspected types of phenomena are of special confirmational status.

The significance of novel predictions was emphasised by Karl Popper. He contrasted the risky predictions of physics with the vague predictions of psychoanalysis, but he also wanted to justify the failure of scientists to abandon Newtonian theory when it was known to be incompatible with certain observations. Often various modifications to background assumptions are made to try and accommodate observed facts that would otherwise refute established theories. Popper, and following him Imre Lakatos and others, argued that this course of action is acceptable only when the new theory produces testable consequences other than the results that motivated it. So for example, the postulation of a new planet to accommodate the observed orbit of a familiar one is legitimate because it ought to be possible to observe the former (or at least its effects on other bodies).

Popper was particularly impressed by the experimental confirmation of Einstein's general theory of relativity in 1917. The latter predicted that light passing close to the Sun ought to have its path bent by the Sun's gravitational field. Another well known example is from optics. In 1818 Fresnel developed a mathematical theory according to which light consists of transverse waves in an optical ether. This theory predicted that in certain circumstances light that was shone on a completely opaque disk would cast a shadow with a bright white spot in its centre. However, Fresnel knew nothing of this phenomenon when he developed his theory, and indeed did not even derive the result himself. This is more striking than the prediction of the existence of an extra planet, because it is a prediction of a completely new and unexpected type of phenomenon.

The most straightforward idea of novelty is that of temporal novelty. A prediction is temporally novel when it is of something that has not yet been observed. The problem with attributing special confirmational status to this kind of novel predictive success is that it seems to introduce an element of arbitrariness into the theory of confirmation. When exactly in time someone first observes some phenomenon entailed by a theory may have nothing to do with how and why the theory was developed. It seems implausible that it should be relevant to the degree of confirmation of a theory provided by some evidence whether or not someone independently observed the evidence before the theory was produced but didn't tell anyone about it. As it turns out, the white spot phenomenon had been observed independently prior to its prediction by Fresnel's theory. A temporal account of novelty would make whether a result was novel for a theory a matter of mere historical accident and that this would undermine the epistemic import novel success is supposed to have for a particular theory.

It is more plausible to argue that what matters in determining whether a result is novel is whether a particular scientist knew about the result before constructing the theory that predicts it. Call this epistemic novelty. The problem with this account of novelty is that, in some cases, that a scientist knew about a result does not seem to undermine the novel status of the result relative to their theory, because they may not have appealed to the former in constructing the latter. For example, many physicists regarded the success of general relativity in accounting for the well-known, previously anomalous orbit of Mercury as highly confirming. Consider again the case of Fresnel. If we say that the fact that the white spot phenomenon was known about is irrelevant, because Fresnel was not constructing his theory to account for it but it still predicted it, then we seem to be saying that the intentions of a theorist in constructing a theory determine in part whether the success of the theory is to be counted as evidence for its truth. Arguably, this undermines the objectivity of theory confirmation.

This motivates the idea of use novelty. A result is use-novel if the scientist did not explicitly build the result into the theory or use it to set the value of some parameter crucial to its derivation. For example, many physicists regard the success of general relativity in accounting for the orbit of Mercury, which was anomalous for Newtonian mechanics, as highly confirming, because the reasoning that led to the theory appealed to general principles and constraints that had nothing to do with the empirical data about the orbits of planets. Even though

Einstein specifically aimed to solve the Mercury problem, the derivation of the correct orbit was not achieved by putting in the right answer by hand.

There is also a modal account of novel prediction, according to which, that a theory could predict some unknown phenomenon is what matters, not whether it actually did so predict. In any case, scientific methodology includes far broader criteria for empirical success, such as providing explanations of previously mysterious phenomena. Indeed, Darwin's theory of evolution and Lyell's theory of uniformitarianism were accepted by the scientific community because of their systematising and explanatory power, and in spite of their lack of novel predictive success.

(V) Falsificationism

Popper argues that it is just too easy to accumulate positive instances which support some theory, especially when the theory is so general in its claims that its seems not to rule anything out. Similarly, some theories that have great explanatory power are scientifically dubious precisely because so much can be explained by them. Popper concludes that the 'confirmation' that a theory is supposed to get from observation of an instance which fits the theory only really counts for anything when it is an instance which was a risky prediction by the theory, that is if it is a potential falsifier of the theory. Even then it doesn't count as positive evidence for the theory, it merely shows that the theory has survived an attempt at refutation. The appropriate response is to try and find another way to try and refute it.

The problem of induction arises because no matter how many positive instances of a generalisation are observed it is still possible that the next instance will falsify it. Popper's solution to the problem of induction is simply to argue that it does not show that scientific knowledge is not justified because science does not depend on induction after all. There is a logical asymmetry between confirmation and falsification of a universal generalisation: a generalisation like all ravens are black would be falsified by a single observation of a raven that is not black. Popper argued that science is fundamentally about falsifying rather than confirming theories, and so he thought that science could proceed without induction because the inference from a falsifying instance to the falsity of a theory is purely deductive. If a theory or hypothesis is in principle unfalsifiable by experience then according to Popper is it unscientific (although it may still be meaningful).

According to falsificationism, science proceeds not by testing a theory and accumulating positive inductive support for it, but by trying to falsify theories. If it is falsified then it is abandoned, but if it is not falsified this just means it ought to be subjected to more attempts to falsify it. Popper says that the scientific method is that of 'conjectures and refutations'. Bold conjectures are those from which novel predictions can be deduced. On this view, science proceeds by natural selection and scientific knowledge is learned only from mistakes. Even the most successful theories could be falsified in the future and so they too should be regarded as conjectures. Popper argued that scientists must state clearly the conditions under which they would give up their theories rather than being committed to them come what may. It is important that on his view, no theories, no matter how well tested, ought to be regarded as even probably true. Popper may have come to this view by thinking about Newtonian mechanics, which must have seemed as well confirmed as a theory could be to a scientist in the early nineteenth century, but by the early twentieth century had been overthrown by special relativity and quantum mechanics, according to which it is quite wrong about the fundamental details of how the world works.

Nonetheless, the falsificationist does not view all scientific theories equally. Some theories are falsifiable but the phenomena they predict are not interesting or surprising. Bold conjectures that make novel predictions are the hypotheses that are scientifically valuable. Popper thought that theories could be ranked according to their degree of falsifiability and that this is the true measure of their empirical content. The more falsifiable a theory is the better it is because if it is highly falsifiable it must make precise predictions about a large range of phenomena. Popper also argued that new theories ought to be more falsifiable than the theories they replace.

Given the Duhem-Quine problem that is discussed in section 2 of this chapter, it is clear that there is no such thing as completely conclusive refutation of a theory by experiment. Popper concedes this point and so claims that as well as a set of observation statements which are potential falsifiers of the theory, there must also be a set of experimental procedures, so that the relevant group of scientists agree on a way in which the truth or falsity of each observation statement can be established. Falsification is only possible in science if there is agreement among scientists about what is being tested on any given occasion. Furthermore, Popper argues that whenever a high-level theoretical hypothesis is in conflict with a basic observation statement, it is the high-level theory that should be abandoned.

There are several problems with Popper's account of falsificationism including the following:

(i) Some legitimate parts of science seem not to be falsifiable

These fall into three categories:

(a) Probabilistic statements

The predictions derived from scientific theories are sometimes statements about the probability of some occurrence. However, such statements cannot be falsified because an improbable experimental outcome is consistent with the original statement. Any statement about the probability of a single event is not falsifiable.

(b) Existential statements

Universal generalisations are part of our scientific knowledge, but so to seem to be statements asserting the existence of things such as black holes, atoms, and viruses. These existential statements cannot be falsified. If a theory asserts the existence of something which is not found this does not deductively entail that the entity does not exist.

(c) Unfalsifiable scientific principles

It is arguable that some unfalsifiable principles may nonetheless be rightly considered part of scientific knowledge. So, for example, the status of the principle of conservation of energy, which states that energy can take different forms but cannot be created or destroyed, is such that it is inconceivable to most scientists that an experiment could falsify it; rather an apparent violation of the principle would be interpreted as revealing that something is wrong with the rest of science and it is likely that a new source, sink or form of energy would be posited.

There are also methodological principles that are arguably central to science but not falsifiable. So, for example, many scientists intuitively regard simple and unifying theories as, all other things being equal, more likely to be true than messy and complex ones. Some people claim that we have inductive grounds for believing in scientific theories that are simple, unified and so on, because in general the search for simple and unifying explanations has been fairly reliable in producing empirically successful theories, but they would add that we should never make simplicity an absolute requirement because sometimes nature is complex and untidy. Another kind of simplicity is that enshrined in the principle known as Ockham's razor, which is roughly the prescription not to invoke more entities in order to explain something than is absolutely necessary (this kind of simplicity is called ontological parsimony). It is not obvious how a falsificationist can justify these methodological rules.

(d) The hypothesis of natural selection

At one time Popper was critical of the theory of evolution because he thought the hypothesis that the fittest species survive was tautological, that is to say true by definition, and therefore not falsifiable, yet evolutionary theory is widely thought to be a prime example of a good scientific theory. Most philosophers of biology would argue that the real content of evolutionary theory lies not in the phrase 'the fittest survive', but in the idea of organisms passing on characteristics, subject to mutation and variation, which either increase or decrease the chances of their offspring surviving long enough to reproduce themselves, and so pass on those characteristics. This is supposed to account for the existence of the great diversity of species and their adaptation to the environment, and also the similarities of form and structure that exist between them. This theory may be indirectly falsifiable but it does not seem to be directly falsifiable.

(ii) Falsificationism is not itself falsifiable

Popper admits this but says that his own theory is not supposed to be because it is a philosophical or logical theory of the scientific method, and not itself a scientific theory so this objection, although often made, misses its target.

(iii) The notion of degree of falsifiability is problematic

The set of potential falsifiers for a universal generalisation is always infinite, so there can be no absolute measure of falsifiability, but only a relative one. The Duhem problem means that judgements about the degree of falsifiability of theories are relative to whole systems of hypotheses, and so our basis for such judgements is past experience and this lets induction in by the back door.

(iv) Falsificationism cannot account for our expectations about the future

Popper says that we are not entitled to believe that our best theories are even probably true. His position is ultimately extremely sceptical, indeed he goes further than Hume, who says induction cannot be justified but that we cannot help but use it, and argues that scientists should avoid induction altogether. But is this really possible, and is it really plausible to say that we never get positive grounds for believing scientific theories?

Our scientific knowledge does not seem to be purely negative and if it were it would be hard to see why we have such confidence in certain scientifically informed beliefs. After all, it is because doctors believe that penicillin fights bacterial infection that they prescribe it for people showing the relevant symptoms. The belief that certain causes do indeed have certain effects is what informs our actions. According to Popper there is no positive inductive support for my belief that if I try to leave the top floor of the building by jumping out the window I will fall hard on the ground and injure myself. If observation of past instances really confers no justification on a generalisation then I am just as rational if I believe that when I jump out of the window I will float gently to the ground. This is an unacceptable consequence of Popper's views for there is nothing more obvious to most of us than that throwing oneself out of high windows when one wishes to reach the ground safely is less rational than taking the stairs. If we adopt Popper's nihilism about induction we have no resources for explaining why people behave the way they do, and furthermore we are obliged to condemn any positive belief in generalisations as unscientific.

Of course, just when and how we can be justified on the basis of experience in believing general laws and their consequences for the future behaviour of the natural world is the problem of induction. But most philosophers think that solving this problem is not a matter of deciding whether it is more rational to take the stairs but why it is more rational to do so. Popper's response to this challenge is to introduce the notion of corroboration; a theory is corroborated if it was a bold conjecture that made novel predictions that were not falsified. Popper says that it is rational to suppose that the most corroborated theory is true because we have tried to prove it false in various ways and failed. The most corroborated theory is not one we have any reason to believe to be true, but it is the one we have least reason to think it is false, so it is rational to use it in making plans for the future, like leaving the building by the stairs and not by jumping. Popper stresses that the fact that a theory is corroborated only means that it invites further challenges.

But the notions of boldness and novelty are historically relative; the former means unlikely in the light of background knowledge and therefore highly falsifiable, and novel means previously unknown, or unexpected given existing corroborated theories, so once again induction based on past experience is smuggled into Popper's account. Furthermore, there is an infinite number of best corroborated theories, because whatever our best corroborated theory is, we can construct an infinite number of theories that agree with what it says about the past, but which say something different about what will happen in the future. The theory that gravity always applies to me when I jump into the air except after today is just as corroborated by all my experience up to now as the alternative that tells me not to jump off tall buildings; again we seem to have no choice but to accept the rationality of at least some inductive inferences despite what Popper says.

(v) Scientists sometimes ignore falsification

In general, contrary to what Popper says, scientists are not prepared to state in advance under what conditions they would abandon their most fundamental assumptions, and indeed many scientists probably would not consider abandoning the idea that species evolve by natural selection, or that ordinary matter is made of atoms. There are also many cases in the history of science where, in the face of falsifying evidence, scientists thought up modifications to save a theory instead of abandoning it. Popper distinguishes between ad hoc and non-ad hoc modifications, where the latter give rise to extra empirical content and the former do not, and argues that only non-ad hoc modifications are acceptable. There are certainly extreme cases where most people will agree that a theory has only been saved from refutation by a gratuitous assumption whose only role or justification is to save the theory.

Unfortunately, it turns out that there are cases in the history of science where a falsifying observation is tolerated for decades despite numerous attempts to account for it. More generally, it often seems to be the case that where scientists have a successful theory, the existence of falsifying observations will not be sufficient to cause the abandonment of the theory in the absence of a better alternative.

For these and other reasons, Popper's falsificationism is probably now more popular among scientists than it is among philosophers.

(VI) Kuhn's Philosophy of Science

The scientific method is supposed to be rational, and to give us objective knowledge of the world. Prior to the work of Thomas Kuhn many philosophers of science agreed with the following statements:

(i) Science is cumulative.

(ii) Science is unified in the sense that there is a single set of fundamental methods for all the sciences, and in the sense that the natural sciences at least are all ultimately reducible to physics.

(iii) There is a epistemologically crucial distinction between the context of discovery and the context of justification.

(iv) There is an underlying logic of confirmation or falsification implicit in all scientific evaluations of the evidence for some hypothesis. Such evaluations are value-free in the sense of being independent of the personal non-scientific views and allegiances of scientists.

(v) There is a sharp distinction (or demarcation) between scientific theories and other kinds of belief systems.

(vi) There is a sharp distinction between observational terms and theoretical terms, and also between theoretical statements and those that describe the results of experiments. Observation and experiment is a neutral foundation for scientific knowledge, or at least for the testing of scientific theories.

(vii) Scientific terms have fixed and precise meanings.

Kuhn argued that many scientists' accounts of the history of their subject considerably simplify and distort the real stories of theory development and change. He argues that the history of science does not consist in the steady accumulation of knowledge, but often involves the wholesale abandonment of past theories. According to Kuhn, the evaluation of theories depends on local historical circumstances, and his analysis of the relationship between theory and observation suggests that theories infect data to such an extent that no way of gathering of observations can ever be theory-neutral and objective. Hence, the degree of confirmation an experiment gives to a hypothesis is not objective, and there is no single logic of theory testing which can be used to determine which theory is most justified by the evidence. He thinks instead that scientists' values help determine, not just how individual scientists develop new theories, but also which theories the scientific community as a whole regards as justified.

There are two closely related ideas of paradigm in Kuhn's work, namely those of paradigm as disciplinary matrix and paradigm as exemplar. Kuhn argues that before scientific inquiry can even begin in some domain, the scientific community in question has to agree upon answers to fundamental questions about, for example: what kinds of things exist in the universe, how they interact with each other and our senses, what kinds of questions may legitimately be asked about these things, what techniques are appropriate for answering those questions, what counts as evidence for a theory, what questions are central to the science, what counts as a solution to a problem, what counts as an explanation of some phenomenon, and so on.

A disciplinary matrix is a set of answers to such questions that are learned by scientists in the course of the education which prepares them for research, and that provide the framework within which the science operates. It is important that different aspects of the disciplinary matrix may be more or less explicit, and some parts of it are constituted by the shared values of scientists, in that they prefer certain types of explanation over others and so on. It is also important that some aspects of it will consist of practical skills and methods that are not necessarily expressible in words.

Exemplars, on the other hand, are those successful parts of science which all beginning scientists learn, and which provide them with a model for the future development of their subject. Anyone familiar with a modern scientific discipline will recognise that teaching by example plays an important role in the training of scientists. Textbooks are full of standard problems and solutions to them, and students are set exercises that require them to adapt the techniques used in the examples to new situations. The idea is that, by repeating this process, eventually, if they have the aptitude for it, students will learn how to apply these techniques to new kinds of problems which nobody has yet managed to solve.

Most science is what Kuhn calls 'normal science', because it is conducted within an established paradigm. It involves elaborating and extending the success of the paradigm, for example, by gathering lots of new observations and accommodating them within the accepted theory, and trying to solve minor problems with the paradigm. Hence, normal science is often said to be a 'puzzle-solving' activity, where the rules for solving puzzles are quite strict and determined by the paradigm. According to Kuhn, most of the everyday practice of science is a fairly conservative activity in so far as, during periods of normal science, scientists do not question the fundamental principles of their discipline. If a paradigm is successful and seems able to account for the bulk of the phenomena in its domain, and if scientists are still able to make progress solving problems and extending its empirical applications, then most scientists will just assume that anomalies that seem intractable will eventually be resolved. They won't give up the paradigm just because it conflicts with some of the evidence.

However, sometimes scientists become aware of anomalies which won't go away no matter how much effort is put into resolving them. These may take the form of conceptual paradoxes or experimental falsifications. Even these will not necessarily cause much serious questioning of the basic assumptions of the paradigm. But when a number of serious anomalies accumulate then, some, often younger or maverick scientists will begin to question some of the core assumptions of the paradigm, and perhaps they will begin speculating about alternatives. This amounts to the search for a new paradigm, which is a new way of thinking about the world. If this happens when successful research within the paradigm is beginning to decline, more and more scientists may begin to focus their attention on the anomalies and the perception that the paradigm is in 'crisis' may begin to take hold of the scientific community. If a crisis happens, and if a new paradigm is adopted by the scientific community, then a 'revolution' or 'paradigm shift' has occurred. On Kuhn's view when a revolution occurs the old paradigm is replaced wholesale. So, for example, the adoption or rejection of each of the examples of paradigms listed above is a scientific revolution.

There are two points about Kuhn's account of this and other scientific revolutions that must be emphasised:

- This is a completely different view of scientific change to the traditional idea of cumulative growth of knowledge, because paradigm shifts or scientific revolutions involve change in scientific theories that is not piecemeal but holistic. In other words, the paradigm does not change by parts of it being changed bit by bit, but rather by a wholesale shift to a new way of thinking about the world, and this will usually mean a new way of practising science as well including new experimental techniques and so on.

- Revolutions only happen when a viable new paradigm is available, and also when there happen to be individual scientists who are able to articulate the new picture to their colleagues.

Kuhn also emphasises the role of psychological and sociological factors in disposing scientists to adopt or a reject a particular paradigm.

Although existing theories guide us in developing new theories, and tell us which observations are significant and so on, the distinction between the context of discovery and the context of justification can be invoked to maintain the idea that scientific theories are tested by observations. Many empiricist philosophers have drawn a sharp distinction between the observational and the theoretical, and both logical positivists, and Popper, at least in his earlier work, assume it. According to the received view, the theory-independence or neutrality of observable facts makes them a suitable foundation for scientific knowledge, or at least for testing theories. The received view incorporated a distinction between observational terms, like 'red', 'heavy', and 'wet', and theoretical terms, like 'electron', 'charge', and 'gravity'. The idea is that the rules for the correct application of observational terms refer only to what a normal human observer perceives in certain circumstances, and that they are entirely independent of theory. So, for example, Ernest Nagel argues that every observational term is associated with at least one overt procedure for applying the term to some observationally identifiable property, when certain specified circumstances are realised. So, for example, the property of being red is applied to an object when it looks red to a normally functioning observer in normal lighting conditions. Many other writers analyse the logic of theory testing relying upon this distinction between observational and theoretical terms.

Incommensurability is a term from mathematics which means 'lack of common measure'. It was adopted by Kuhn, and another philosopher called Paul Feyerabend, both of whom argued that successive scientific theories are often incommensurable with each other in the sense that there is no neutral way of comparing their merits. One of the most radical ideas to emerge from Kuhn's work is that what counts as the evidence in a given domain may depend upon the background paradigm. If this is right then how can it be possible to rationally compare competing paradigms? Kuhn argues that there is no higher standard for comparing theories than the assent of the relevant community, and that, the choice between competing paradigms is a choice between incompatible modes of community life.

In his later work Kuhn sought to distance himself from extreme views which give no role to rationality in the progress of science, and which do not allow for comparison of the merits of theories within different paradigms. He argues that the following five core values are common to all paradigms:

- A theory should be empirically accurate within its domain.
- A theory should be consistent with other accepted theories.
- A theory should be wide in scope and not just accommodate the facts it was designed to explain.
- A theory should be as simple as possible.
- A theory should be fruitful in the sense of providing a framework for ongoing research.

Hence, Kuhn avoids complete irrationalism because these values impose some limits on what theories scientists can rationally accept. On the other hand, these values are not sufficient to determine what decisions they ought to make in most interesting cases, because these values may conflict; a theory may be simple but not accurate, or fruitful but not wide in scope, and so on. Furthermore, a value like simplicity may be understood in different ways depending on background views and so on.

(VII) Bayesianism

Bayesianism is potentially a theory of the relationship between theories and evidence, a theory of rationality, an account of the scientific method, and a theory of probability. It is increasingly being taught and used in place of traditional statistics. Bayesianism employs the mathematical theory of probability. Probabilities of the form $P(A)$ are used to represent a subject's degree of belief that a proposition A is the case. The probabilities must conform to the constraints of the probability calculus ('$\vdash$': entails):

$$0 \leq P(A) \leq 1$$
$$\text{necessarily } A \vdash P(A) = 1$$
$$A \rightarrow \neg B \vdash P(A \text{ or } B) = P(A) + P(B)$$
$$P(\neg A) = 1 - P(A)$$
$$(A \vdash B) \vdash P(B) \geq P(A)$$
$$A \leftrightarrow B \vdash P(A) = P(B)$$
$$A, B \text{ are independent events} \vdash P(A \& B) = P(A) \times P(B)$$

The notation $P(A/B)$ is used to represent the conditional probability of A given B, defined:

$$P(A/B) = P(A\&B)/P(B)$$

which allows Bayes' theorem to be proved:

$$P(h/e) = P(h).P(e/h)/P(e)$$

This equation can be understood as ruling how to update belief in h in the light of new information e. Suppose, h is a scientific hypothesis (or more accurately a combination of hypotheses) and e is the statement that some phenomenon is observed under certain conditions, and that h predicts that e will be true when a test is performed. Let the scientist's prior degree of belief in the hypothesis be $P(h)$, and $P(e)$ the scientist's degree of belief, disregarding h, that the phenomenon will be observed. $P(e/h)$ is the scientist's degree of belief as to how likely e is given that h is true. This is known as 'Bayesian conditionalisation'.

Some intuitive aspects of confirmation are captured by this formalism. Firstly, $P(h/e)$ is proportional to $P(h)$, in other words, the more likely the hypothesis was considered to start with, the more likely it will be considered even in the light of new, possibly disconfirming, evidence. Second, $P(h/e)$ is proportional to $P(e/h)$, in other words, the more closely linked the evidence and the hypothesis are the more observation of the evidence supports the hypothesis. Finally, $P(h/e)$ is inversely proportional to $P(e)$, in other words, the more unlikely the evidence which the hypothesis predicts the more it supports the hypothesis if it is observed. Bayesians claim that evidence confirms a hypothesis if learning the evidence raises the probability of the hypothesis and disconfirms it if learning the evidence lowers that probability.

Bayesianism is alleged to be able to resolve a number of the well-known paradoxes of confirmation, in particular, the Ravens Paradox, the tacking paradox, and Goodman's New Problem of Induction. It is also claimed that Bayesianism deals with the problem of underdetermination.

There are various problems with Bayesianism including the following:

(i) The Problem of Old Evidence.

This is the problem of explaining how it is possible for a theory to be confirmed by evidence that was known about before the theory was formulated. One common response is to argue that the relevant prior probability to be used in Bayes' formula is a counterfactual probability, namely the probability the theory would have been judged to have had the evidence not been known about.

(ii) The Problem of the Priors

According to Bayesianism, how credible a scientific theory is seems to be a function of its prior probability. Must Bayesians therefore provide an account of what the priors ought to be? There are various theorems that show that agents who start with very different prior degrees of belief will nonetheless end up converging in their posterior degrees of belief if they keep updating their degrees of belief on the basis of the same experimental data. So it seems that eventually differences in priors will be irrelevant. On the other hand, these theorems only say what will happen in the (potentially infinitely) long run, and yet we expect scientists to reach agreement about the status of scientific hypotheses in reasonable short amounts of time.

(iii) The Interpretation of Probability

There is some controversy about the nature of the probabilities used in Bayesianism. Some Bayesians believe that they can only represent subjective degrees of belief of agents, while others think that they can be understood as referring to degrees of belief that have to match up to objective probabilities of some kind. Ramsey argued for the former and claimed that degrees of belief can be interpreted as corresponding to the least odds someone would be willing to gamble on.

(iv) The Problem of Psychological Implausibiltity

Consistency with the probability calculus seems to require a lot of agents. Arguably it is psychologically unrealistic to expect people in general and scientists in particular, to have no inconsistencies in their beliefs, and to continually deduce all the logical consequences of their beliefs as they acquire new ones.

(v) The Status of 'Dutch Book' Arguments

There are two kinds of Dutch book arguments namely synchronic and diachronic. The former are intended to show that a rational agents degrees of belief at any given time must satisfy the axioms of the probability calculus, while the latter are intended to show that rational agents ought to update their degrees of belief over time in accordance with Bayesian conditionalisation. The former is relatively uncontroversial and is often regarded as the probabilistic analogue of logical consistency, however the latter has been the subject of intense debate since it appears that the most such arguments can show is that if you reject conditionalisation you would be bound to lose if you bet honestly with someone who knows your strategy for changing your betting quotients.

Bayesianism is actually vastly more complicated than the above discussion suggests since, for example, there are other forms of conditionalisation based on how degrees of belief ought to change in the light of uncertain evidence.

BIBLIOGRAPHY

1. ONTOLOGICAL POSITIONS

(I) Natural Kinds

[Bird, 1998] A. Bird. *Philosophy of Science*. London: UCL Press, chapter 3, 1998.

[Dupre, 1993] J. Dupre. *The Disorder of Things: Metaphysical Foundations of the Disunity of Science*. Cambridge MA: Havard University Press, 1993.

[Dupre, 1981] J. Dupre. Natural kinds and biological taxa. *Philosophical Review*, 90: 66–90, 1981.

[Mellor, 1977] D. H. Mellor. Natural kinds. *The British Journal for the Philosophy of Science*, 28: 299–312, 1977; and in his *Matters of Metaphysics*, Cambridge: Cambridge University Press, 1991.

[Putnam, 1975] H. Putnam. The meaning of meaning. In his *Philosophical Papers, volume ii: Mind, Language and Reality*, Cambridge: Cambridge University Press, 1975.

[Quine, 1969] W. v. O. Quine. Natural kinds. In N. Rescher (ed.), *Essays in Honour of Carl Hempel*, Dordrecht: Reidel, 1969

(II) Truth and Mind-independence

[Blackburn and Simmons, 1999] S. Blackburn and K. Simmons (eds). *Truth*. Oxford: Oxford University Press, 1999.

[Devitt, 1989] M. Devitt. *Realism and Truth*. Oxford: Blackwell, 1989 (second edition 1991).

[Dummett, 1978] M. Dummett. *Truth and other enigmas*. London: Duckworth, 1978.

[Engel, 2002] P. Engel. *Truth*. Chesham: Acumen, 2002.

[Goodman, 1978] N. Goodman. *Ways of Worldmaking*. Indianapolis, IL: Hackett, 1978.

[Horwich, 1990] P. Horwich. *Truth*. Oxford: Oxford University Press, 1990 (second edition 1998).

[Kukla, 2000] A. Kukla. *Social Constructivism and the Philosophy of Science*. London: Routledge, 2000

[James, 1897] W. James. *The Will to Believe and Other Essays*. New York: Dover, 1897.

[James, 1907] W. James. *Pragmatism*. New York: Longmans and Green, 1907.

[Putnam, 1981] H. Putnam. *Reason, Truth and History*. Cambridge: Cambridge University Press, 1981.

[Wright, 1987] C. Wright. *Realism, Meaning and Truth*. Oxford: Oxford University Press, 1987 (second edition 1983).

[Wright, 1992] C. Wright. *Truth and Objectivity*. Harvard: Harvard University Press, 1992.

(III) Properties and Universals

[Armstrong, 1989] D. M. Armstrong. *Universals: An Opinionated Introduction*. Boulder, Colorado: Westview Press, 1989.

[Loux, 2002] M. Loux. *Metaphysics: a contemporary introduction*. London Routledge, chapters 1&2, 2002.

[Mellor and Oliver, 1997] D. H. Mellor and A. Oliver (eds.). *Properties*. Oxford: Oxford University Press, 1997.

[Russell, 1912] B. Russell. *The Problems of Philosophy*. Oxford: Oxford University Press, chapters 9–10, 1912.

[Kim and Sosa, 1999] J. Kim and E. Sosa (eds.). *Metaphysics: An Anthology*. Oxford: Blackwell, part IV, 1999.

(IV) Identity and Individuality

[Black, 1952] M. Black. The identity of indiscernibles. *Mind*, 61: 153–164, 1952.

[Loux, 2002] M. Loux. *Metaphysics: A Contemporary Introduction*. London Routledge, chapter 3, 2002.

[Kim and Sosa, 1999] J. Kim and E. Sosa (eds.). *Metaphysics: An Anthology*. Oxford: Blackwell, part II, 1999.

[Laurence and Macdonald, 1998] S. Laurence and C. Macdonald. *Contemporary Readings in the Foundations of Metaphysics*. Oxford: Blackwell, part IV, 1998.

(V) Matter and Motion

[Salmon, 2001] W. Salmon (ed.). *Zeno's Paradoxes*. Indianapolis, IL: Hackett, 2001.

[Jammer, 1961] M. Jammer. *Concepts of Mass in Classical and Modern Physics*. New York: Dover, 1961.

(VI) Causation

[Hume, 1963] D. Hume. *An Enquiry Concerning Human Understanding.* La Salle: Open Court, section VII, 1963.

[Hume, 1978] D. Hume. *A Treatise of Human Nature.* Oxford: Oxford University Press, I.3, 14/15, 1978.

[Mackie, 1980] J. L. Mackie. *The Cement of the Universe: A Study of Causation.* Oxford: Clarendon Press, 1980.

[Sosa and Tooley, 1993] E. Sosa and M. Tooley (eds.). *Causation.* Oxford; Oxford University Press, 1993.

[Kim and Sosa, 1999] J. Kim and E. Sosa (eds.). *Metaphysics: An Anthology*, Oxford: Blackwell, part VII, 1999.

(VII) Laws of Nature

[Bird, 1998] A. Bird. *Philosophy of Science.* London: UCL Press, chapter 1, 1998.

[Carroll, 2004] J. Carroll (ed.). *Readings on Laws of Nature.* Pittsburgh: University of Pittsburgh Press, 2004.

[Armstrong, 1983] D. Armstrong. *What is a Law of Nature?.* Cambridge: Cambridge University Press, 1983.

[Cartwright, 1983] N. Cartwright. *How the Laws of Physics Lie.* Oxford: Oxford University Press, 1983.

(VIII) Probability, Propensity and Dispositions

[Mellor, 2005] D. H. Mellor. *Probability: a philosophical introduction.* London: Routledge, 2005.

[Ellis, 2001] B. Ellis. *Scientific Essentialism.* Cambridge: Cambridge University Press, 2001.

(IX) Reductionism, Emergence and Supervenience

[Churchland, 1990] P. M. Churchland. *Matter and Consciousness.* Cambridge, MA: MIT Press, chapter 2, 1990.

[Jackson, 1998] F. Jackson. *From Metaphysics to Ethics: A Defence of Conceptual Analysis.* Oxford: Oxford University Press, 1998

[Kim, 1998] J. Kim. *Mind in a Physical World.* Cambridge, MA: MIT Press, 1998.

(X) Space, Time and Spacetime

[Dainton, 2001] B. Dainton. *Time and Space.* Chesham: Acumen, 2001.

[Huggett, 1999] N. Huggett. *Space from Zeno to Einstein: Classic Readings with a Contemporary Commentary.* Cambridge, MA: MIT Press, 1999.

[le Poidevin and McBeath, 1993] R. le Poidevin and M. McBeath (eds.). *The Philosophy of Time.* Oxford: Oxford University Press, 1993.

[Lockwood, 2005] M. Lockwood. *The Labyrinth of Time.* Oxford: Oxford University Press, 2005.

[Mellor, 1998] H. Mellor. *Real Time II.* London: Routledge, 1998.

[Sklar, 1974] L. Sklar. *Space, Time and Spacetime*, Berkeley: University of California Press, 1974.

(XI) Events and Processes

[Casati, and Varzi, 1996] R. Casati, and A. C.Varzi (eds.). *Events.* Dartmouth, Aldershot, 1996.

[Davidson, 1980] D. Davidson. *Essays on Actions and Events.* Oxford: Clarendon Press, 1980.

[Rescher, 1996] N. Rescher. *Process Metaphysics: An Introduction to Process Philosophy.* New York: SUNY Press, 1996.

[Rescher, 2000] N. Rescher. *Process Philosophy: A Survey of Basic issues.* Pittsburgh, Pa.: University of Pittsburgh Press, 2000.

[Whitehead, 1919] A. N. Whitehead. *An Enquiry Concerning the Principles of Natural Knowledge*. Cambridge: Cambridge University Press, 1919; reprinted New York: Kraus Reprints, 1982.

2. EPISTEMOLOGICAL POSITIONS

(I) Rationalism

[Cottingham, 1988] J. Cottingham. *The Rationalists*. Oxford: Oxford University Press, 1988.

[Boghossian and C.Peacock, 2000] P. Boghossian and C. Peacock (eds). *New Essays on the A Priori*. Oxford: Oxford University Press, 2000.

(II) Empiricism

[Ayer, 1952] A. J. Ayer. *Language, Truth and Logic*. New York: Dover, 1952.

[Gower, 1997] B. Gower. *Scientific Method: An Historical and Philosophical Introduction*. London: Routledge, 1997.

[Hanfling, 1981] H. Hanfling, (ed.). *Essential Readings in Logical Positivism*. Oxford: Blackwell, 1981.

[Shapin, 1996] S. Shapin. *The Scientific Revolution*. Chicago: Chicago University Press, 1996.

[Woolhouse, 1988] R. Woolhouse. *The Empiricists*. Oxford: Oxford University Press, 1988.

(III) Induction

[Russell, 1912] B. Russell. *The Problems of Philosophy*. Oxford: Oxford University Press, chapter 6, 1912.

[Goodman, 1973] N. Goodman. *Fact, Fiction and Forecast*. Indianapolis, IL: Bobbs-Merrill, 1973.

[Ladyman, 2002] J. Ladyman. *Understanding Philosophy of Science*. London: Routledge, chapter 2, 2002.

[Hume, 1963] D. Hume. *An Enquiry Concerning Human Understanding*. La Salle: Open Court, section IV, 1963.

[Hume, 1978] D. Hume. *A Treatise of Human Nature*. Oxford: Oxford University Press, part III, 1978.

[Swinburne, 1974] R. Swinburne (ed.). *Justification of Induction*. Oxford: Oxford University Press, 1974.

(IV) Scientific Realism

[Kitcher, 1993] P. Kitcher. *The Advancement of Science: Science without Legend, Objectivity without Illusions*. Oxford: Oxford University Press, 1993.

[Ladyman, 2002] J. Ladyman. *Understanding Philosophy of Science*. London: Routledge, chapter 5, 2002.

[Psillos, 1996] S. Psillos. *Scientific Realism: How Science Tracks Truth*. London: Routledge, 1996.

[Van Fraassen, 1980] B. C. Van Fraassen. *The Scientific Image*. Oxford: Oxford University Press, 1980.

[Churchland and Hooker, 1985] P. Churchland, and C. A. Hooker (eds). *Images of Science*. Chicago: University of Chicago Press, 1985.

(V) The Duhem-Quine Problem and Underdetermination

[Duhem, 1906] P. Duhem. *The Aim and Structure of Physical Theory.* Translated by P. Wiener 1954, Princeton: Princeton University Press, chapter 6, 1906.

[Quine, 1953] W. Quine. Two dogmas of empiricism. In his *From a Logical Point of View*, Cambridge, MA: Harvard University Press, 1953.

[Harding, 1976] S. Harding, (ed.). *Can Theories be Refuted?: Essays on the Duhem-Quine Thesis.* Dordrecht, Netherlands: D.Reidel, 1976.

[Ladyman, 2002] J. Ladyman. *Understanding Philosophy of Science.* London: Routledge, chapter 6, 2002.

[Kukla, 1996] A. Kukla. Does every theory have empirically equivalent rivals? *Erkenntnis*, 44: 137–66, 1996.

[Kukla, 1993] A. Kukla. Laudan, Leplin, empirical equivalence, and underdetermination. *Analysis*, 53: 1–7, 1993.

[Kukla, 1998] A. Kukla. *Studies in Scientific Realism.* Oxford: Oxford University Press, 1998.

[Laudan and Leplin, 1991] L. Laudan, and J. Leplin. Empirical equivalence and underdetermination. *Journal of Philosophy*, 88: 269–85, 1991.

[Laudan and Leplin, 1993] L. Laudan, and J. Leplin. Determination underdeterred. *Analysis*, 53: 8–15, 1993.

[Hoefer and Rosenberg, 1994] C. Hoefer and A. Rosenberg. Empirical equivalence, underdetermination, and systems of the world. *Philosophy of Science*, 61: 592–607, 1994.

(VI) Inference to the Best Explanation

[Harman, 1965] G. Harman. Inference to the best explanation. *Philosophical Review.* 74: 88–95, 1965.

[Ladyman, 2002] J. Ladyman. *Understanding Philosophy of Science.* London: Routledge, chapter 7, 2002.

[Ladyman *et al.*, 1997] J. Ladyman, I. Douven, L. Horsten, and van B. C. Fraassen. In defence of Van Fraassen's critique of abductive reasoning: A reply to Psillos. *Philosophical Quarterly*, 47: 305–321, 1997.

[Lipton, 1991] P. Lipton. *Inference to the Best Explanation.* London: Routledge, 1991.

[Van Fraassen, 1989] B. C. Van Fraassen. *Laws and Symmetry.* Oxford: Oxford University Press, 1989.

[Psillos, 1996] S. Psillos. *Scientific Realism: how science tracks truth.* London: Routledge, chapter 9, 1996.

(VII) Arguments from Theory Change

[Hardin and Rosenberg, 1982] C. L. Hardin and A. Rosenberg. In defence of convergent realism. *Philosophy of Science*, 49: 604–615, 1982.

[Kitcher, 1993] P. Kitcher. *The Advancement of Science: Science without Legend, Objectivity without Illusions.* Oxford: Oxford University Press, 1993.

[Ladyman, 2002] J. Ladyman. *Understanding Philosophy of Science.* London: Routledge, chapter 8, 2002.

[Ladyman, 1998] J. Ladyman. What is structural realism? *Studies in History and Philosophy of Science*, 29: 409–424, 1998.

[Laudan, 1981] L. Laudan A confutation of convergent realism. *Philosophy of Science*, 48: 19–49, 1981; reprinted in D. Papineau (ed.), *Philosophy of Science*, Oxford: Oxford University Press, 1996.

[Laudan, 1984] L. Laudan. Discussion: Realism without the real. *Philosophy of Science*, 51: 156–162, 1984.

[Psillos, 1996] S. Psillos. *Scientific Realism: How Science Tracks Truth.* London: Routledge, 1996.

[Worrall, 1989] J. Worrall. Structural realism: The best of both worlds? *Dialectica*, 3: 99–124; reprinted in D. Papineau (ed.), *Philosophy of Science*, 1996.

(VIII) Contemporary Empiricism

[Ladyman, 2002] J. Ladyman. *Understanding Philosophy of Science*. London: Routledge, chapter 6, 2002.

[Ladyman, 2000] J. Ladyman. What's really wrong with constructive empiricism?: van Fraassen and the metaphysics of modality. *The British Journal for the Philosophy of Science*, 51: 837–856, 2000.

[Monton and van Fraassen, 2003] B. Monton and B. C. van Fraassen. Constructive empiricism and modal nominalism. *The British Journal for the Philosophy of Science*, 54: 405–422, 2003.

[Rosen, 1984] G. Rosen. What is constructive empiricism? *Philosophical Studies*, 74: 143–178, 1984.

[Van Fraassen, 1981] B. C. Van Fraassen. *The Scientific Image*. Oxford: Oxford University Press, 1981.

[Van Fraassen, 1989] B. C. Van Fraassen. *Laws and Symmetry*. Oxford: Oxford University Press, 1989.

[Van Fraassen, 2002] B. C. Van Frassen. *The Empirical Stance*. New Haven: Yale University Press, 2002.

(IX) Pragmatism

[Hacking, 1983] I. Hacking. *Representing and Intervening*. Cambridge: Cambridge University Press, 1983.

(X) Naturalism and Normativity

[Papineau, 1993] D. Papineau. *Philosophical Naturalism*. Oxford: Blackwell, 1993.

3. METHODOLOGICAL POSITIONS

(I) Inductivism

[Achinstein, 1991] P. Achinstein. *Particles and Waves*. Oxford: Oxford University Press, 1991.

[Gower, 1997] B. Gower. *Scientific Method: An Historical and Philosophical Introduction*. London: Routledge, chapter 3, 1997.

[Ladyman, 2002] J. Ladyman. *Understanding Philosophy of Science*. London: Routledge, chapter 1, 2002.

[Urbach, 1987] P. Urbach. *Francis Bacon's Philosophy of Science: An Account and a Reappraisal*. LaSalle, IL: Open Court, 1987.

[Woolhouse, 1988] R. Woolhouse. *The Empiricists*. Oxford: Oxford University Press, chapter 2, 1988.

(II) The Context of Discovery and the Context of Justification

[Popper, 1934-1959] K. Popper. *The Logic of Scientific Discovery*. London: Hutchinson, 1934-1959.

[Newton-Smith, 1987] W. Newton-Smith. *The Rationality of Science*. London: Routledge and Kegan Paul, chapter III, 1987.

(III) The Paradoxes of Confirmation

[Brown, 1977] H. Brown. *Perception, Theory and Commitment*. Chicago: University of Chicago Press, 1977.

[Glymour, 1980] G. Glymour. *Theory and Evidence*. Princeton, N.J.: Princeton University Press, 1980.

(IV) Explanation versus Prediction

[Hempel, 1965] C. Hempel. *Aspects of Scientific Explanation*. New York: Free Press, 1965
[Nagel, 1961] E. Nagel. *The Structure of Science*. New York: Harcourt, Brace, 1961.
[Salmon, 1989] W. C. Salmon Four decades of scientific explanation. In P. Kitcher, and W. C. Salmon (eds), *Scientific Explanation: Minnesota Studies in the Philosophy of Science*, Volume XIII, pages 3–219, 1989.
[Friedman, 1974] M. Friedman. Explanation and scientific understanding. *Journal of Philosophy*, LXXI: 5–19, 1974.

(V) Falsificationism

[Popper, 1934-1959] K. Popper. *The Logic of Scientific Discovery*. London: Hutchinson, 1934-1959.
[Popper, 1969] K. Popper. *Conjectures and Refutations*. London: Routledge and Kegan Paul, 1969.
[Ladyman, 2002] J. Ladyman. *Understanding Philosophy of Science*. London: Routledge, chapter 3, 2002.
[Lakatos, 1968] I. Lakatos. Criticism and the methodology of scientific research programmes. *Proceedings of the Aristotelian Society*, 69: 149–86, 1968.
[Lakatos and Musgrave, 1970] I. Lakatos and A. Musgrave (eds.). *Criticism and the Growth of Knowledge*. Cambridge: Cambridge University Press, 1970.
[Newton-Smith, 1987] W. Newton-Smith. *The Rationality of Science*. London: Routledge and Kegan Paul, chapter III, 1987.

(VI) Kuhn's Philosophy of Science

[Feyerabend, 1977] P. Feyerabend. *Against Method*. London: New Left Books, 1977.
[Hacking, 1981] I. Hacking (ed.). *Scientific Revolutions*. Oxford: Oxford University Press, 1981.
[Hoyningen-Huene, 1993] P. Hoyningen-Huene. *Reconstructing Scientific Revolutions: Thomas Kuhn's Philosophy of Science*. Chicago: University of Chicago Press, 1993.
[Kuhn, 1957] T. S. Kuhn. *The Copernican Revolution: Planetary Astronomy in the Development of Western Thought*. Cambridge, Mass.: Harvard University Press, 1957.
[Kuhn, 1962] T. S. Kuhn. *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press, 1962 (second edition 1970).
[Kuhn, 1977] T. S. Kuhn. *The Essential Tension*. Chicago: University of Chicago Press, 1977.
[Ladyman, 2002] J. Ladyman. *Understanding Philosophy of Science*. London: Routledge, chapter 4, 2002.
[Lakatos and Musgrave, 1970] I. Lakatos and A. Musgrave (eds.). *Criticism and the Growth of Knowledge*. Cambridge: Cambridge University Press, 1970.
[Laudan, 1977] L. Laudan. *Progress and its Problems*. Berkeley: University of California Press, 1977.

(VII) Bayesianism

[Bovens and Hartmann, 2003] L. Bovens and S. Hartmann. *Bayesian Epistemology*. Oxford: Clarendon Press, 2003.
[Horwich, 1982] P. Horwich. *Probability and Evidence*. Cambridge: Cambridge University Press, 1982.
[Howson and Urbach, 1993] C. Howson and P. Urbach. *Scientific Reasoning: The Bayesian Approach*. Open Court, 2^{nd} edition, 1993.
[Glymour, 1980] G. Glymour. *Theory and Evidence*. Princeton, N.J.: Princeton University Press, 1980.

REDUCTION, INTEGRATION, AND THE UNITY OF SCIENCE: NATURAL, BEHAVIORAL, AND SOCIAL SCIENCES AND THE HUMANITIES

William Bechtel and Andrew Hamilton

1 A HISTORICAL LOOK AT UNITY

The notion that science is unified in one way or another dates back at least to Aristotle, though unity claims since then have been diverse and variously motivated. By way of introduction to the modern discussion of unity, disunity, and integration, in this first section we examine five historical attempts to unify knowledge: Aristotle's metaphysical and hierarchical unity; the Enlightenment project of the French Encyclopedists; the systematic unity of *Naturphilosoph* Lorenz Oken; the methodological unity of the Vienna School's *Encyclopedia of Unified Science*; and finally, the organizational unity of cybernetics and general systems theory. We treat these unification projects not only as context, but also because, as we shall see, something of their momentum carries over into the modern discussion.

1.1 Aristotle's Metaphysical and Hierarchical Unity

Aristotle arranged the 'sciences' into three divisions: the theoretical sciences (metaphysics, mathematics, and physics): the practical sciences (e.g., ethics and politics), and the productive sciences (e.g., poetry and rhetoric). That is, he divided sciences according to their purposes. Theoretical sciences are concerned with knowledge alone and for its own sake, practical sciences are for doing, and productive sciences are for making. Despite these divisions, however, Aristotle's image of the sciences was one of a unified hierarchy. In the *Metaphysics*, he made clear that the theoretical sciences — most particularly metaphysics or 'theology' — are at the top of the hierarchy. These are the sciences that investigate first causes, and the people who know them know universally and in the highest degree, as well as "understand... all the underlying subjects" (*Metaphysics* A.2).

Aristotle argued that the theoretical sciences are the most basic. It is by virtue of theoretical knowledge that one has true command of practical and productive matters. Without theory, one merely has experience. With theory, one has art (*techné*). Consider Aristotle's example of the physician who treats Callias.

Handbook of the Philosophy of Science: General Philosophy of Science - Focal Issues, 377–430.
Volume editor: Theo Kuipers. Handbook editors: Dov M. Gabbay, Paul Thagard and John Woods.

Medicine for Aristotle is a practical science, but its practice is enhanced by a grasp of theory. The better physician will not be one who knows only how to treat Callias, or men of a certain age, or those with the specific ailment afflicting Callias. Rather, the better physician will be one who understands disease *qua* disease, according to its principles and causes, and understands people *qua* people.

This consequence of a science's rank in the hierarchy applies even within the theoretical sciences. It is by virtue of doing metaphysics, the highest theoretical science, that one truly grasps the lesser two theoretical sciences. That is, the better physicist or mathematician is one who understands metaphysics. As Aristotle makes clear in the middle books of the *Metaphysics*, he thinks there are causes and substances that are beyond the reach of physics. For him, physics is the science of sensible substances and their causes, but there is a more fundamental substance (*ousia*) as well as a more fundamental source of motion. The study of this substance and of the first motion inform physics rather than the other way around. Mathematics has the same relationship with metaphysics as physics: the study of surfaces and quantity depends upon and is informed by the more universal questions of metaphysics (*Metaphysics M* and *N*): Are there mathematical objects? Do numbers exist? Are numbers causes? Are they substances? Aristotle does not take up these questions by asking what we know about mathematics, but rather by asking what we know universally.

1.2 French Encyclopedists

When we think of comprehensive accounts of knowledge today, we often think of encyclopedias. These modern works have their origin in the period after the scientific revolution, when the integration of knowledge achieved by Aristotle and maintained by the Scholastics was fundamentally undercut. Historically the most famous encyclopedia was *Encyclopédie, ou dictionnaire raisonné des sciences, des arts et des métiers* (Encyclopedia, or Reasoned Dictionary of the Sciences, Arts, and Trades). Its 17 volumes (plus 11 volumes of illustrations) were produced over the period 1751–1772 under the editorship of Denis Diderot along with the mathematician Jean Le Rond d'Alembert. The project had its origins in a French translation undertaken by John Mills in 1743-1745 of Ephraim Chambers's *Cyclopaedia, or Universal Dictionary of Arts and Sciences.* The French publisher wrested control from Mills and, intending speedy publication, engaged two editors in succession who instead expanded the project's contours. The second, Diderot, undertook a monumental effort to outline the present state of knowledge in the sciences, arts, and practical crafts and to make this knowledge widely accessible. Originally each topic was to be covered by a scholar or craftsperson expert in it, and contributors included such prominent Enlightenment figures as Voltaire and Rousseau. In the end, though, Diderot and d'Alembert wrote many of the 71,818 entries themselves.

Although clearly embracing a philosophical perspective, the *Encyclopédie* served more to bring together different domains of knowledge than to unify or even systematize them. To the extent that there was a unifying theme, it lay in the Enlightenment's reliance on reason and empirical observation to provide knowledge. Even religion was presented as an object of human reason, not as a source of knowledge via revelation. The *Encyclopédie* thus stood in opposition to the scholastic tradition, which maintained Aristotle's legacy but subordinated it to Christian theology. The entry on philosophy emphasizes the role of reason:

> Reason is to the *philosopher* what grace is to the Christian. ... Other men are carried away by their passions, without their actions being preceded by reflection: these are men who walk in the shadows; whereas the *philosopher*, even in his passions, acts only after reflection; he walks in the night, but he is preceded by a torch. The *philosopher* forms his principles on the basis of an infinite number of discrete observations. ... He certainly does not confuse it with probability; he takes as true that which is true, as false that which is false, as doubtful that which is doubtful, and as probable that which is only probable. He goes further — & here is a great perfection of the *philosopher* — when he has no proper motive for judging, he remains undecided. (Translation by Dena Goodman from *The Encyclopedia of Diderot and d'Alembert Collaborative Translation Project*, http://www.hti.umich.edu/d/did/)

Not surprisingly, this emphasis on reason and empirical knowledge and criticism of claims for revealed truth ran afoul of the Church, so after the first seven volumes were published in Paris under a royal privilege, the remainder were published under the false imprint of Samuel Faulche, Neuchâtel (in fact they were published in Paris).

Reflecting the great diversity of human pursuits that involve acquisition of knowledge, the *Encyclopédie* represents a compilation of knowledge rather than an integration of it. In many respects, this reflects our contemporary situation. But in the wake of the enlightenment, other theorists resumed the pursuit of systematic unity.

1.3 Oken's Systematic Unity

Lorenz Oken (1779–1851) was an anatomist and a leader of the *Naturphilosophie* movement in Germany. A student and follower of Friedrich Schelling, Oken applied the precepts of *Naturphilosophie* to his thinking about biological systematics. The metaphysics he learned from Schelling — a Pantheistic view by which everything in nature could be deduced from a first principle, namely God — led him not only to treat the biological world as a part of God, but also to articulate a hierarchical classification of everything [Oken, 1809; 1831; Ghiselin and Breidbach, 2002]. Oken treated the organization of the world as a divine code that could be read by understanding the systematic relations between each thing and everything else.

Oken's approach to systematics was essentially that of the *scala natura*. His *Lehrbach der Naturphilosophie* offers an account in which his philosophical, theological, numerological, and biological assumptions were all tied together to produce a single, unified 'anatomy' of the world. There was first an argument that God is nothing, since all comes from nothing. This is just to say, of course, that God is (the source of) everything [Ghiselin, 2004]. After this theological argument was some numerological reasoning relating the four basic elements of the world (fire, air, water, and earth) to processes like electricity and crystallization. The book culminated in an argument that war-making is the highest art.

The thoroughgoing unity of Oken's classification is well illustrated by his theory of color.[1] For numerological, theological, alchemical, and scientific reasons, red corresponds to fire, then to love, and then to God the Father. Blue, as we might expect, corresponds to air, then to faith, and then to God the Holy Spirit. Yellow corresponds to earth, vice, and Satan. The colors of natural entities fit into, and are regarded as explained by, this overarching system. For example, animals are predominantly red because they correspond to fire (and the cosmos). Plants have green leaves because they correspond to water (and the planets). Flowers get a three-way classification: those of lower plants are most often yellow, the intermediate ones blue, and the highest ones red.

1.4 Encyclopedia of Unified Science

Whereas Oken attempted to build unity in terms of conceptual (semantic) ideas, other approaches to systematizing knowledge appealed to logic (syntax) for the bridges between bodies of knowledge. Logical positivism, later known as logical empiricism, developed in the 1920s in Austria (Verein Ernst Mach in Wien, commonly known as *the Vienna Circle*), Germany (Gesellschaft für Wissenschaftliche Philosophie Berlin, commonly known as *the Berlin Circle*), and Poland. The term and basic doctrine of *positivism* originated with August Comte, an early 19^{th} century French philosopher who was skeptical of philosophical systems and metaphysics generally and emphasized positive knowledge — that is, knowledge grounded on observation and experimentation. A more immediate influence was the positivism of Ernst Mach, a professor of physics in Prague and Vienna until his retirement in 1901. He adopted a radical empiricism in which the only source of knowledge was sensory experience, and scientific laws were instrumental, serving to describe and predict phenomena available to the senses. Most of the early logical positivists adopted Mach's emphasis on the experiential grounding of knowledge, although most did not share his extreme instrumentalism. The adjective *logical* identifies the chief resource to which the logical positivists appealed in advancing beyond individual observations to generalized scientific claims. The logic to which they appealed was not traditional Aristotelian logic, but rather the modern mathematical logic developed in the late 19^{th} and early 20^{th} centuries by Frege, Peano, Russell, Whitehead, and others. Many of the logical positivists were themselves scientists who were concerned about clarifying the foundations

[1]This example is due to Michael Ghiselin, and is spelled out in more detail in [Ghiselin, 2004].

of science, especially in light of major developments in physics and other sciences contemporaneous with the rise of mathematical logic.

Although many of the logical positivists focused on physics, their emphasis was on providing a general account of knowledge, which they equated with scientific knowledge. They also, as discussed in more detail below, articulated a vision of how different sciences could be unified into a theoretical whole through theory reduction. One motivation was to counter a view, widespread at the time, that psychology addressed an inner world that was discontinuous with the outer world studied by the other sciences. Initially Rudolf Carnap [1928] proposed to overcome this discontinuity by treating all science as grounded on private experience, from which the world was *constructed*. This project, however, was unsuccessful. An alternative proposal for unification was offered by Moritz Schlick, who distinguished the content of experience (specific sensations) from its structure (relations between experiences). He maintained that the structure of experience was objective and could be investigated empirically. These and other attempts to provide a common account of the methodology of all sciences and link them into a common theoretical edifice gave rise to the *International Encyclopedia of Unified Science*, edited jointly by Otto Neurath, Rudolf Carnap, and Charles Morris.[2]

Neurath envisaged that the encyclopedia would grow to hundreds of volumes, with one entry issued each month in a subscription series. In the end only 20 entries were published in two volumes, the first under the original title and the second under the more modest title *Foundations of the unity of science; toward an international encyclopedia of unified science*. The goal, according to Neurath [1938, 24], was "to integrate the scientific disciplines, so to unify them, so to dovetail them together, that advances in one will bring about advances in the others." The main tool for such dovetailing of different sciences was logical analysis, which would serve to relate the concepts and ultimately the theoretical claims of various sciences. Although the editors envisioned an axiomatized integration of the great body of knowledge provided by the various sciences, they adopted a piecemeal strategy. They fully expected this procedure would uncover inconsistencies whose eventual resolution would improve each science as well as the prospects for their integration.

Since the account of unity advanced by the logical positivists has been the chief focus of philosophical accounts of the unity of science ever since, we will return in greater detail to this account in part 2 of this chapter. First, however, we consider one last proposal for unity which, although receiving less attention in philosophy, had and continues to have considerable influence in the sciences themselves.

[2]The term *unified science* was first invoked in 1938 when *Erkenntnis*, which had been the house organ of the Vienna Circle since 1930, was moved to the Hague and renamed the *Journal of Unified Science*. Just two years later, however, it ceased publication.

1.5 Cybernetics and General Systems Theory

Beginning in the 1940s, cybernetics and general systems theory advanced a very different conception of how to unify science that focused primarily on the organization found in phenomena the sciences seek to explain, especially the biological and social sciences. The term *cybernetics* was coined by mathematician Norbert Wiener from the Greek word for 'helmsperson', and was applied to systems that could steer themselves [Wiener, 1948]. Working during World War II, Wiener initially focused on a practical problem: developing a system for improving the accuracy of anti-aircraft guns. His desired solution invoked feedback control; that is, the accuracy of previous shots would be used to adjust gun controls before taking the next shot. Challenges he faced in getting the idea to work led him to collaborate with an engineer, Julian Bigelow, and a physiologist, Arturo Rosenblueth. In a paper in *Philosophy of Science* [Rosenblueth *et al.*, 1943], the three developed the idea that feedback enabled both biological and artificial systems to be goal-directed. They regarded this as resuscitating a notion that was anathema to the positivists: that of *teleology*. Subsequently Wiener organized a multi-year conference series. He initially called it Conference on Circular Causal and Feedback Mechanisms in Biological and Social Systems but, beginning in 1949, the conference adopted Wiener's term *cybernetics* for its name. As the initial name suggests, the participants regarded the idea of feedback organization as having the potential to unify biological and social systems.

Around the same time, biologist Ludwig von Bertalanffy [1951] advanced General Systems Theory as an antireductionist yet unifying perspective. Rather than focusing on the particular components out of which different things were made, systems theory emphasized the organization of parts into wholes and maintained that the same principles of organization, such as negative feedback, would be found applicable in physics, chemistry, biology, the social sciences, and technology.

Although there is an International Society for Systems Sciences that is still active and runs large international meetings, cybernetics and general systems theory have declined into niche specializations. Today the strongest influence of these approaches is indirect, funneled through successors with new ways to identify general principles of organization and use them towards unifying science. The new work goes under such rubrics as the sciences of complexity, complexity theory, and self-organizing systems and emphasizes systems with non-linear interactions. Tools for describing such systems were first developed in physics by Poincaré and others in the late 19^{th} century, giving rise to Dynamical Systems Theory (DST) in the 20^{th} century. DST was initially applied to physical phenomena such as eddies in a stream [Landau, 1944], but also was used to elucidate phenomena in biology (see [Kaufmann, 1993]) and then psychology. The earliest psychological accounts focused on motor coordination [Kelso, 1995] and its development [Thelen and Smith, 1994], but gradually DST has expanded to other domains. Indeed, some proponents have presented DST as a revolutionary, overarching alternative to other approaches to cognition [Port and van Gelder, 1995; Keijzer, 2001].

One particularly interesting offshoot of complex systems research has been the introduction of a number of important ideas about the structure of networks and how they can be used to characterize phenomena in the world. Most traditional investigations of networks focused either on regular lattices, in which only neighboring units are connected, or on random networks (the focus of pioneering investigations by Erdös and Rényi). Such organization is very different from the "small world" networks first articulated by Stanley Milgram [1967], who discovered empirically that while individual humans are primarily connected to those around them (as in regular lattices; this feature is known as *high clustering*), they are indirectly connected to a vast number of others via relatively short paths of direct connections through people who know each other (as in random networks; this is known as *short path length* and provided the premise for the play and movie *Six Degrees of Separation*). Duncan Watts and Steven Strogatz [1998] showed that minimal changes to a regular lattice can transform it into a small-world network and explored real-world phenomena exhibiting this form of organization — including collaborations between actors in feature films, the electrical power grid of the western U.S., and the neural network in a nematode. Moreover, Albert-Lászlo Barabási and Réka Albert [1999] discovered that many networks in the real world are *scale free*, in that connections exhibit a power-law distribution (the majority of units are connected to only a few others, but a few are connected to a *very* large number of others). (The term *scale free* is used to reflect the fact that power-law distributions lack any intrinsic scale.) Barabási and his collaborators have attempted to account for the occurrence of scale-free networks as a result of historically earlier nodes having a longer time to attract attachments and to new nodes preferentially attaching to already highly connected nodes [Albert and Barabási, 2002]. More recently Cees van Leeuwen and his collaborators have shown how scale-free small-world networks can evolve through coupling of chaotic oscillators [Gong and van Leeuwen, 2003]. These developments potentially provide a powerful set of tools for analyzing organization in a wide variety of natural and social systems.

2 FIELD GUIDE TO MODERN CONCEPTS OF REDUCTION AND UNITY

In the 20^{th} century, claims about unity of science were commonly tied to claims about theory reduction. In particular, the strategy was to reduce the theories of higher-level sciences such as biology to the laws and theories of lower-level sciences such as physics and chemistry. (Spelling out the notion of levels is challenging and we will return to this issue at several junctures below.) Claims about reduction were, in turn, treated as claims about deductive relations between theories. Recently, strong dissent has been raised on both scores, with some philosophers rejecting both reduction (see below, 2.3) and unity of science (see below, Section 4). Other philosophers, more sanguine about unity, have advanced alternative conceptions that emphasize integration more than unity and detach these issues from questions of theory reduction. In addition, accounts of reduction that do not tie it to deductive relations between theories have been advanced. Although the more recent alternative treatments of both integration and reduction offer much promise for providing more adequate accounts of both of these notions, we will start by laying out the traditional accounts of both positions.

2.1 The Theory-Reduction Model

The logical positivists advanced the theory-reduction model as part of their effort to provide an account of science that avoided entanglement with metaphysical issues. To accomplish this they focused on the knowledge claims of science and emphasized the role of logical relations between these. A crucial move was to represent two kinds of knowledge claims in the same format, yielding sets of propositions encompassing both observation statements (reports of empirical observations such as "The marble is rolling down the incline") and theoretical statements like Newton's law of universal gravitation, which says that the attractive force between any two bodies is equal to the product of their masses divided by the square of the distance between them. Nagel identified an intermediate category of *experimental laws*, which provide an empirical summary of the phenomena observed. Galileo's law that the distance a falling object travels is proportional to the square of the time it is in motion is an example. These experimental laws are contrasted with *theoretical laws*, such as Newton's, which go beyond the observed phenomena by positing theoretical entities like forces and masses to account for the experimental laws. The power of laws or theories to explain observations could then be rooted in the ability to derive new observation statements — predictions — from laws. This is the well-known deductive-nomological (D-N) or covering-law model of explanation [Hempel and Oppenheim, 1948; Hempel, 1965]. To account for the relations between the laws or theories of different sciences, the logical empiricists proposed simply generalizing this account, and argued that it should be possible to derive the laws or theories of one discipline or science from those of another [Nagel, 1961; see also Woodger, 1952; Quine, 1964; Kuipers, 2001, chapter 3].[3]

Two fundamental challenges arose in developing this generalization of the D-N model. First, the laws in the different sciences are typically presented in different

[3]Kemeny and Oppenheim [1956], see also [Oppenheim and Putnam, 1958], advanced an alternative account of reduction that did not derive the reduced theory from the reducing theory but only required generating identical observable predictions from the reducing theory as the reduced theory. This account of reduction is far more liberal, since it allows for the reduction of what are regarded as false theories (e.g., phlogiston chemistry) from what are taken to be true theories (e.g., Lavoisier's oxygen-based chemistry) as long as the predictions made by the reducing theory include all those made by the reduced theory. Yet another alternative was put forward by Patrick Suppes, who required an isomorphism between any model (in the model-theoretic sense) of one theory and a model of the reduced theory: "To show in a sharp sense that thermodynamics may be reduced to statistical mechanics, we would need to axiomatize both disciplines by defining appropriate set theoretical predicates, and then show that given any model T of thermodynamics we may find a model of statistical mechanics on the basis of which we may construct a model isomorphic to T" [Suppes, 1957, 271].

vocabularies.[4] Laws in physics, for example, might employ terms such as *mass* and *attractive force*, whereas those in chemistry would invoke names of elements and molecules and types of chemical bonds. But logical inferences are only possible between statements using the same vocabulary, in much the same way as certain algebraic problems can be solved only when the units of time, length, or weight are expressed using the same measure. To address this issue, advocates of the theory reduction model appealed to *bridge principles* (Nagel called them *rules of correspondence*) that equated vocabulary in the two laws. Sklar emphasized that these correspondence claims are really identity claims: "Light waves are not correlated with electromagnetic waves, for they *are* electromagnetic waves" [Sklar, 1967, 120]. Applied to the context of relating psychology to neuroscience, this contention that the terms in the different theories picked out the same entity became the foundation of the celebrated mind-brain identity theory [Place, 1956; Feigl, 1958/1967; Smart, 1959]. Although such bridge principles might seem unproblematic, we will see that they are the target of one of the more powerful objections to unification through reduction.

The second challenge confronting advocates of the theory reduction model is the fact that the regularities captured in higher-level laws arise only under a particular range of conditions. To accommodate this, they proposed that reduction also required statements of boundary conditions. With these components in place, a reduction was then conceived to have the form of the following deduction:

Lower-level laws (in the basic, reducing science)
Bridge principles
Boundary conditions

∴ Higher-level laws (in the secondary, reduced science)

An oft-cited example is the derivation of the Boyle-Charles' law from the kinetic theory of gases, as part of an overall reduction of classical thermodynamics to the newer and more basic science of statistical mechanics [Nagel, 1961, 338–366]. This law states that the temperature (T) of an ideal gas in a container is proportional to the pressure (P) of the gas and volume (V) of the container. Because the term *temperature* does not appear in statistical mechanics, to achieve the reduction a linkage to a term in that science is required. This is expressed in a bridge principle (rule of correspondence) stating that the temperature of a gas is proportional to the mean kinetic energy (E) of its molecules. A number of boundary conditions also must be specified, such as those limiting the deduction to monotonic gases in a temperature range far from liquefaction. With the appropriate bridge principles and boundary conditions included as premises, the Boyle-Charles' law can be derived from laws of statistical mechanics. Here is a key part of the full derivation:

[4]Nagel did consider cases in which the same vocabulary was employed in the reducing and reduced theories. He referred to such reductions as *homogeneous*.

Laws of statistical mechanics (including the theorem $PV = 2E/3$)
Bridge principles ($2E/3 = kT$)
Boundary conditions (monotonic gas; T in specified range)

$\therefore$ Boyle-Charles' law ($PV = kT$)

Notice that behind the unity-as-reduction conception is a view of the natural world as comprised of levels, often referred to as 'levels of organization'. Given their reluctance to engage ontological issues, the logical empiricists tended to construe levels in terms of the disciplines that investigate them. On this view there are levels of organization in the world that correspond to such disciplines as physics, chemistry, biology, psychology, and sociology. Unification consists of reducing the theories of each higher discipline to theories of a lower discipline. Some philosophers regard this as thereby achieving a reduction between the disciplines themselves. So, for instance, if biological theory were reduced to physical and chemical theory, the science of biology would also thereby be reduced to the sciences of physics and chemistry, and biology would no longer be an autonomous science.

While embracing many features of the logical empiricists' account of reduction, Robert Causey [1977] advanced a more ontologically committed interpretation of levels wherein higher levels resulted from the structuring of lower-level entities. On this view, theories at the lower level primarily describe the operation of parts of the structured wholes, while those at the higher level focus on the behavior of the structured wholes themselves. For a reduction to be possible, the lower-level theory must itself have the resources to describe the structured wholes and their behavior. (Although this is quite problematic, assume for a moment that it is possible.) We then have two descriptions of the higher-level entity, one as a whole unit in the vocabulary of the higher-level science, and one as an entity structured out of lower-level components. For Causey, reduction then requires bridge principles that relate terms in the higher-level theory referring to the wholes to those terms in the lower-level theory that characterize them as composed, structured wholes. Assuming that the lower-level theory has laws that describe the behavior of the structured wholes, one can try to derive the upper-level theory from the lower-level one.

An important feature of the theory-reduction model is that it requires the lower-level theory (or science) to have all the resources required to derive the upper-level theory (when bridge principles and boundary conditions are supplied). Below we will consider whether this is plausible. But noting this feature of the model allows us to consider what many practitioners of higher-level sciences find problematic in philosophical accounts of reduction: successful reduction apparently obviates any need for any laws or theories specific to the higher-level sciences. At least in a hypothetical final picture of science, higher-level sciences would be expendable or redundant: by supplying the appropriate boundary conditions, any higher level regularity could be derived directly from the lower-level theory. In practice, at a given stage in the development of science, appeals to the higher-level sciences may be required because the reduction base may not yet have been developed. Higher-level sciences may even play a heuristic role in the development of the lower-level sciences; for example, they may reveal regularities (laws) in the behavior of the structured wholes that must be accounted for. In this respect, there may even be a co-evolution of higher- and lower-level sciences [Churchland, 1986]. In the end, however, the theories of the lower-level science will be complete, and the only reason for invoking the vocabulary and laws of the higher-level science will be that they provide a convenient shorthand for referring to what, in the lower-level theory, may be unmanageably complex statements.

2.2 Revisionist Accounts of Theory Reduction

Among the early challenges to the theory-reduction model, one of the most influential focused on the possibility of establishing the appropriate bridge principles. Paul Feyerabend [1962; 1970], as a result of adopting an account that characterized the meaning of scientific vocabulary in terms of the theory in which they were used, argued that words in different theories, even if they have the same form, are *incommensurable* with one another. In classical thermodynamics, for example, *temperature* can be defined in terms of Carnot cycles and its behavior described by the non-statistical version of the second law of thermodynamics. But in statistical thermodynamics, temperature is characterized in statistical terms. Given the important differences in the surrounding theory and hence the different entailments of the meanings of 'temperature', it would seem impossible to construct bridge principles that would adequately relate 'temperature' as used in these two theories. At the same time, Thomas Kuhn [1962/1970] focused on other examples of putative reduction, such as Newtonian to Einsteinian mechanics, and maintained that words like 'mass', were used incommensurably in the two theories. On the basis of such examples, Kuhn challenged the account of progress implicitly assumed by the logical empiricists, in which sciences progress towards better theories through a process of continual extension and refinement. Kuhn argued instead that the history of science is a history of revolutions in which new theories replace, rather than build upon older, incommensurable theories.

One specific context in which Feyerabend maintained that reduction would fail involved attempts to relate psychological theories presented in mentalistic vocabulary to accounts of brain function in neuroscience. Although Feyerabend later came to champion the position that incompatible theories ought both to be maintained [Feyerabend, 1975], in his early writing on mind-brain relations he advanced a position known as *eliminative materialism* [Feyerabend, 1963]. The key claim of Feyerabend and subsequent eliminativists [Rorty, 1970; Churchland, 1981; Churchland, 1986] is that instead of reducing the old (folk) psychological theory to the new neuroscientific theory, the old psychological theory is *replaced* by the new theory and *eliminated* from the corpus of science. The model of such replacement is the replacement of Ptolemy's astronomical theory by Copernicus'. The old, Ptolemaic, theory accounted for the observed motions of the planets by assuming that they moved on epicycles whose centers themselves orbit around the earth (the epicycles explained the apparent retrograde motion of the planets when viewed from earth). Copernican astronomy, as in Figure 1b, explains the same phenomenon by assuming that the earth and other planets orbit the sun. Since the two astronomical accounts are inconsistent and we assume that the Copernican account is fundamentally correct, eliminativists conclude that Ptolemy's account is wrong and should be discarded and replaced. Such replacement befell not only historical theories such as Ptolemaic astronomy, the impetus theory of motion, and phlogiston chemistry but also, on this view, awaits folk psychology and other mentalistic accounts.[5]

Although Feyerabend and Kuhn viewed themselves as opposing reduction, and eliminativists such as the Churchlands held that elimination awaits when reduction fails, other philosophers such as Kenneth Schaffner treated Feyerabend and Kuhn as advancing an alternative account of reduction. On this alternative, even when deduction fails (as it must when the reducing theory is true and the reduced theory is false), one can still relate the old theory to the new one. In the late 16^{th} century, for example, Tyco Brahe developed a way to map the Copernican model of the solar system onto the Ptolemaic one (Figure 1c). This showed that all the empirical observations that had supported Ptolemy's model also fit Copernicus's and offered support to it. Thus, one reason for exploring such relations between an old theory and its replacement is to enable the replacing theory to claim much of the empirical support that had been developed for the old theory.

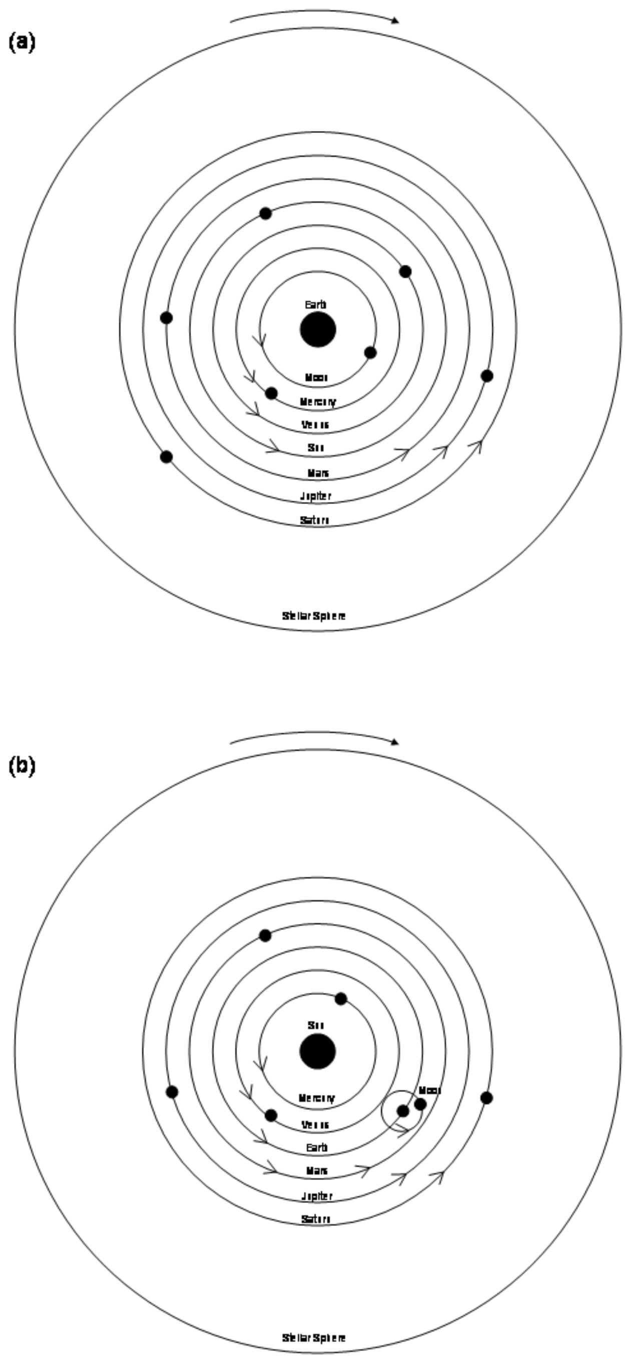
After construing the discovery of such similarities as a kind of reduction that differs from the traditional model in interesting ways, Schaffner [1967] suggested that these two kinds of reduction need not be regarded as competitors. Instead, he proposed a comprehensive account in which reduction by deduction and reduction by replacement each play a role. In particular, a frequent consequence of a new lower-level theory (T_1) is that an old upper-level theory (T_2) gives way to a revised one (T_2*). T_2* should be deducible from T_1, just as envisaged in the standard theory-reduction model, but Schaffner thought its relation to T_2 should also be recognized. He suggested that the T_2-T_2* relation was one of analogy:

> T_2* corrects T_2 in the sense of providing more accurate experimentally verifiable predictions than T_2 in almost all cases (identical results cannot be ruled out however), and should also indicate why T_2 was incorrect (e.g., crucial variable ignored), and why it worked as well as it did. ... The relations between T_2 and T_2* should be one of strong analogy — that is (in current jargon) they possess a large "positive analogy". [p. 144]

Subsequently Schaffner [1969] amended his model to incorporate revision of an existing lower-level theory (T_1) to obtain a corrected lower-level theory (T_1*) in addition to the revision of old higher-level T_2 into T_2* (see Figure 2).

Schaffner provided little guidance as to what counted as a strong analogy. Thomas Nickles [1973] argued that in many instances these analogies could be understood mathematically as limit relations. At specific limit values for variables in the new theory, he argued, the new theory will yield the older theory, nearly enough. Nickles give the example of Einstein's formula for momentum reducing to the Newtonian formula by taking the limit as velocity approaches zero. Such

[5]Although most commonly the Churchlands have targeted folk psychology for their eliminativist claims, they also on occasion target contemporary cognitive psychology: "There is a tendency to assume that the capacities at the cognitive level are well defined ... As we see in the case of memory and learning, however, the categorial definition is far from optimal, and remembering stands to go the way of impetus" [Churchland, 1986, 373].


(a)
Earth
Moon
Mercury
Venus
Sun
Mars
Jupiter
Saturn
Stellar Sphere
(b)
Sun
Mercury
Moon
Venus
Earth
Mars
Jupiter
Saturn
Stellar Sphere

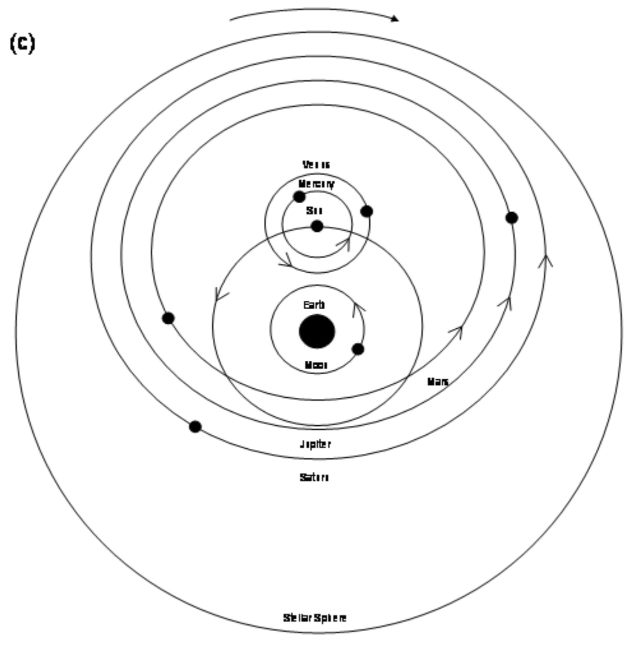


Figure 1. The orbits of the planets according to (a) Ptolemy, (b) Copernicus, and (c) Tycho Brahe. In all cases the planets orbit counter-clockwise while the stellar sphere moves clockwise. In Brahe's account, the earth is at the center of the solar system, as it was for Ptolemy, but the planets other than the moon revolve around the sun, as in Copernicus's account.

limit relations enable researchers to appreciate why the older theory worked as well as it did — most velocities Newtonian scientists considered were sufficiently small that the actual momentum differed only minutely from that predicted by the Newtonian formula.

The strategy of using a limit relation to capture the analogy between a revised theory and its predecessor will not work in all cases, however. William Wimsatt [1976a] argued that Schaffner's conception of strong analogy should be understood in terms of pattern-matching, in which a limit relation is just one of a number of ways to construct a match. Moreover, he extended Nickles' account of the function of such matches by focusing on the differences remaining after the match. These differences not only mark points at which evidence may show the revised theory to be an improvement; they may also, in cases where the predictions from the new theory are not as successful as those from the old theory, point to loci where yet further work is needed to amplify and extend the new theory.

Nickles further argued, convincingly, that advocates of the traditional theory re-

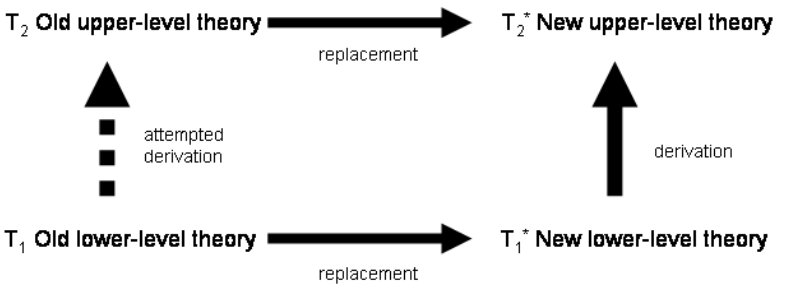


Figure 2. Schaffner's (1969) model of reduction, in which a new upper-level theory (T_2^*) is derived from a new lower-level theory (T_1^*) and each new theory replaces an older theory at the same level.

duction model were talking about a very different relation than were such critics as Kuhn and Feyerabend. He labeled reduction as envisaged by the theory-reduction model *reduction*$_1$ and argued that it is particularly relevant in explaining domain-combining types of reduction (which Wimsatt [1976a] characterized as interlevel reductions). But the relation between predecessor and successor theories is a domain-preserving relation which he labels *reduction*$_2$ (and Wimsatt construed as intralevel reduction). One feature that Nickles identified as distinguishing the two types of reduction is that they tend to be invoked for different reasons: reductions$_2$ serve heuristic and justificatory roles, while reductions$_1$ are unifying and explanatory. He also noted that the two reductions point in opposite directions with respect to theories differing in their generality. In reduction$_1$ the more specific upper-level theory is reduced to the more general lower-level one (e.g., the reduction of gas laws to the more general theory of statistical mechanics). In reduction$_2$ the more general theory is a newer one that reduces to the older theory, now recognized to be incorrect (e.g., the reduction of Einstein's formula for momentum to Newton's). In sum, in reduction$_1$ the move is from specific to general, whereas in reduction$_2$ it is from general to specific. (See Figure 3)

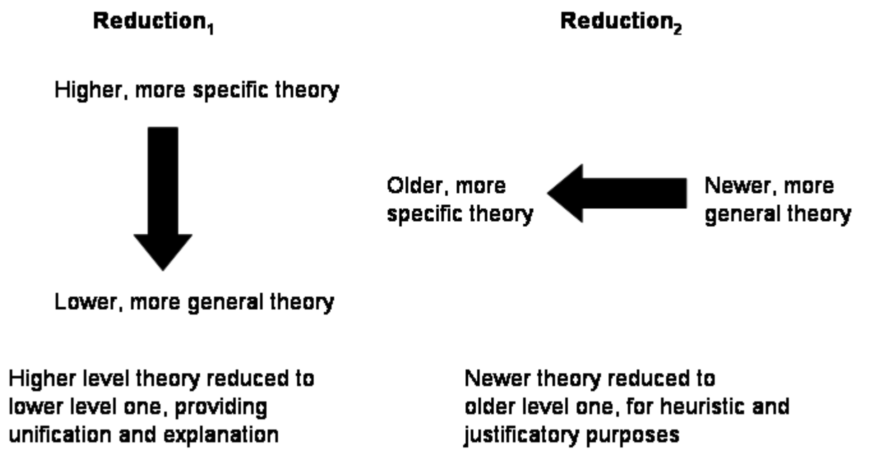


Figure 3. Nickles' two senses of reduction. In reduction$_1$ a higher-level theory is reduced to a lower-level one, whereas in reduction$_2$ a new, more general theory is reduced (e.g., in the limit) to an older, more specific theory.

In his development of Nickles' position, Wimsatt offered a novel reading of when new theories eliminate older ones. In cases for which there is a close pattern match between the old theory and the new one, the older theory might well continue to be employed because it is simpler or easier to use. But reductions between successively introduced theories are unlikely to be transitive. Rather, "*intralevel reductions should be intransitive* — ... a number of intralevel *reductions* could 'add up' to an intralevel *replacement*. ... Relativistic [Einsteinian] mechanics may reduce to classical mechanics (etc.) but it clearly replaces (rather than reduces to) Aristotelian physics" [Wimsatt, 1976a, 217–219].

Wimsatt's distinction between interlevel and intralevel reductions reveals interesting consequences for the eliminativist argument as applied to the relation between neuroscience and psychology. Whereas "eliminative materialism seems ... to derive its inspiration from intralevel reduction," Wimsatt contended, "the proper model for the mind-body problem is interlevel reduction" [Wimsatt, 1976a, 215]. This critique was further developed by McCauley [1986; 1996], who showed that historical cases exemplifying the replacement and elimination of an old theory have all involved a revised theory that is at the same level as the old theory.

McCauley suggested that the same would be true in the case of a psychological theory: elimination would be expected only when it was superseded by a replacement theory that lay at the same level — i.e., another psychological theory rather than a neural one). As for interlevel reductions, McCauley distinguished cases in which there is a tight fit between upper- and lower-level theories and cases in which there is not. Loose fit may result from the very nature of theorizing at the upper and lower level. In some cases, the finer grain of an account at the lower level may enable it to explain what appear to be deviations at the higher level. But the advantage is not always with the lower level. In other cases,

> the upper-level theory lays out regularities about a subset of the phenomena that the lower-level theory encompasses but for which it has neither the resources nor the motivation to highlight. That is the price of the lower-level theory's generality and finer grain. [McCauley, 1996, 31]

McCauley thus advocates a pluralistic approach that would allow theorists a fair degree of autonomy. Theories at higher and lower levels could be developed independently, with no immediate need to force the levels to relate in a reductionistic manner.

2.3 Criticism of Theory Reduction

Revisionists presented the difficulty of providing bridge principles as arising principally with cases involving successive theories at the same level (Wimsatt, as we will see below, is an exception), leading them to invoke a different account of intralevel and interlevel relations. Some influential critics, however, see the problem as arising even in the interlevel case and as providing the death-knell for the theory reduction account of interlevel relations. Similar arguments were advanced independently by two such critics, David Hull regarding biology and Jerry Fodor regarding psychology. The strategy in both cases was to maintain that the same term used in the laws of the higher-level theory must be related on different occasions with different terms and fall under different laws at the lower level. As it is sometimes expressed, one type of entity as characterized in the higher-level theory is *realized* by multiple different types of lower level entities on different occasions.

Hull [1972; 1974] focuses on the notion of *gene* as it figures in both Mendelian genetics and molecular genetics. One challenge to providing a reductive account in this case is that genes in Mendelian accounts are characterized in terms of phenotypic traits for which they code (e.g., a pea plant is tall, not short). Genes in molecular genetics are characterized in terms of their molecular constitution. Any one of a number of distinct molecular mechanisms could produce the same phenotypic trait (this is often referred to as *multiple realizability*, to which we will return below). Although the complicated nature of the phenotype-genotype map makes developing the reduction difficult, it does not necessarily block it. To achieve a reduction, what is needed is "to discover one or more molecular mechanisms which correspond to the various predicate terms of Mendelian genetics, such that the resulting classification of traits into types corresponds fairly well with the classification of these traits according to the principles of Mendelian genetics" [Hull, 1972, 497].[6] Hull went on to point out that even this modest goal cannot be reached; instead, scientists have found that "the same molecular mechanism can produce different phenotypic effects". This is just the reverse of multiple realizability, as it involves multiple different effects produced by the same mechanism. The reason is not a mystery: other conditions vary. Which conditions and combinations of conditions produce different phenotypic effects can be determined empirically by researchers, if desired. However, to bring such detailed findings into molecular genetics, so an adequate reduction of Mendelian genetics could be accomplished, would result is a radical expansion in scope: "We are no longer correlating Mendelian predicate terms with molecular mechanisms but with the entire molecular milieu" (p. 498). One possible conclusion is that reduction fails in the case of Mendelian genetics, but Hull pointed the blame instead at the account of reduction offered by philosophers:

> If the logical empiricist analysis of reduction is correct, then Mendelian genetics cannot be reduced to molecular genetics. The long-awaited reduction of a biological theory to physics and chemistry turns out not to be a case of "reduction" after all but an example of replacement. But given our pre-analytic intuitions about reduction, it *is* a case of reduction, a paradigm case. [Hull, 1974, 44]

If a paradigm case of reduction fails to go through on the theory-reduction model, Hull reasoned, the philosophical framework would seem to have failed. However, some philosophers of biology drew a different conclusion from the difficulties identified by Hull: they treat the failure as pointing to fundamental deficiencies in biology. In particular, Alexander Rosenberg [1994] argued that because natural selection selects for function rather than structure, the relations between Mendelian genetics (phenotypic features characterized functionally) and molecular genetics (genotypes characterized structurally) are so complex that any attempt to construct bridge laws between them will yield disjunctions too long to be useful for creatures of our mental capacity. Focusing just on the multiple realizability of traits, not the reverse relation, Rosenberg observes that given an environmental 'problem' to solve, selection can achieve the same phenotypic function by any number of molecular pathways. The phenotypic or functional features 'tallness' and 'roundness' are, in other words, multiply realizable from the point of view of molecular genetics. Offering bridge laws, then, will amount to making a list of all the possible pathways. This process, Rosenberg argues, leads to intractably

[6] As Richardson [1979] noted, Nagel actually allowed for such multiple realizations of the same higher-level property as long as it was possible to explain why the different lower-level properties realized the same higher-level one. Differences in context may determine whether a particular lower-level property realizes a higher-level one.

long lists rather than a better understanding (which is what a true science would provide) of the molecular underpinnings of Mendelian genetics or of the operation of natural selection. Since the theories of functional biology are not reducible to molecular foundations, they provide only problematic access to the biological world.

The other critic of theory reduction, Fodor [1974], focused on psychological predicates and argues that they cannot be linked via bridge principles to neuroscientific ones. Invoking an analogy with finance, he noted that money does not correspond to any natural kind of physical stuff. In the right circumstances, pieces of paper, gold, silver, bronze, or even patterns of electrons can each serve as money; hence, money is multiply realizable The example nicely draws out Fodor's primary point that the factors that determine kinds in behavioral and societal realms, such as finance, are very different from those determining kinds in the physical realm. In particular, Fodor, as well as Hilary Putnam [1978], maintained that psychological kinds should be identified functionally in terms of how they interact in the generation of behavior. For example, hunger will interact with cognitive states, such as beliefs, in generating particular food-seeking behaviors. Given the differences in their nervous systems, a functional state such as hunger will arise as a result of different neural processes in species such as octopi and humans, although in both cases the state will result in food-seeking behavior (this example is due to Putnam). Accordingly, both Fodor and Putnam reject the project of reducing psychology to neuroscience, instead advocating the autonomy of what Fodor refers to as the *special sciences*.[7]

A second response is advocated by Causey and Hooker. They recommend acknowledging multiple realizability and accepting that a different reduction will be needed for different lower-level realizations of a given higher-level law. Far from promoting unity, this response may actually result in greater disunity when phenomena that appear very similar in high-level terms turn out to be reduced to ver y different lower-level theories. Pylyshyn [1984], for example, argued that folk psychology successfully groups diverse behaviors under the same regularities, enabling us to predict behavior effectively, but that virtue would be lost if one tried to reduce it to the diverse behaviors that realized the regularity. For example, in our folk idiom we make generalizations about people's propensity to answer the phone,

[7] Fodor also maintains that in developing their taxonomies and relating states, special sciences will commonly appeal to very different principles than those that are typical in more basic sciences. For example, in seeking a psychological account of human decision making, we will prefer one that renders people and their decisions as rational; whereas this is not an objective in developing neuroscientific accounts. Charles Taylor [1967, 206] made essentially the same argument: "... if human behavior exhibits lawlike regularity, on the physiological level, of the sort which enables prediction and control, and a rougher regularity of a less all-embracing kind on the psychological level, it does not follow that we can discover one-one or even one-many correspondences between the terms which figure in the first regularities and those which figure in the second. For we can talk usefully about a given set of phenomena in concepts of different ranges, belonging to different modes of classification, between which there may be no exact correspondence, without denying that one range yields laws which are far richer in explanatory force than the others."

yet on different occasions that activity can involve different motor systems (e.g., picking up a handpiece and talking; sending a text message). By treating each instance separately, we lose the generality provided in the folk idiom "answering the phone".

Although many philosophers have assumed that multiple realizability is rampant and undermines the prospects of relating higher-level kinds to those of the more basic sciences, drawing such connections has been a key strategy in biological investigation. While recognizing that the mechanisms underlying physiological and psychological processes in different species do differ, investigators nonetheless draw extensively on what they have learned in one taxon to understand others. For example, much of what is now known about mechanisms of visual processing in humans was secured through research on other mammalian species, especially the cat and monkey [Bechtel, 2001]. Although neuroscientists fully realize that there are differences between brains of different organisms, especially of organisms from different taxa, they also expect and have found extensive commonality. This should not be surprising — it has long been known that biological mechanisms at all levels are often highly conserved, attributed in part to the high cost in fitness for large changes. Also — and this point has not received sufficient attention in the philosophical literature — biologists generalize from mechanisms, processes, and features identified in one taxon, to others by means of what might be called *phylogenetic reasoning*. Where such mechanisms, processes, and features can be shown to be carried through lineages, investigators expect fundamental similarities [Hennig, 1966]. Accordingly, the underlying mechanisms are not likely to be as radically different as advocates of multiple realizability assume. Researchers also expect differences between taxa and seek these out, but these will often be variations on a common structure: in the language of cladistics systematics, these similarities (and dissimilarities) will be both *shared* because of membership in a common lineage and *derived* due to the differential influences of evolution.[8] Given the conservative nature of evolution, we should not be surprised that human brains retain much of what is found in cat and monkey brains (and indeed, even the brains of invertebrates).

Those who view multiple realizability as an obstacle for reduction often neglect a further factor — just as there are neural differences between organisms and especially between species, there are psychological differences as well. The behavior of a hungry octopus is very different from that of a hungry human. Putnam ignores these differences when he applies the same psychological predicate to both. But these differences often matter as well in developing psychological theory. In both psychology and neuroscience, researchers can select a coarse-grained analysis, lumping together instances that differ in many respects, or a fine-grained analysis,

[8]There are, of course, examples of convergent evolution in which similar adaptations arise in different lineages (e.g., wings in bats, birds, and pterodactyls). But these are typically readily distinguishable functionally in a variety of ways (e.g., the amount of weight that can be supported or response to turbulence in the case of wings) and so typically do not provide good examples of the *same* function being multiply realized. For further criticisms of the assumption of multiple realizability, see [Bickle, 2003; Polger, 2004; Shapere, 2004].

splitting similar instances into different kinds. For different purposes, they may select one or the other. Putative examples of multiple realizability, however, often trade on invoking coarse-grained analyses of psychological kinds and fine-grained analyses of neural kinds. When the same grain is employed in lumping brains in the same category as is employed in lumping mental states into the same category, the alleged problems induced by multiple realizability for reduction seem to vanish [Bechtel and Mundale, 1999].

Leaving behind these worries about multiple realizability, a critical feature of theory reduction accounts, either in their original or revisionist versions, is the assumption that the lower-level theories have sufficient resources from which to derive all the laws of the higher-level science. This assumption is radically implausible. A first objection is that the lower-level theories to which higher-level ones could be successfully reduced would have to be rather different from those currently under development in the lower-level sciences. We can appreciate this by returning to Causey's version of the theory-reduction model. In his discussion, although not in his formal treatment, Causey suggests that researchers will study the behavior of the components of structured wholes when they are not part of the whole (his *non-bound condition*) and then derive their behavior when part of the structured whole from this information plus specification of the boundary conditions prevailing when they are bound. Yet, in real science, researchers frequently find that what they know about the behavior of entities in their non-bound condition fails to reveal how they will behave in various complex environments. The behavior of atoms as they behave independently reveals little of how they will behave when bound into molecules; likewise, the behavior of amino acid strings reveals little of how they will behave when folded into proteins. Instead, how such entities will behave in bound situations has to be determined empirically. (One indication of this is that when research teams include scientists from both lower-level and higher-level disciplines, the relationship is not one in which the lower-level scientist provide general theories and the higher-level scientist derives the consequences. Rather, all recognize they must discover new information and that what the lower-level scientist often has to offer are techniques that can help reveal how the component parts are behaving in the more complex environment.)

An alternative strategy is simply to incorporate into the lower-level theory everything that is learned about lower-level entities as they are bound into various structured wholes. Clifford Hooker adopts this view:

> First, the mathematical development of statistical mechanics has been heavily influenced precisely by the attempt to construct a basis for the corresponding thermodynamical properties and laws. For example, it was the discrepancies between the Boltzmann entropy and thermodynamical entropy that led to the development of the Gibbs entropies, and the attempt to match mean statistical quantities to thermodynamical equilibrium values which led to the development of ergodic theory. Conversely, however, thermodynamics is itself undergoing a process of enrichment through the injection "back" into it of statistical mechan-

> ical constructs, e.g., the various entropies can be injected "back" into thermodynamics, the differences among them forming a basis for the solution of the Gibbs paradox. [Hooker, 1981, 49]

The idea that lower-level theories need to be enriched to account for what is learned at the higher level leads to a view that reduced and reducing theories co-evolve, a view that Patricia Churchland [1986] espouses for the relation between psychology and neuroscience. The difficulty with this approach is that lower-level accounts of the behavior of entities when they are bound in complex structures may share little with accounts of how they behave in isolation. The resulting lower-level theory may be so complex and its various claims sufficiently unrelated to one another that little unity will have been achieved.

Before leaving criticisms of the theory-reduction account, we should note one feature of the account not often discussed — the role played by boundary conditions. It is only under specific boundary conditions that, on this account, higher-level laws can be derived from lower-level ones. But where do these boundary conditions come from? They are not themselves derived from the lower-level laws. Rather, they must be determined empirically as investigators try to develop the reduction. This has significant consequences for the claims that reduction unifies all higher-level laws in terms of basic-level ones. In fact, the higher-level laws are derived from lower-level theories *plus* bridge principles and boundary conditions. Even if the rest of the theory reduction account proved adequate, it would not promote as much unity between the various sciences as is often suggested.

3 KITCHER'S REVISIONIST ACCOUNT OF UNIFICATION

Pursuing a line of argument first formulated by Michael Friedman [1974], Philip Kitcher has argued for more than two decades that we should be interested in the unity of science because of the tight connection between unification and explanation. Kitcher [1981] defends this view as a means to offering an account of explanation that both builds on the work of some of the logical empiricists (particularly Hempel and Feigl) and overcomes some shortcomings of the covering-law (D-N) model of explanation (and by extension, the theory-reduction model). Three of these inadequacies are of chief importance. First, according to Kitcher [1981, 508], the covering-law model does not make clear just how it is that scientific explanation advances understanding. Second, the covering-law model does not offer a means to weigh the explanatory power of some theory, or of some theory as against another one. Third, the quality of the covering-law model depends on there being a good way to distinguish between laws and accidental generalizations, but this distinction has been famously problematic since Goodman [1955].

Kitcher's emphasis on unification is meant to be a way to retain the logical empiricists' commitment to explanation as derivation. Kitcher is able to avoid the problems discussed above by arguing that successful explanations are part of a "system" or "store" of explanations, such that no putative explanation can be evaluated individually, but rather must be assessed (at least partly) by reference to the rest of the explanations science accepts at a time.

> Science supplies us with explanations whose worth cannot be appreciated by considering them one-by-one but only by seeing how they form part of a systematic picture of the order of nature. [Kitcher, 1989, 430]

The central move here is to accept, with the logical empiricists, that explanations are derivations, but to deny that such derivations can be assessed in a piecemeal fashion. Rather, they must be part of the best systematization of the set of statements accepted by the scientific community at a given time. "Best systematization" here means, roughly, the set of derivations that minimizes the number of argument patterns while maximizing the number of conclusions. The number of argument patterns can be obtained by giving a classification of argument patterns based on inferential characteristics.

The change from individual derivations to a best system of derivations circumvents the three problems noted above by making no use of the law-accidental generalization dichotomy, by providing a means of assessment for the explanatory power of a candidate explanation (a better explanation is one that leads to more conclusions while adding the least number of argument patterns), and finally, by showing how explanations lead to understanding. The unification approach accomplishes the latter by "showing us how to derive descriptions of many phenomena using the same patters of derivation ... and it teaches us to reduce the number of types of facts we have to accept as ultimate (or brute)" [Kitcher, 1989, 432]. On this view, unificatory power is a criterion by which new explanations can be evaluated against old ones, and a means to force explanations to advance our understanding by making them cumulative parts of an over-arching system.

Prompted by critics of unity (see below), Kitcher seems to have softened his view in recent years to one that he calls "modest unificationism" [Kitcher, 1999]. The essential scheme — "finding as much unity as we can by discovering perspectives from which we can fit a large number of apparently disparate empirical results into a small number of schemata" [Kitcher, 1999, 339] — is the same, but Kitcher now acknowledges that the world may indeed be a messy place and that we may have to "employ concepts that cannot be neatly integrated" into a single best system. Still, Kitcher is not willing to abandon unification entirely, as he thinks that explanatory unification functions well as a "regulative ideal".

4 CRITICS OF UNITY

In the late 1970s and on through the early- and mid-1980s, the idea that science is or can be unified even in Kitcher's revisionist sense met with powerful criticisms from a group of philosophers centered around Stanford University. In "The Plurality of Science", Patrick Suppes [1981] offers a short argument to the effect that unity of science theses as conceived by philosophers and scientists down the ages have been poorly supported by theory and practice. The several forms of reductionism upon which these theses rely, Suppes claims, are untenable. What is left is a kind of pluralism of scientific language, practice, and subject matter. These, Suppes argues, are diverging rather than converging, and this is as it should be.

At about the same time as Suppes published his piece on pluralism, Nancy Cartwright was developing her view that the empirical success of our best physical theories argues against, rather than for, the universality of our theories and the unity of science [Cartwright, 1980; 1983; 1999]. John Dupré also [Dupré, 1983; 1993] mounted an attack on the unity of science that was motivated by his understanding of biological science, particularly regarding how natural kinds are identified and differentiated.

Cartwright's opposition to the unity of science works by turning the observations that fund views like the one voiced by Oppenheim and Putnam and Nagel on their heads. Cartwright grants that science can often provide predictions of impressive accuracy and can be used to manipulate certain systems very precisely. She argues that in order to do so, however, the laboratory scientist or mathematical modeler must abstract in crucial ways from the world as we usually encounter it. The charge, at base, is that scientists often describe and model systems that are constituted as much by human engineering as they are by the world. Research systems such as a sealed beaker in a laboratory incubator, or an insulated housing to be sent aloft in a spacecraft, are highly circumscribed and shielded from intrusions. But outside the beaker or box, in the universe at large, the models may very well fail to apply. Cartwright emphasizes that the world is a good deal messier than our theoretical descriptions of carefully and artificially isolated systems in it would lead us to believe.

According to Cartwright, the more restricted relevance of theoretical models suggested by this view should not be cause for concern. We do not usually try to apply models outside their domain of applicability, so this view is not really asking us to give up anything with respect to our use of models for prediction, manipulation, and control. Our models of the mechanics of falling objects do not offer good counsel on what, exactly, will happen even to fairly solid, relatively heavy, though oddly shaped objects dropped from the Golden Gate Bridge into the water below. It's possible for a person to jump or fall from the bridge and be retrieved just beneath it very much alive, as happened to a real estate agent in 1988. More often, one does not survive the fall, as happened to the same real estate agent in 2003. Neither models of mechanics nor of biology will tell us exactly which outcome will result — even for the very same 'object' — because there is no good model that includes all the relevant forces. In this case, mechanical and biological models apply only partially at best.

What Cartwright does ask us to give up is what she takes to be the unsupported assumption that there *could be* such a model — that mechanics can *in principle* be universalized to be useful in those cases where it is currently of limited applicability. In order for models of (for instance) falling objects to be universalized, it must be the case that all instances of falling are relevantly similar. Whether some real case is enough like the model case, Cartwright argues, will have to be worked out for each new application. On this view it is anything but clear that we can build a model to fit every real or imagined situation. This is not a claim about our cognitive limits — Cartwright is not claiming that we cannot build models of some systems because their dynamics are too complex for us to measure or describe. She is arguing, rather, that we ought to consider in such cases whether what we have is a system that is genuinely and relevantly different than the ones we know how to deal with. Where this is so, we should not expect there to be any one small set of theories or models that will come to include all others. The best we can hope for is a patchwork of theories and models that will sometimes be compatible and sometimes will not.

By contrast to Cartwright's focus on models and their applicability, John Dupré's opposition to unity of science arguments focuses on the concepts used in different disciplines of science and is motivated by his view that essentialism about kinds is indefensible and thus that kind-membership is a much messier affair than we usually allow. He argues that most things objectively belong to more than one kind. Moreover, he thinks that privileging one kind-membership claim over another for the same individual is always unprincipled. Take a chicken (or all chickens), for example. Chickens are noticed by both biological taxonomists and cooks, but are chickens more fundamentally members of the taxonomic class 'Aves', or of the kind 'gustatory objects'? Both kinds, Dupré says, are objective, and there is no principled way to prefer one taxonomy to the other or to take one to be more basic. It will do no good, of course, to retreat to the position that one of these kinds is scientific while the other is not: we have neither a principle of demarcation nor reasons to think that science is more basic than cuisine.

For Dupré, though kind-membership is objective, it is also context relative. Is the thing I now have before me a common and domesticated instance of the taxonomic class 'Aves', or the sort of thing that a lot of people like to eat when it has been sautéed with mushrooms and port wine? One answer to this question that Dupré will endorse is 'yes'. Another is that arriving at a 'correct' or unambiguous division of objects into kinds requires one to specify one's underlying intent or theoretical perspective in carrying out the classification.

The upshot of Dupré's ontology for the unity of science debate is that the kind of hierarchical ordering that some unity theses rely upon is essentialist or idealist by his lights, and is therefore not to be found in the world. Sometimes one will get nice orderings, but only for a particular purpose, and the very same objects will often belong to some non-hierarchical ordering as well. Dupré points out that the parts of an automobile are hierarchically ordered only so long as we are interested in them *qua* parts of a car. Old pistons with their rings and wrist pins removed very often end up on the desks of autoshop managers and serve as instances of the kind 'ashtray' and 'paperweight'. When they do, they seem not to be part of a hierarchical ordering of parts.

Those unity of science theses that rely on seeing in past and present science some progress toward identifying the most basic kinds — the few microkinds in terms of which many or all macrokinds can or will be described, derived, or explained — will be frustrated if Dupré's ontology is accepted. On Dupré's picture of the world, identifying some kind of thing as most basic for some pursuit will not make it the most basic for all pursuits or even for all scientific pursuits. Put simply, Dupré's anti-unity thesis is that the world itself is radically disordered. We should not, then, expect any science that accurately describes the world to be itself so ordered as to be unified.

5 INTEGRATION INSTEAD OF UNITY

The underlying idea of both the theory-reduction model and Kitcher's revisionist account is that science will be unified through deductive relations. But a variety of scientific enterprises involve constructing bridges between theories without either one being reduced to the other. Lindley Darden and Nancy Maull saw the importance of integration without reduction and incorporated this characteristic when they advanced the notion of an *interfield theory*. Foundational to their account is the notion of a *field*, which they characterized in terms of the following elements:

> a central problem, a domain consisting of items taken to be facts related to that problem, general explanatory facts and goals providing expectations as to how the problem is to be solved, techniques and methods, and sometimes, but not always, concepts, laws and theories which are related to the problem and which attempt to realize the explanatory goals. [1977, 144]

By downplaying concepts, laws, and theories while emphasizing expectations, techniques, and methods, Darden and Maull departed significantly from traditional philosophical accounts. Their starting point was a field (this notion was first developed by Dudley Shapere [1974]) and its diverse characteristics, not theories that may or may not be part of what the field has to offer.[9] In examining cases in which two different fields became integrated, they arrived at the further notion of an *interfield theory*, "a different type of theory ... which sets out and explains

[9]Darden and Maull's notion of a field focused primarily on cognitive features: "a central problem, a domain consisting of items taken to be facts related to that problem, general explanatory facts and goals providing expectations as to how the problem is to be solved, techniques and methods, and sometimes, but not always, concepts, laws and theories which are related to the problem and which attempt to realize the explanatory goals [1977, 144]. But, as sociologists of science have emphasized, fields are also characterized by social structures — laboratories, departments, funding agencies, journals, and professional societies. There are also various informal networks, such as Derek de Solla Price sought to characterize with the notion of *invisible colleges* [1961; see also Crane, 1972; Chubin, 1982]. Recently techniques such as analysis of citation and co-authorship have been used to identify such networks [Wasserman and Faust, 1994]. These aspects of fields are shaped in part by social considerations but often play an important role in determining, for example, what problems are taken to be serious or what methods are accepted for addressing them. As a result, interfield connections involve more than just interfield theories but interfield communities, which often end up transforming the fields from which they originated.

the relations between fields". They identified several types of interfield relations: (a) structure-function, e.g., physical chemistry targets the structure of molecules while biochemistry describes their function; (b) physical location of a postulated entity or process, e.g., the chromosomes identified in cells by cytologists provide the physical location of the genes postulated by geneticists (a case that also exemplifies structure-function and part-whole relations); (c) physical nature of a postulated entity or process, e.g., biochemistry specifies the physical realization of entities postulated by the operon theory in genetics; (d) cause-effect; e.g., biochemical interactions are a cause of heritable patterns of gene expression.[10]

Such relations between different fields are not always obvious or straightforward to develop, since fields may conceptualize the phenomena they investigate in very different terms. Consider the construction of the interfield theory of vitamins, which successfully integrated research on nutritional requirements with the biochemistry of metabolism. Most B vitamins are either coenzymes or precursors of coenzymes that serve to transport hydrogen or phosphate groups from one macromolecule to another. But prior to the 1930s, neither nutrition researchers nor biochemists could recognize this function. For nutrition researchers, vitamins were a puzzle because they were required in the diet, but only in minute quantities. The working conception of nutrition from the mid-19th century was that nutrients were either burned to liberate energy or recruited into the structure of the animal's body (this was especially true of proteins, but also of fats). The minute quantity of vitamins required in a diet, however, would not provide for generating much energy or building much structure. Moreover, the only known components involved in metabolic reactions were carbohydrates, fats, and proteins and the enzymes that broke them down (catabolized them) into a succession of smaller molecules including pyruvate and succinate. With the rise of biochemical laboratory methods early in the 20th century, researchers learned that such reactions could be maintained in extracts of cells in the laboratory, but only if the substances that became known as coenzymes were provided. No one knew why until it was discovered in the 1930s that the energy released in catabolic metabolic reactions was harvested and stored by reversible reactions in active chemical groups of the coenzymes. For example, carrying hydrogen involved a reduction reaction (picking up hydrogen from a donor) followed by oxidation (handing off the hydrogen to a recipient). Since each active chemical group could reduce and oxidize repeatedly, it made sense that a great deal of work could be done under conditions of minimal replenishment. With this reconceptualization of biochemistry, an interfield theory relating nutrition and metabolism could be developed which helped guide further research in each field. For example, vitamin B_2 was a major component of the flavin nucleotide coenzymes and, in particular, contributed the active group that played such an essential role in harvesting energy. (For further discussion of this case see [Bechtel, 1984].)

[10]See Darden [1986] for an extension of this account to the multidisciplinary integration achieved by the synthetic theory of evolution in the 1930s.

Interfield theories sometimes serve simply to bridge existing disciplines, allowing practitioners in each discipline to utilize techniques developed and knowledge procured in the other. In the most interesting cases, however, constructing a bridge between fields or disciplines results in the construction of a new discipline. For example (see [Bechtel, 2006]), cell biology emerged after World War II from what had been a *terra incognita* between biochemistry and classical cytology. Its visionary pioneers developed techniques for using new instruments to tackle new problems. For instance, the electron microscope was used to identify cell components at a much smaller scale than previously possible and the ultracentrifuge was used to localize particular biochemical reactions in the newly discovered components. The methodological and theoretical bridges constructed between cytology and biochemistry gave rise to cell biology as a new discipline. Not all cases of successful interfield interaction result in new disciplines, however. If the existing disciplines are well-established and there is no uncharted territory requiring new instruments, interdisciplinary clusters such as cognitive science are more likely to result [Bechtel, 1986].

6 REDUCTION VIA MECHANISMS

Although philosophers have generally construed reduction as theory reduction, this notion fits poorly with what is scientists typically call 'reduction'. As Wimsatt [1976b] put it: "At least in biology, most scientists see their work as explaining types of phenomena by discovering mechanisms, rather than explaining theories by deriving them or reducing them to other theories, and *this* is seen as reduction, or as integrally tied to it."[11] To appreciate Wimsatt's claim, it is necessary to understand what is meant by a mechanism and by mechanistic explanation. These notions have been pursued since the late 1980s by an emerging school of philosophers of science focusing on biology rather than physics [Bechtel and Richardson, 1993; Glennan, 1996; 2002; Machamer *et al.*, 2000]. The following provides a basic conception of mechanism:

> A mechanism is a structure performing a function in virtue of its components parts, component operations, and their organization. The orchestrated functioning of the mechanism is responsible for one or more phenomena. [Bechtel and Abrahamsen, 2005]

A central feature of mechanistic explanations, and the one that makes them reductive, is that they involve decomposing the system responsible for a phenomenon into component parts and component operations. Given that parts and their operations are at a lower level of organization than the mechanism as a whole, mechanistic explanations appeal to a lower level than the phenomenon being explained. For most scientists and non-philosophers, such appeals to lower levels

[11] For Wimsatt, the complexity of mappings between lower- and upper-level entities establishes both the failure of translation as required in bridge principles and of reduction as a relation between theories [Wimsatt, 1975, 221].

are the hallmark of reduction. As we will see, though, lower-level components of a mechanism do not work in isolation and do not individually account for the phenomenon. Rather, they must be properly organized in order to generate the phenomenon. The most important feature of mechanistic explanation to bear in mind is that it seeks to explain why a mechanism as a whole behaves in a particular fashion under specific conditions. This strategy in no way undermines the reality of the phenomenon being explained; rather, it begins by treating the phenomenon as something that really occurs when the mechanism operates in a particular set of environments.

It is most convenient to introduce the mechanistic perspective on reduction by considering an example. One of the major activities of cells is the manufacture and export of proteins. Beginning around the middle of the 20^{th} century, cell biologists together with biochemists and molecular biologists set out to explain how cells carry out this activity. Philosophers examining this case have focused especially on how DNA is transcribed into RNA, which then *codes* for the sequence of amino acids that comprise a protein. Even this part of the mechanism is extremely complex. For example, three types of RNA are involved. The sequence information is transcribed (by a complicated set of operations) into the sequence of base pairs comprising messenger RNA (mRNA). But to synthesize proteins, these must be *read* by ribosomes, which are complex structures composed of ribosomal RNA (rRNA) and proteins. They temporarily attach to mRNA strands and move along them. A third kind of RNA, transfer RNA (tRNA) forms bonds with particular free amino acids and transports them to the ribosome. There the ribosome creates peptide bonds between the last added amino acid and this new one before moving down the mRNA and repeating the process. (For an account of the discovery of this mechanism, see [Darden and Craver, 2002].) But this is only part of the mechanism. When proteins are synthesized for export from the cell, the ribosomes are attached to the membrane of the endoplasmic reticulum. The emerging strands are pushed across the membrane into the inner space of the endoplasmic reticulum and then transported to the Golgi apparatus. There they are encapsulated in another membrane and transported across a series of sacs (the saccules of the Golgi stack). There carbohydrates are combined with the proteins to create secretory particles, which are then excreted from the cell through the process of exocytosis [Whaley, 1975; Bechtel, 2006].

One important point to note from this example is that the components of the mechanism do different things than what the mechanism as a whole does. Individual lower-level components do not explain the overall performance of the mechanism. Individual enzymes, for example, catalyze particular reactions. They do not perform whole physiological activities such as protein synthesis. Only the mechanism as a whole is capable of generating the phenomenon, and then only under appropriate conditions. Herein lies the explanation for the need for bridge principles in the theory-reduction account — different vocabulary is needed to describe what the parts of a mechanism do than is required to describe what the mechanism as a whole does. The appropriate bridge in this case, however, is not a set of translation rules, but an account of how the operations of the parts of the mechanism are organized so as to yield the behavior of the whole mechanism.

One consequence of taking apart a mechanism that depends on organization to generate the phenomenon is that the investigator destroys the phenomenon itself. A not uncommon situation in science is that after investigators decompose a system they find they cannot readily put it back together again. Sometimes this is because they have neglected some important component. But more frequently it is because they have failed to recognize the specific mode of organization that was involved in the functioning mechanism. The simplest mode of organization is to relate the operations of different parts in a linear series. Understanding more than this simplest mode of organization has presented a serious challenge to humans [Bechtel and Richardson, 1993].

A simple but extremely powerful organizational principle is a negative feedback loop in which the product of an operation feeds back into an earlier operation, allowing for its regulation. (Recall that negative feedback was the central principle advanced by the cyberneticists and general systems theorists in their proposals to unify science.) We are all familiar with this kind of organization from mechanical systems in the home. In the heating system, for example, a thermostat monitors the output of an operation (the heating of the air) and, when the desired temperature has been reached, sends back a signal that stops the furnace from generating more heat. As familiar as negative feedback is today, it was a very difficult concept for engineers and scientists to acquire. It was reinvented numerous times, each in a specific application (for a discussion of the history of re-discovery of negative feedback, see [Mayr, 1970]). Ancient water clocks, for example, required that the water-supply tank be maintained at a constant level; in approximately 270 BCE, Ktesibios invented a feedback control system for such clocks. Windmills need to be pointed into the wind, and British blacksmith Edmund Lee developed the fantail as a feedback system to keep the windmill properly oriented. A temperature regulator for furnaces was developed by Cornelis Drebbel around 1624. Finally, James Watts' invention of a governor for his steam engine helped establish the principle as a general one for use in engineering. This was in large part a result of the mathematical analysis of such control systems in terms of differential equations developed by James Clerk Maxwell.

Recognizing negative feedback control in biological systems was equally difficult. Vitalists in the 19^{th} century objected to mechanist accounts in physiology on the grounds that they could not conceive how a mechanism could behave in the manner biological organisms were known to behave.[12] In particular, organisms maintain themselves in the face of various assaults of their environment. Claude Bernard [1865] developed a framework for answering such objections by distinguishing between an *inner environment* in which the organs of an organism

[12]Bichat [1805], provides some of the most compelling arguments of such a type for vitalism. He focused, for example, on the apparent indeterminism in the responses of organisms to external stimuli and the tendency of organisms to behave in ways that resisted external forces that would kill them.

function and the *outer environment* in which the organism lives. He proposed that each organ in the body was designed to respond to specific changes in the internal environment so as to help maintain the *constancy of the internal environment*. As a result of the actions of the various organs, the inner environment provided a buffer against conditions in the external environment. Bernard, however, was not able to characterize in any detail how the organs each helped to maintain the constancy of the internal environment. Walter Cannon [1929] picked up this thread from Bernard and introduced the term 'homeostasis' (from the Greek words for 'same' and 'state') for the capacity of living systems to maintain a relatively constant internal environment. He also sketched a taxonomy of strategies through which animals are capable of maintaining homeostasis. The simplest involve storing surplus supplies in time of plenty, either by simple accumulation in selected tissues (e.g., water in muscle or skin), or by conversion to a different form (e.g., glucose into glycogen) from which reconversion in time of need is possible. Cannon noted that in most cases such conversions are under neural control. A second means of maintaining homeostasis is through negative feedback — measuring the effects of a continuous process and using that to alter the rate of its performance (e.g., measuring internal temperature and when it is too high or too low increasing or decreased the rate of blood flow by modifying the size of peripheral blood vesicles).

Negative feedback is frequently realized in biological systems as a result of cyclic organization in which the products of several successive chemical operations ultimately combine with some new input to produce an earlier intermediate. The citric acid cycle, first advanced by Krebs and Johnson [1937], provides an illustrative example (see figure 4). The ultimate function of the citric acid cycle is to enable synthesis of ATP, the macromolecule in which energy is stored in animal cells for use in such activities as muscle contraction. Specifically, energy is stored in a high-energy bond created by adding a phosphate group to ADP. A small amount of ATP is generated within the citric acid cycle itself (substrate-level phosphorylation), and a larger amount using the energy that is released by oxidative reactions in the cycle and transported, in the form of NADH or FADH, to another mechanism (oxidative phosphorylation). There is no point in performing the oxidations in the citric acid cycle at a rate that exceeds the system's capacity to synthesize ATP from ADP. Hence, when this happens, NADH and FADH build up and there is no NAD or FAD available to support further oxidations in the citric acid cycle. Thus, the rate of the citric acid cycle is regulated by means of negative feedback. The less ADP available, the less NAD and FAD is available, and therefore the less oxaloacetic acid is available to react with acetyl-CoA, the substrate that typically enters the cycle from other metabolic processes.

Although once the citric acid cycle was discovered its functional significance became apparent, the work leading to its discovery had other motivations. The spur to develop this and other cycles was the realization that the initially conceptualized linear pathway of reactions resulted in a product that, lacking hydrogen, could not be further oxidized directly. Recombination with something else was

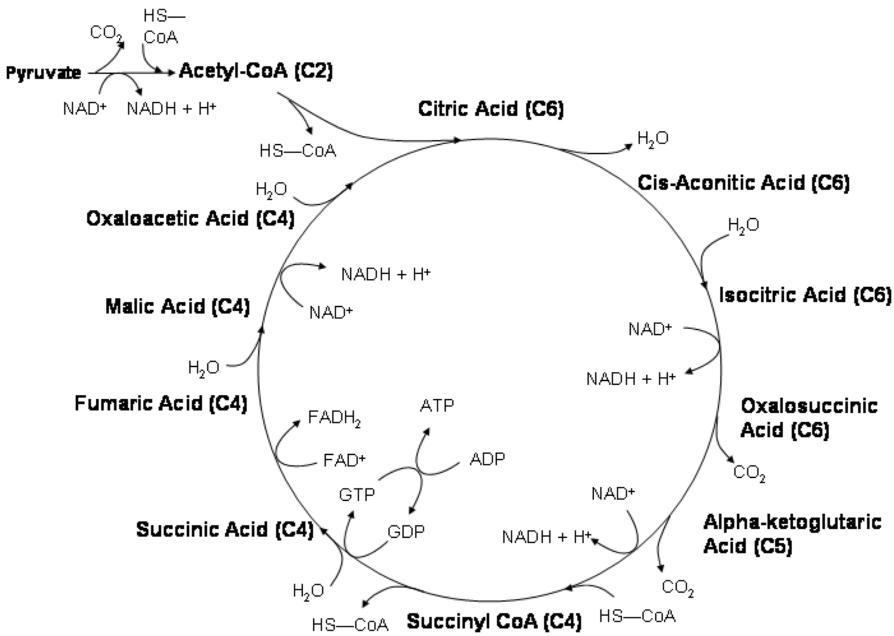


Figure 4. The citric acid cycle, a central biochemical reaction in cell metabolism. The crucial oxidation reactions are shown in the interior. When energy demands are low, there is no ADP available, which in turn means there is no NAD^+ or FAD^+ available (all supplied being taken up in NADH or $FADH_2$). This will result in no accumulation of oxaloacetic acid to react with acetyl-CoA, thereby bringing the reactions in the cycle to a halt. Trough such feedback, critical metabolites are conserved until they are needed to synthesize new ATP from ADP.

an expedient to overcome this obstacle. In short order biochemists discovered a number of cycles, such as the citric acid cycle, and began to appreciate cyclic organization as a common design principle in living organisms. But this was a hard-won battle since the focus remained on the overall production of the end product from the input, not the organization in between.

As difficult as it was to understand the significance of negative feedback, the importance of positive feedback was even more difficult to appreciate. At first positive feedback seemed not to be very functional since it appeared to lead to run-away mechanisms. That is, if the product of a mechanism spurred the mechanism to produce yet more of it, the process would continue until all supplies were exhausted. Yet, there are constrained contexts in which positive feedback is desirable. Particularly important are sets of reactions that function autocatalytically, with one reaction producing a catalyst for a second reaction, and it in turn producing a catalyst for the first reaction [Kaufmann, 1993; Maturana and Varela, 1980]. Theorists interested in the origins of life have been the leaders in exploring these ideas (see, for example, the intriguing models of Gánti [1975; 2003]), but they have yet to achieve major uptake in the broader scientific community.

It is easiest to recognize the role of organization in generating higher levels by considering the perspective of an engineer who has been asked to organize existing components in a new way to accomplish some task. When she has finished, she has built something new, perhaps something for which she could secure a patent. We would not expect the patent office to deny her a patent because all of the components were already known to her — they were also known to the others who failed to have the insight needed to develop the new mechanism. Thus, invention of a new organization alone is noteworthy. (In real life, an engineer would more often invent some of the components as well as their organization. However, at some level of decomposition the invented components would themselves be built from existing ones.)

Beyond organization, the environment is often key to understanding how a mechanism works. Mechanisms are not isolated systems, but depend on conditions in their environment. This is particularly the case for biological mechanisms as against physical machines that may be engineered to perform in an identical fashion over a wide range of conditions. With biological mechanisms evolved to operate in a specific range of environments, features of the environment may be co-opted into the mechanism's operation. Evolution is an opportunist, and if something can be relied upon in the mechanism's environment, then it does not have to be generated by the mechanism. Vitamins provide just one well-known example. Because our ancestors could generally count on the availability of vitamins in their foods, there was no evolutionary pressure for us to retain the ability to synthesize them. Nonetheless, insofar as such environmental factors are necessary for the functioning of the mechanism, mechanistic explanations need to focus on the mechanism's context, not just its internal configuration.

With this account of mechanisms and mechanistic explanation in place, we can consider further how they offer a fresh perspective. Unlike theory-reduction accounts, mechanistic reductionism neither denies the importance of context or of higher levels of organization nor appeals exclusively to the components of a mechanism in explaining what the mechanism does. The appeal to components, in fact, serves a very restricted purpose of explaining how, in a given context, the mechanism is able to produce a particular phenomenon. There are other differences as well. Whereas theory reduction is often treated as transitive, with higher-level theories ultimately being reduced to those at the lowest level, mechanistic reductions often proceed for only one or two iterations. Once investigators understand the operations performed by the parts and how the organization orchestrates their operation to produce the phenomenon, they generally have neither the desire nor the tools to pursue a further round of decomposition into subparts and suboperations. Moreover, it is not the case that detailed knowledge of how the component parts or subparts operate will already be available in lower-level disciplines, since, as we discussed above, these parts will be operating in specialized contexts not typically studied by practitioners of the lower-level science. While the study of mechanisms is reductionistic and can promote integration of knowledge from various disciplines, it does not promote a grand unificationist vision.

6.1 Rethinking Levels

The notion of levels plays a central role in all accounts of reduction, but it has not been fully explicated in any of them. In the early accounts of theory reduction, levels were associated with broad scientific disciplines, so that one sees reference to the physical level, the chemical level, etc. But just why the objects of physics, which range in size from the sub-atomic to the universe, comprise a level is left unspecified. Although still committed to the theory reduction framework, philosophers such as Causey approached levels from a more ontological perspective, emphasizing that lower levels deal with the parts of wholes studied at higher levels. Wimsatt develops this mereological perspective, making part-whole relations fundamental in distinguishing levels:

> By level of organization, I will mean here compositional levels — hierarchical divisions of stuff (paradigmatically but not necessarily material stuff) organized by part-whole relations, in which wholes at one level function as parts at the next (and at all higher) levels .... [Wimsatt, 1976a]

One limitation of compositional relations from Wimsatt's perspective is that they do not permit ordering of entities not part of the same part-whole hierarchy. Accordingly, Wimsatt also appeals to interactions between entities in identifying levels — entities interact principally with others at their own level and with entities at lower levels in terms of the complexes of which they are part. People, for example, interact primarily with other people, animals, plants, computers, furniture, etc., not the cells of other people or the chips of computer. Accordingly, Wimsatt comments: "*Levels of organization can be thought of as local maxima of regularity and predictability in the phase space of alternative modes of organization of matter*". [Wimsatt, 1994]

Wimsatt notes that the neat layering of levels breaks down at higher levels — individual humans do engage in relations with entities several times larger or smaller than themselves. Accordingly, he introduces the notions of *perspectives* and *causal thickets* for cases in which neat layering into levels breaks down. But the problems go deeper and calls into question the general project of conceiving of the natural world as layered in terms of levels. In biology it is routine for things of very different size-scales to interact. The transfer of energy released in basic metabolism to ATP, for example, is mediated by the transport of protons across the inner mitochondrial membrane, and its diffusion back. Yet the very membrane that is maintaining the proton gradient is also composed in part of protons. Protons are thus part of the very structure through which the protons are being transported. Thinking in terms of the operation of the mechanism, it is correct to say that the protons in the membrane are at lower level than those being transported across it.

Thinking in terms of mechanisms allows one to articulate a more limited but less problematic conception of levels. From the point of view of a given mechanism performing a particular function, the component parts into which a researcher decomposes it constitutes a lower level. If researchers decomposed these parts, they reach yet a lower level. This account allows for the denizens of a level to be of different sizes as long as they are working parts of the same mechanism. Moreover, it is compatible with viewing two structurally identical entities as at different levels if one performs its operations in a sub-mechanism of another — a proton that is being pumped across a membrane is at a higher level than one that is part of the membrane. But an important feature of this account of levels is that they are limited to the scope of the original mechanism.

One advantage of construing and limiting the notion of levels to levels of organization in mechanisms is that it permits a coherent account of the important idea that lies behind the problematic notion of downward causation [Campbell, 1974]. The important idea behind appeals to downward causation is that causal effects of interactions of higher-level entities have consequences for their component parts. Your DNA is a passenger on all your travels and some of your neurons are altered every time you learn something new. The notion of downward causation is problematic, though, since it seems to result in a problem of causal overdetermination — if we assume that there is a comprehensive account of causal interactions of entities at a lower level, then the effect is already determined regardless of any putative top-down effect [Kim, 1998]. One solution to this problem is to keep the notion of causation univocal by restricting it to intralevel cases and provide a different, constitutive account of interlevel relations within a mechanism [Craver and Bechtel, in press]. The intuition behind top-down causation can be maintained, but expressed in terms other than causation: the causal interactions of a mechanism with its environment (including other mechanisms) alters the mechanism itself. The changed condition of the parts and operations within the mechanism then propagate causal effects within the mechanism.[13]

A consequence of the mechanistic approach is surrendering the view that a complete causal story can be told at the lower level — all one can account for is changes in the mechanism as the parts operate and interact with each other under the conditions in which the mechanism is operating (some of these being set by the interaction of the environment with its environment). Since it does not have the resources to describe the way in which the mechanism engages its environment, the lower-level account of goings-on inside the mechanism cannot provide a complete account of all that is happening. Our discussion of the problems with global unity theses, though, suggests that the aspirations for a complete theory should

[13]On this view, so-called *bottom-up* causation works in the same manner — the operation of parts within the mechanism alters the condition of the mechanism itself, thereby altering the manner in which it engages its environment.

be surrendered anyway. What a mechanist requires is only that the causal effects at a given level within a mechanism can be explained — for example, that one can explain how, given the impingements on the brain from the environment, neural changes within it occur. This is precisely what molecular accounts of learning and memory strive to do [Craver and Darden, 2001]. The level of neural processes inside the brain is locally constituted — it is not part of a broad level that crosses mechanisms.

6.2 Within Level identities: Heuristic Identity Theory

In characterizing mechanisms we identified both parts and their operations. The research tools for decomposing mechanisms into their parts and operations are often different. As a result, the decompositions are often developed in different disciplines. For example, cytologists using various microscopes, identified various organelles in the cell, whereas biochemists, preparing homogenates and using various assays, identified chemical reactions. One of the accomplishments of modern cell biology was to establish that different cell functions were performed by specific cell structures, thereby localizing the function [Bechtel, 2006].[14] Since localization claims maintain that it is the same entity that constitutes a particular structure and has performs a specific operation, they are identity claims in the sense advanced by the mind-brain identity theory [Place, 1956; Feigl, 1958/1967; Smart, 1959] noted above. The identity theory is often construed as advancing a reduction of psychology to neuroscience, since neuroscience is at a lower level than psychology. From the point of view of mechanistic explanation, however, we can recognize that accounts of the part of the system and the operation it is performing are at the same level. For example, initial encoding of information to be stored as long-term episodic memories (an operation described by psychology) is an operation of the hippocampus (a structure identified by neuroscientists).

Although not themselves vehicles of reduction, since they are intralevel claims, identity claims play an important role in mechanistic research and ultimately help advance mechanistic reductions. One way to see this is to consider one of the major objections that critics raised to the mind-brain identity claim. They charged that at best empirical investigation could establish a correlation between the psychologically characterized phenomenon and a brain process, an objection that has been pressed anew in recent discussions of consciousness [Chalmers, 1996]. Despite the prevalence of the language "neural correlates" in recent presentations of empirical research concerning consciousness [Crick and Koch, 1998], most empirical researchers do not make a distinction between establishing a neural correlate and identifying the neural substrate. It is philosophers who insist in emphasizing that the empirical evidence cannot decide between correlation and causation. One import of making such a distinction is that a dualist can maintain that conscious

[14]Linking structural and functional accounts developed in different fields was one of Darden and Maull's major examples of an interfield theory. In general, interfield theorizing often culminates in accounts of mechanisms.

states are not material phenomena at all, but are simply correlated with brain processes.

When considered in the context of how identity claims typically figure in empirical research, however, the attempt to reconstrue them as correlation claims appears radically misguided. The reason is that they typically are not the conclusions of scientific investigations but heuristics for guiding further scientific discovery [McCauley, 1981]. Once an identity claim is made between a structural and a functional characterization of an entity, researchers use each characterization as a guide to elaborating the other. Discovery of an operation that cannot be linked to a part of the structure poses the question of whether that operation is indeed being performed and if so, by what component. Discovery of a component of a structure that does not seem to be performing any operation raises the question of whether it really is a working part and if so, what operation has been missed in extant functional decompositions. Such research invokes the converse of Leibniz's law of the identity of indiscernables, focusing instead on the indiscernability of identicals: what is learned about a structure or a function under one description must apply to it under the other, or one must revise the identity claim. Correlational claims, by contrast, impose no such burden. To indicate its constructive role in guiding further research, Bechtel and McCauley [Bechtel and McCauley, 1999; McCauley and Bechtel, 2001] speak of *heuristic identity theory*. Once an identity claim has fulfilled its heuristic function of guiding discoveries both on the structural and functional sides, the identity has been woven into the science and investigators who had taken advantage of the heuristic would not be tempted to consider it a mere correlation.

As noted above, identity claims are not themselves reductive since they relate different accounts of the same entity. They do, however, directly contribute to integration between different accounts of the phenomenon, often ones developed in different disciplines with different research techniques.

7 CASE STUDIES IN REDUCTION AND UNIFICATION ACROSS THE DISCIPLINES

Although we noted examples from various sciences to illustrate points in the previous sections, the focus was on the conceptual account and its continuity. Looking at actual cases of reduction and unification/integration reveals that they are quite diverse. In this final section we examine four cases that have been important in the discussion of reduction and unity. In each case we ask how the foregoing discussions applies and, in the last cases, identify foci that have not been sufficiently developed in accounts to date and should serve as topics for further philosophical investigation.

7.1 Temperature: Thermodynamics and Statistical Mechanics

At the end of Section 2.3 above, we pointed to the importance of the role played by boundary conditions and bridge principles in carrying out theory reductions of higher-level laws to lower-level ones. In this first case study we revisit this feature of reductions by a deeper look at the relationship between thermodynamics and statistical mechanics, the standard example of successful theory reduction since Nagel [1961]. As we saw in section 2.1, temperature in particular has long been regarded by many in the scientific and philosophical communities as completely explained in terms of the mean kinetic energy of lower-level particles (molecules): $2E/3 = kT$. Indeed, we now learn from some standard high school and university textbooks and from renowned physicists that temperature *just is* mean kinetic energy of the molecules that constitute the gas [Feynman, 1963, 39].

Several problems with this identity claim have been noted by philosophers and physicists, many of them having to do with boundary conditions. Philosopher Mark Wilson reminds us, for instance, that while the simple equality claim holds in the case of classical gases — the case Nagel emphasized — it is not anything like universal: "in point of fact, this temperature equation is generally false; the proportionality between temperature and kinetic energy is substance specific" [Wilson, 1985, 228].[15] As Nagel pointed out in developing his example, the kinetic theory of matter includes both the general postulates of statistical mechanics and more specific postulates appropriate to classical gases — those that are thermodynamically isolated, dilute, and in which the particles influence each other only by perfectly elastic collisions. The kinetic theory, of course, gives excellent predictive results for substances that are relevantly like those described by its postulates. But what about other kinds of substances or even non-dilute gases? Because of the way solids are constituted, for instance, the molecules cannot collide as they do in gases, but can only vibrate. Similar problems arise for other states of matter. It turns out that the observable macrophenomenon we call temperature is multiply realizable at the microlevel.

What this means for the quality of the reduction generally is not quite clear — except that there is good reason to think, as Lawrence Sklar puts it, that we "do not expect to 'deduce' or 'derive' thermodynamics from statistical mechanics in any simple minded way ..." [Sklar, 1974, 16]. In the case of temperature, there will not be just one reduction, but several, as boundary conditions for several states of matter, types of gases and for fluctuating energy situations will have to be specified. Some have argued that this situation causes no real problem for the reduction — we just need to be careful about specifying the boundaries of the reduction.

In addition, as we pointed out above, such specification relies importantly on empirical, rather than deductive, evidence. The descriptions of various states of matter and how they behave has been achieved experimentally, not deduced from

[15] As Lord Kelvin pointed out, it is possible, of course, to construct an absolute temperature scale — a scale on which what is being measured is not relative to what is being used to measure it. This is a separate issue from the one we are raising here.

the relevant lower-level theory. While statistical mechanics has thrown light on the knowledge gained from experiment, it is not the case that the relevant boundary conditions for temperature can be read off the axioms of statistical mechanics. Neither is it immediately clear how far from the 'ideal' boundary conditions a system can be before the lower-level laws cease to offer acceptably good predictions of the behaviour of that system at the higher level. This, too, must be investigated empirically, at least until standards are articulated.[16]

Given all this, even a 'successful' reduction in this seemingly simple case will turn out not to be as unificatory as many proponents of the theory-reduction model would have hoped. The reduction will be complicated, disjunctive, and empirically informed, rather than simple, general, and purely deductive. Indeed, the more general and unifying principles are actually those of classical thermodynamics, not the reductive bases.

It is worth noting that mechanistic reduction may provide a superior way to understand this case. The main problems noted above can be side-stepped: mechanistic reduction does not deny the importance of specifying the relevant context, neither does it demand that relations be deductive. Instead of an attempt at reduction that issues in a simple and powerful proportionality that fails to achieve full generality, a mechanistic explanation will be sensitive to boundary conditions in addition to the relations between higher- and lower-level phenomena and entities. This argues against unity, not for it, because we should not expect the physicist who works with concentrated gases to consult the physicist who works with dilute gases when she defines temperature for the systems on which she works. The simpler, better understood case has no obvious claim to epistemic superiority. On the contrary, *each* mechanistic explanation will be relatively substance specific and it is anything but clear that one is the best or more appropriate model for all the others.

The prospects for Darden and Maull-style integration also seem more promising than those for unity by theory reduction. Indeed, a great amount of integration has already taken place. Structure-function and cause-effect accounts on which relations between micro and macroproperties are specified are at the heart of thermal physics. So too are accounts from the perspective of the microlevel of the nature of features and processes at the macrolevel. These descriptions and accounts often represent the integration of different fields, of which thermodynamics and statistical mechanics are just one example.

7.2 Genes: Molecular Biology and Developmental Systems Theory

From what has been the primary exemplar case in philosophical accounts of reduction, we turn to one that we have also alluded to above and is currently capturing

[16]We have focused on temperature because of its familiarity and centrality in the reductionism literature, but problems with entropy have also been widely discussed as a possible confounder for the reduction of thermodynamics to statistical mechanics. For discussion see Sklar [1993] and Callender [1999].

both scientific and popular attention in the life sciences. Very near the end of the famous paper in which the outcome of their work on the structure of DNA is announced, Watson and Crick offer the following single-sentence paragraph: "It has not escaped our notice that the specific pairing [of bases] we have postulated immediately suggests a possible copying mechanism for the genetic material" [Watson and Crick, 1953]. With this was born a new emphasis on DNA as the ultimate source for knowledge about the macrofeatures of organisms. Biology soon had a new "central dogma" — DNA makes RNA makes protein — and with it an explicitly reductionist (gene-based) approach to accounting for all sorts of biological phenomena, including phenotypes [Dawkins, 1976], the evolution of morality [Ruse and Wilson, 1986], and even human belief in God [Hamer, 2004]. This approach quickly led to widespread accounts of macroproperties of organisms or groups of organisms in terms of genes. Some property P could be explained by or deduced from the presence (or absence) of the gene for P. Dean Hamer's recent claims about the gene for belief in god, or "self-transcendence," are a good example. Hamer argues that whether or not one believes in god is best predicted by whether or not one inherits the VMAT2 gene, the 'gene for' belief.

The gene-based approach, however, has important problems. As Oyama, Griffiths, and Gray [2001] have pointed out, privileging DNA's role in biological processes makes inheritance, evolution, and development, for instance, the mere passing on of DNA. On this view, DNA becomes the only relevant causal factor in these and other biological processes, and the locus of explanation for them. Richard Lewontin has pointed out on several occasions and at some length, however, that the central-dogma view cannot be the whole picture, because DNA can have no such causal efficacy. DNA, he contends, "is not self-reproducing", "makes nothing", and does not determine much, if anything, about organisms [Lewontin, 2000]. Without the rest of the cellular machinery of proteins and enzymes, DNA produces nothing at all. To extend a well-used metaphor, if DNA *codes for* this or that protein, there must be something that *reads* the code, something that *builds* what the code specifies, and perhaps most importantly, something that *writes* the code for the next iteration. DNA cannot do all this.

Another significant problem with the gene-based approach to accounting for macrofeatures is that being in possession of the full genome sequence does not by itself tell researchers much about the properties of the organism. Far from having a gene-for map that offers one-to-one correspondence of molecules to macrofeatures, we have learned that a great many genes have regulatory functions — they 'switch' other genes on and off rather than code for the manufacture of particular proteins. It is worth quoting the following passage from Karola Stotz and Adam Bostanci [2005]:

> *Gene regulation* means that there is always more involved in the production of the product than the coding sequence. In the case of *alternative* cis-*splicing* of exons and introns, one structure contains several modules that can be alternatively spliced together. One stretch of DNA may therefore give rise to several proteins. *Overlapping genes* and *al-*

> *ternative reading frames* entail that the "same" DNA sequence can yield different products. *Cotranscription* of adjacent DNA sequences blurs the boundaries between structural "genes". In the case of *trans-splicing*, one might say that two "genes" (if a gene is defined as a unit of transcription), are involved in coding for a single protein (or more than one products [sic] as in the case of *alternative* trans-*splicing*). Mechanisms such as *exon scrambling*, *exon repetition*, or *antisense*-trans-*splicing* further increase the divergence of DNA sequence and protein product. *mRNA editing* exchanges single nucleotides in the linear sequence. Last but not least, *protein splicing* changes the final product once more, but in this case by splicing so-called 'inteins' in and out of the final polypeptides of which proteins are composed.

The phenomenon of gene regulation clearly shows that in order to have good explanations of what genes are doing, we need to know what is being regulated and how. These explanations ask for more context than is available at the level of the molecular gene alone, and often come from physical chemistry rather than from genetics. This further suggests that privileging the gene as the locus of explanation is premature in at least some cases. There are also higher levels to consider: How did the genotype-phenotype map get to be the way it is? Why and how is it stable across generations?

Recently, developmental systems theory has emerged as a competitor for gene-based thinking about developmental biology. Proponents of developmental systems theory argue that development cannot be understood outside the framework of its neighbor disciplines and processes, and thus that the causal contexts of heredity and evolution cannot safely be ignored if developmental processes are to be explained. On this view, molecular genetics is just one part of a long and complex story — a story in which genetic goings-on do not make up the only plot.

The developmental systems approach rejects simple reduction of macrofeatures to molecular genetics and urges that there are very often several causal factors in a given developmental process. This viewpoint makes room for the kinds of alternatives to reduction discussed above. Mechanistic reduction, in particular, seems useful for explaining developmental processes in ways that do not neglect epigenetic influences. Mechanistic explanations, by their nature, account for phenomena in context and across levels or organization, rather than privileging a particular level.

This approach is exemplified by recent work on heterochrony — changes in the timing of events or processes during organismal development — as it applies to evolution. Researchers who have investigated differences in organisms that arise as a result of heterochrony have recognized that heterochrony is often not driven by the mere presence of some gene or other. Rather, there may be differences in the timing of gene expression or of the rates of expression. These processes are very often described in mechanistic terms (see, for instance, [Wray and Love, 2000; Tautz, 2000] and the review article by [Smith, 2003]), and researchers have not generally assumed or argued that in those cases where heterochrony can be mechanistically related to particular genes, gene products, or differences in the timing of gene expression, the observed differences can be explained at the molecular level. Even with the molecular part of the story in hand, if we are to apply what we know to evolutionary development, we will still want to know whether and how heterochrony leads to major evolutionary transitions, how the developmental process is regulated for embryos, and at what level(s) of organization this regulation is orchestrated. It is interesting to note that at present the best-known candidate for a developmental regulator in at least some organisms is the so-called somite clock. It is a kind of feedback mechanism responsible for the timing of segmentation in the vertebrate embryo that is usually described as operating at the cellular, rather than molecular, level [Pourquié, 1998; Dale and Pourquié, 2000].

There is also a strong case to be made that the proponents of developmental systems theory are calling for an explanatory strategy like the one advocated by Darden and Maull. We can see molecular genetics, embryology, cell biology, and other disciplines as fields that all have some relation to development, and the search for a better understanding of developmental systems as an attempt to specify interfield relations for particular developmental processes. There is no reason, however, to assume beforehand that the field concerned with the lowest level of organization is epistemically prior or more basic. Take, as a simple example, the well-known case of inheritance among diploid organisms. Studied from a molecular level, we only learn about gene variation at certain loci. Couple this knowledge, though, with the study of cellular mechanisms and we can begin to see why Mendel's second law holds: the process of meiosis regularly distributes each allele such that the assortment is independent of every other allele. Population genetics tells us still more of the story, informing us as to what the distributions of alleles will be when no outside forces are operating.

Choosing any one of these levels as primary artificially limits the inquiry in ways that may not be heuristically justifiable. At the cellular level, we can ask structure-function questions of the molecular level, as well as cause and effect questions. From the molecular and cellular perspectives we can ask about the physical processes that underlie the regularities captured by population genetics. We can also hope, as developmental systems theorists do, that not limiting ourselves to a single perspective will result in interfield theories that parlay knowledge at these various levels into a more thoroughgoing account of evolutionary development.

It is important to note that in the case of heterochrony and in the case of diploid inheritance, molecular genetics does not provide a sufficient account on its own. Rather, it requires interfield connections with developmental and evolutionary biology or explanations that pay attention to the important connections between the molecular, cellular, phenotypic, and population levels.

7.3 Historical Archaeology: Physical and Social Sciences

So far we have focused on the explanatory gain that results from integration of fields — interfield theories and accounts of mechanisms enable investigators to answer a multitude of questions that they could not otherwise address. But there is an additional virtue, one that has been clearly brought out by Alison Wylie [1999] in her account of historical archaeology. Drawing upon the insights of Ian Hacking [1983] on how scientists triangulate independent research techniques to secure reliable evidence even when they cannot directly establish the reliability of any one technique, Wylie shows how historical archaeologists are affecting such triangulation. The approaches of traditional history, which relies primarily on the analysis of documents, and archaeology, which has relied on the analysis of material remains of societies, are radically different. In many cases there is no potential for integrating them. Prehistoric civilizations have left no written documents and they have been the province of archaeologists. The material remains of more recent societies are often destroyed and historians have relied primarily on the analysis of documents to describe their history. But there are a range of early human societies for which both documents and material remains can be recovered. While practitioners of traditional history and traditional archaeology have tended to insist on the primacy of their own tools of investigation, starting after World War II a number of investigators attempted to integrate the two and have adopted the name *historical archaeology* for this integrated investigation.[17] In the U.S., for example, historical archaeologists tended to focus on early European settlement and the effects of these on native American peoples as well as subsequent expansion of the frontier and urbanization of the continent. Its institutional structure did not materialize until the late 1960s. They have attempted to weave together results from analysis of documents and archaeological remains.

As Wylie notes in describing the sometimes tempestuous relations between historical archaeologists and their home disciplines,

> A recurrent theme [sounded by advocates of historical archaeology] ... is an insistence that when events and conditions of life or historic periods are at issue, vastly more can be achieved by making conjoint use of the evidential, methodological, and theoretical resources of archaeology and documentary history than can be achieved by either field working in isolation from the other. [Wylie, 1999, 305]

What is significant is that the attempts to integrate sources often forced revisions in the accounts compiled from one source alone. By drawing upon archaeological methods to study the artifacts of a society, one is not just a filling in the historical record but procuring "substantially different, potentially transformative insights about the recent past" (p. 305). This stems from the fact that archaeology can provide evidence of people who do not show up in documentary records, illustrating the ways they lived their lives, which then provides a different perspective on the documents left by the cultures in question.

[17]The Society for Historical Archaeology was established in 1967 and began publishing the journal *Historical Archaeology* that year (see [Schuyler, 1978], for a discussion of these events in the U.S. and related developments in other countries during the same period).

Wylie's particular interest in historical archaeology is its potential to provide an illuminating example of how integrating the modes of investigation from multiple disciplines can both provide epistemic warrant beyond what each alone can produce and serve as a heuristic to encourage new inquiry. The key idea behind increased epistemic warrant is Whewell's [1840] notion of consilience of induction according to which results secured through independent lines of inquiry are more likely to be true than those relying on just one line of investigation. Wylie notes, however, that one cannot just assume that because evidence is advanced in two different disciplines that it represents independent evidence and emphasizes the need to tease apart difference in causal processes, independence of background knowledge and theories invoked, and disciplinary independence. These must be evaluated case by case. But she argues that historical archaeology does offer cases of such independent convergence of evidence and offers the convergence in dating by reliance on tree ring counts, radio-carbon decay, magnetic orientation, and evolution of stylistic traditions in documents:

> The disciplines that supply the relevant technologies of detection are certainly institutionally autonomous, and the content of their theories is substantially independent; it is unlikely that the assumptions that might produce error in the reconstruction of a date using principles from physics will be the same as those that might bias a date based on background knowledge from botany or socio-cultural studies of stylistic change. Finally, this independence in the content of the auxiliaries and in their disciplinary origins is especially compelling because it is assumed to reflect a genuine causal independence between the chemical, biological, and social processes that generated and transmitted the distinct kinds of material trace exploited by different dating techniques. [p. 310]

Securing different forms of evidence that can be used to evaluate and revise claims made by any one form of evidence is clearly an important aspect of integrating sciences that applies broadly. In entering the *terra incognita* [de Duve, 1984, 11] that then existed between classical cytology and biochemistry, pioneers in cell biology drew upon two new tools recently developed in physics and chemistry — the electron microscope and the ultracentrifuge. Each presented its own risk of artifact but their combined use, including the use of one to calibrate results from the other, provided investigators with the opportunity to develop an integrated structural and functional account of many basic cell mechanisms [Bechtel, 2006]. Integration thus can serve both an explanatory and an evidential role.

7.4 Language: Linguistics and Psycholinguistics

So far our examples have stemmed predominately from the physical and biological sciences, but we end with one that bridges into traditional areas of the humanities. This case also provides us a glimpse into the dynamics of integrating research efforts across disciplines. Many disciplines in the humanities, social sciences, and engineering focus their attention on products created, intentionally or unintentionally, by human beings. Literary, artistic, philosophical, and technical products typically are constructed intentionally by their authors. Languages and other symbol systems are typically not constructed intentionally, but are nonetheless the products of human activity. How do the disciplines that study these products relate to other disciplines in the physical, biological, and behavioral sciences? We will follow the analysis of Abrahamsen [1987] to discuss one such case: the relationship between linguistics (concerned with the formal structure of human languages) and psychology, especially cognitive psychology (concerned with the mental processes that enable cognitive systems, including humans, to perform their activities). Note that these are different enterprises and typically try to account for different phenomena using different theoretical constructs and appealing to different sources of evidence. Linguists are principally concerned with the structure of language, advance grammars to account for such structure, and test their grammars by their capacity to generate all and only the sentences of a particular language. Psychologists, on the other hand, attempt to explain the mental processes that enable individual language users to comprehend or produce sentences of their language.

Abrahamsen [1987] identifies three patterns in the relationship between linguistics and psychology in the 20^{th} century: (1) boundary maintaining, in which the two disciplines pursued their in quiries independently, (2) boundary breaking, in which one discipline tried to usurp the territory of the other, and (3) boundary bridging, in which practitioners of the disciplines collaborated rather than competing for the same territory. Boundary-breaking episodes often attract the greatest attention. At the turn of the 20^{th} century, psychology was a new and rapidly advancing discipline that attracted a number of young linguists seeking to move beyond the older traditions in their own discipline. What they encountered in psychology, however, was not a single view they could take back to linguistics but competing conceptual frameworks — notably the mechanistic cognitive framework of Johann Herbart and the antimechanistic idealist perspective of Wilhelm Wundt. Wundt [1900] himself addressed a host of issues in both linguistics proper (grammatical structure, phonological systems) and psycholinguistics (language acquisition, speech errors) whereas Herbart influenced linguistics through the applications of his work by the linguist Hermann Paul [1880]. As Blumenthal [1987] describes, these two approaches conflicted — Hobart's approach proceeded bottom-up from sentence elements invoking association techniques whereas Wundt's started with unified, often creative, mental representations and proceeded top-down. The conflict within psychology, according to Blumenthal, soon left linguists disillusioned and many opted to divorce linguistics from psychology [McCauley, 1987].

The second round of boundary-breaking interactions followed Chomsky's introduction of transformational grammar [Chomsky, 1957]. Chomsky viewed his approach to grammar not only as a revolution against structuralism in linguistics proper but also as a revolution against behaviorism in psychology [Chomsky, 1959]. Many psychologists, themselves striving to break free of the behaviorist tradition, eagerly followed Chomsky's lead. Notably, Miller [1962] sought to provide evidence for the psychological reality of transformations. This time it was psychologists who were to be disillusioned, as Chomsky repeatedly revised his grammars regardless of the evidence psychologists offered for their psychological reality ([Reber, 1987]; see also [McCauley, 1987]). Chomsky continued to break boundaries by characterizing many of his ideas as contributions to psychology, including his nativism, competence-performance distinction, and construal of linguistic grammars as accounts of human linguistic competence ([Chomsky, 1965; 1966; 1986], see discussion in [Abrahamsen, 1987]).

Abrahamsen contrasts such instances of boundary-breaking relations with ongoing boundary-bridging interaction between linguistics and psychology. She proposes that a boundary-bridging relation often holds between psycholinguistics, as a subdiscipline of psychology, and linguistics. In this boundary-bridging research, psycholinguists rely on linguists to provide specialized descriptions of, for example, phonemes, distinctive features, and phonological rules, while psycholinguists provide linguists with explanations (e.g., of universal characteristics of phonological systems) and evidence (e.g., for the psychological reality of certain linguistic accounts).[18] Abrahamsen observes, however, that the psycholinguist must often reformat the account provided by the linguist in order to make use of it. Some linguistic theories (e.g., augmented transition network grammars; lexical-functional grammars) require less adjustment than others (e.g., Chomsky's Standard Theory). Abrahamsen comments:

> The psychological studies benefit from ongoing involvement of linguists who are willing to consider psychological goals in addition to their own native goals as linguists. When these linguists carry out their work of linguistic description, they must satisfy two sets of constraints simultaneously, producing descriptions that can be easily applied in behavioral research as well as satisfy criteria of linguistic adequacy. [p. 373]

While boundary-breaking research as characterized by Abrahamsen would promote a unificationist conception of science, boundary-bridging research has far more limited aspirations. In some cases a cultural product discipline such as linguistics may simply provide a description of the phenomena for which psychologists then offer a mechanistic explanation. In other cases the understanding of the mechanism may explain certain linguistic phenomenon (e.g., multiply center embedded sentences such as *the dog the cat the mouse squeaked ate chased* are uncommon because they exceed the working memory capacity of humans). The results are interfield theories, not theory reductions.

[18] Abrahamsen generalizes this framework to many interdisciplinary relations. Subdisciplines of the physical sciences obtain specialized descriptions from the biological sciences, while biological sciences in turn appeal to these subdisciplines for explanation and evidence. The same, she proposes, is true of subdisciplines of the biological sciences with respect to the behavioral sciences, and of the subdisciplines of the behavioral sciences with respect to the cultural product disciplines (mathematics and engineering, humanities, and social sciences).

8 CONCLUSIONS

Visions of unifying all the sciences have been popular ever since the work of the ancient Greek philosophers. Such aspirations were prevalent in many of the historical proposals for unity with which we began this chapter. But the quest for unity can take make forms, often achieving integration rather than true unification. Perhaps the strongest vision of unity appeared in the theory-reduction model of the logical empiricists. This model was attractive because it suggested that logic might provide a powerful way to unite the results all scientific inquiries by showing higher-level theories to be derivable from lower-level ones. Not only were serious objections raised against this model, but as we have seen, much of the unity that appears to result is illusory. Even in the exemplar case of temperature, the bridge principles and boundary conditions have to be established empirically for each type of material in which heat is realized. For many years worries about multiple realizability provided the principal objections to the applicability of the theory-reduction account. A more troubling concern is that any lower-level theory that will provide a foundation from which to derive all higher-level theories will look very unlike contemporary lower-level theories, since it will have to incorporate all knowledge acquired at the higher levels. Altogether, the various objections to the theory-reduction have succeeded in moving it off center-stage in discussions about unity of science.

The problems confronting the theory-reduction model have led some philosophers to abandon the ideal of unity altogether. Cartwright emphasizes the plurality of models that investigators need to deal with the actual world, while Dupré focuses on the need for multiple different ways of categorizing phenomena, each of which is useful for different purposes. Kitcher remains a strong defender of the objective of theoretical unity, but even he has reduced it to the status of a regulative ideal. Still other philosophers, as we have shown, have adopted a reversionary perspective of advocating integration rather than advocating unity. This was the point of Darden and Maull's notion of an interfield theory — it integrates by bridging fields rather than establishing one complete unified theory. It is also exemplified in the notion of reduction which we have identified in the new mechanistic accounts of scientific explanation.

On mechanistic accounts, explanation consists in demonstrating how the orchestrated operation of the components of a mechanism enable the whole mechanism to perform a function in its environment. The conditions imposed on the mechanism from its environment remain a critical part of the explanation, so the higher-level account remains an autonomous component of any explanation. Further, there is no promise that the knowledge of how components behave in a mechanism will be unified with knowledge about how those components behave in other conditions. Lastly, organization turns out to be crucial in getting mechanisms to perform their function, and despite some key theoretical advances in understanding how negative and positive feedback systems enable dynamically organized mechanisms to maintain themselves, this inquiry is still in an early stage. Nonetheless, as the developments in the life sciences in the 20^{th} century illustrate, there is great explanatory gain to developing models of mechanisms that integrate knowledge over several levels of organization. In discussing the more restrictive type of reduction that is achieved through understanding a mechanism, we also noted the need to rethink levels from the rather global perspective embraced in the theory-reduction account to a far more restricted sense in which the constituents of a given level are only determined as one takes a mechanism apart and establishes its working parts. Further, we noted that not all integration in mechanistic explanations is reductive — sometimes claims linking two characterizations of the same entity (e.g., a functional and a structural account) play an important heuristic role in fostering the development of science.

The kind of knowledge that results when investigators focus on mechanism is illustrated in the developmental systems account of how genetic information is linked to knowledge of biological traits — it is linked via an understanding of genetic regulation that relies on knowledge of the cellular machinery (especially the machinery of protein synthesis) which makes development possible. Our last two brief case studies bring out yet other important aspects of integration: the use of integration to overcome epistemic limitations and advance the epistemic warrant of research techniques and theories in each discipline and the dynamics of the process of interdisciplinary exchange (including boundary breaking as well as boundary bridging endeavors). Although we cannot follow up on these threads here, they point to very promising directions for further philosophical investigations of scientific integration.

ACKNOWLEDGEMENTS

We thank Adele Abrahamsen, Elihu Gerson, and Theo Kuipers for their very helpful comments and suggestions on earlier drafts of this paper.

BIBLIOGRAPHY

[Abrahamsen, 1987] A. A. Abrahamsen. Bridging boundaries versus breaking boundaries: Psycholinguistics in perspective. *Synthese*, 72(3): 355–388, 1987.

[Albert and Barabási, 2002] R. Albert, and A.-L. Barabási. Statistical mechanics of complex networks. *Review of Modern Physics*, 74: 47–97, 2002.

[Barabási and Albert, 1999] A.-L. Barabási, and R. Albert. Emergence of scaling in random networks. *Science*, 286: 509–512, 1999.

[Bechtel, 1984] W. Bechtel. Reconceptualization and interfield connections: The discovery of the link between vitamins and coenzymes. *Philosophy of Science*, 51: 265–292, 1984.

[Bechtel, 1986] W. Bechtel. The nature of scientific integration. In W. Bechtel (ed.), *Integrating Scientific Disciplines*. Dordrecht: Martinus Nijhoff, pages 3–52, 1986.

[Bechtel, 2001] W. Bechtel. Decomposing and localizing vision: An exemplar for cognitive neuroscience. In R. S. Stufflebeam (ed.), *Philosophy and the Neurosciences: A Reader*. Oxford: Basil Blackwell, pages 225–249, 2001.

[Bechtel, 2006] W. Bechtel. *Discovering Cell Mechanisms: The Creation of Modern Cell Biology*. Cambridge: Cambridge University Press, 2006.

[Bechtel and Abrahamsen, 2005] W. Bechtel, and A. Abrahamsen. Explanation: A mechanist alternative. *Studies in History and Philosophy of Biological and Biomedical Sciences*, 36: 421–441, 2005.

[Bechtel and McCauley, 1999] W. Bechtel, and R. N. McCauley. Heuristic identity theory (or back to the future): The mind-body problem against the background of research strategies in cognitive neuroscience. In S. C. Stoness (ed.), *Proceedings of the 21st Annual Meeting of the Cognitive Science Society*. Mahwah, NJ: Lawrence Erlbaum Associates, pages 67–72, 1999.

[Bechtel and Mundale, 1999] W. Bechtel, and J. Mundale. Multiple realizability revisited: Linking cognitive and neural states. *Philosophy of Science*, 66: 175–207, 1999.

[Bechtel and Richardson, 1993] W. Bechtel, and R. C. Richardson. *Discovering Complexity: Decomposition and Localization as Strategies in Scientific Research*. Princeton, NJ: Princeton University Press, 1993.

[Bernard, 1865] C. Bernard. *An Introduction to the Study of Experimental Medicine*. New York: Dover, 1865.

[Bichat, 1805] X. Bichat. *Recherches Physiologiques sur la Vie et la Mort* (3rd ed.). Paris: Machant, 1805.

[Bickle, 2003] J. Bickle. *Philosophy and Neuroscience: A Ruthlessly Reductive Account*. Dordrecht: Kluwer, 2003.

[Blumenthal, 1987] A. L. Blumenthal. The emergence of psycholinguistics. *Synthese*, 72(3), 313–323, 1987.

[Callender, 1999] C. A. Callender. Reducing statistical mechanics to thermodynamics: The case of entropy. *The Journal of Philosophy*, 96: 348–373, 1999.

[Campbell, 1974] D. T. Campbell. 'Downward causation' in hierarchically organised biological systems. In Dobzhansky (ed.), *Studies in the Philosophy of Biology*, Macmillan Press Ltd., 1974.

[Cannon, 1929] W. B. Cannon. Organization of physiological homeostasis. *Physiological Reviews*, 9: 399–431, 1929.

[Carnap, 1928] R. Carnap. *Der logische Aufbau der Welt*. Berlin: Weltkreis, 1928.

[Cartwright, 1980] N. Cartwright. Do the laws of physics state the facts? *Pacific Philosophical Quarterly*, 61: 64–75, 1980.

[Cartwright, 1983] N. Cartwright. *How the Laws of Physics Lie*. Oxford: Oxford University Press, 1983.

[Cartwright, 1999] N. Cartwright. *The Dappled World: A Study of the Boundaries of Science*. Cambridge: Cambridge University Press, 1999.

[Causey, 1977] R. L. Causey. *Unity of Science*. Dordrecht: Reidel, 1977.

[Chalmers, 1996] D. Chalmers. *The Conscious Mind*. Oxford: Oxford University Press, 1996.

[Chomsky, 1957] N. Chomsky. *Syntactic Structures*. The Hague: Mouton, 1957.

[Chomsky, 1959] N. Chomsky. Review of *Verbal Behavior*. *Language*, 35: 26–58, 1959.

[Chomsky, 1965] N. Chomsky. *Aspects of a Theory of Syntax*. Cambridge, MA: MIT Press, 1965.

[Chomsky, 1966] N. Chomsky. *Cartesian Linguistics: A Chapter in the History of Rationalist Thought*. Cambridge, MA: MIT Press, 1966.

[Chomsky, 1986] N. Chomsky. *Knowledge of Language: Its Nature, Origin, and Use*. New York: Praeger, 1986.

[Chubin, 1982] D. E. Chubin. *Sociology of Sciences: An Annotated Bibliography on Invisible Colleges, 1972-1981*. New York: Garland, 1982.

[Churchland, 1981] P. M. Churchland. Eliminative materialism and propositional attitudes. *The Journal of Philosophy*, 78: 67–90, 1981.

[Churchland, 1986] P. S. Churchland. *Neurophilosophy: Toward a Unified Science of the Mind-Brain*. Cambridge, MA: MIT Press/Bradford Books, 1986.

[Crane, 1972] D. Crane. *Invisible Colleges*. Chicago: University of Chicago Press, 1972.

[Craver and Bechtel, in press] C. Craver and W. Bechtel. Top-down causation without top-down causes. *Biology and Philosophy*, in press.

[Craver and Darden, 2001] C. Craver and L. Darden. Discovering mechanisms in neurobiology: The case of spatial memory. In P. McLaughlin (ed.), *Theory and Method in Neuroscience*. Pittsburgh, PA: University of Pittsburgh Press, pages 112–137, 2001.

[Crick and Koch, 1998] F. Crick and C. Koch. Consciousness and neuroscience. *Cerebral Cortex*, 8: 97–107, 1998.

[Dale and Pourquié, 2000] J. K. Dale and O. Pourquié. A clock-work somite. *Bioassays*, 22: 72–83, 2000.

[Darden, 1986] L. Darden. Relations amongst fields in the evolutionary synthesis. In W. Bechtel (ed.), *Integrating Scientific Disciplines*. Dordrecht: Martinus Nijhoff, pages 113–123, 1986.

[Darden and Craver, 2002] L. Darden and C. Craver. Strategies in the interfield discovery of the mechanism of protein synthesis. *Studies in the History and Philosophy of the Biological and Biomedical Sciences*, 33: 1–28, 2002.

[Darden and Maull, 1977] L. Darden and N. Maull. Interfield theories. *Philosophy of Science*, 43: 44–64, 1977.

[Dawkins, 1976] R. Dawkins. *The Selfish Gene*. Oxford: Oxford University Press, 1976.

[de Duve, 1984] C. de Duve. *A Guided Tour of the Living Cell*. New York: Scientific American Library, 1984.

[Dupré, 1983] J. Dupré. The disunity of science. *Mind*, 92: 321–346, 1983.

[Dupré, 1993] J. Dupré. *The Disorder of Things: Metaphysical Foundations of the Disunity of Science*. Cambridge, MA: Harvard University Press, 1993.

[Feigl, 1958/1967] H. Feigl. *The 'Mental' and the 'Physical': The Essay and a Postscript*. Minneapolis: University of Minnesota Press, 1958/1967.

[Feyerabend, 1962] P. K. Feyerabend. Explanation, reduction, and empiricism. In G. Maxwell (ed.), *Minnesota Studies in the Philosophy of Science*. Minneapolis, MN: University of Minnesota Press, Vol. III, pages 28–97, 1962.

[Feyerabend, 1963] P. K. Feyerabend. Mental events and the brain. *The Journal of Philosophy*, 60: 295–296, 1963.

[Feyerabend, 1970] P. K. Feyerabend. Against method: Outline of an anarchistic theory of knowledge. In M. R. a. S. Winokur (ed.), *Minnesota Studies in the Philosophy of Science*. Minneapolis, MN: University of Minnesota Press, Volume IV, pages 17–130, 1970.

[Feyerabend, 1975] P. K. Feyerabend. *Against method*. London: New Left Books, 1975.

[Feynman, 1963] R. P. Feynman. *The Feynman Lectures on Physics*. Reading, MA: Addison-Wesley Publishing Compan, 1963.

[Fodor, 1974] J. A. Fodor. Special sciences (or: the disunity of science as a working hypothesis). *Synthese, 28*, 97-115, 1974.

[Friedman, 1974] M. Friedman. Explanation and scientific understanding. *Journal of Philosophy*, 71: 5–19, 1974.

[Gánti, 1975] T. Gánti. Organization of chemical reactions into dividing and metabolizing units: The chemotons. *BioSystems*, 7: 15–21, 1975.

[Gánti, 2003] T. Gánti. *The Principles of Life*. New York: Oxford, 2003.

[Ghiselin, 2004] M. T. Ghiselin. Lorenz Oken and In T. Bach and O. Breidbach (eds.), *Naturphilosophie nach Schelling*. Stuttgart: Frommann-Holzboog, 2004, pages 433–457.

[Ghiselin and Breidbach, 2002] M. T. Ghiselin and O. Breidbach. Lorenz Oken and *Naturphilosophie* in Jena, Paris, and London. *History and Philosophy of the Life Sciences*, 24: 219–247, 2002.

[Glennan, 1996] S. Glennan. Mechanisms and the nature of causation. *Erkenntnis*, 44: 50–71, 1996.

[Glennan, 2002] S. Glennan. Rethinking mechanistic explanation. *Philosophy of Science*, 69: S342–S353, 2002.

[Gong and van Leeuwen, 2003] P. Gong and C. van Leeuwen. Emergence of a scale-free network with chaotic units. *Physical A: Statistical Mechanics and its Applications*, 321: 679–688, 2003.

[Goodman, 1955] N. Goodman. *Fact, Fiction, and Forecast*. Cambridge, MA: Harvard University Press, 1955.

[Hacking, 1983] I. Hacking. *Representing and Intervening*. Cambridge: Cambridge University Press, 1983.

[Hamer, 2004] D. Hamer. *The God Gene*. New York: Doubleday, 2004.

[Hempel, 1965] C. G. Hempel. Aspects of scientific explanation. In C. G. Hempel (ed.), *Aspects of Scientific Explanation and other Essays in the Philosophy of Science*. New York: Macmillan, pages 331–496, 1965.

[Hempel and Oppenheim, 1948] C. G. Hempel and P. Oppenheim. Studies in the logic of explanation. *Philosophy of Science*, 15: 137–175, 1948.

[Hennig, 1966] W. Hennig. *Phylogenetic Systematics*, (R. Zangerl, Trans.). Urbana: University of Illinois Press, 1966.

[Hooker, 1981] C. A. Hooker. Towards a general theory of reduction. *Dialogue*, 20: 38–59; 201–236; 496–529, 1981.

[Hull, 1972] D. L. Hull. Reduction in genetics — Biology or philosophy? *Philosophy of Science*, 39: 491–499, 1972.

[Hull, 1974] D. L. Hull. *The Philosophy of Biological Science*. Englewood Cliffs, NJ: Prentice-Hall, 1974.

[Kaufmann, 1993] S. A. Kaufmann. *The Origins of Order*. Oxford: Oxford University Press, 1993.

[Keijzer, 2001] F. Keijzer. *Representation and Behavior*. Cambridge, MA: MIT Press, 2001.

[Kelso, 1995] J. A. S. Kelso. *Dynamic Patterns: The Self Organization of Brain and Behavior*. Cambridge, MA: MIT Press, 1995.

[Kemeny and Oppenheim, 1956] J. G. Kemeny and P. Oppenheim. On reduction. *Philosophical Studies*, 7: 6–19, 1956.

[Kim, 1998] J. Kim. *Mind in a Physical World*. Cambridge, MA: MIT Press, 1998.

[Kitcher, 1981] P. Kitcher. Explanatory unification. *Philosophy of Science*, 48: 507–531, 1981.

[Kitcher, 1989] P. Kitcher. Explanatory unification and the causal structure of the world. In W. C. Salmon (ed.), *Scientific Explanation*. Minneapolis, MN: University of Minnesota Press, Vol. XIII, pages 410–505, 1989.

[Kitcher, 1999] P. Kitcher. Unification as a regulative ideal. *Perspectives on Science*, 7: 337–348, 1999.

[Krebs and Johnson, 1937] H. A. Krebs and W. A. Johnson. The role of citric acid in intermediate metabolism in animal tissues. *Enzymologia*, 4: 148–156, 1937.

[Kuhn, 1962/1970] T. S. Kuhn. *The Structure of Scientific Revolutions*, (Second ed.), Chicago: University of Chicago Press, 1962/1970.

[Kuipers, 2001] T. A. F. Kuipers. *Structures in Science*. Dordrecht: Kluwer, 2001.

[Landau, 1944] L. Landau. On the problem of turbulence. *Comptes Rendus d'Academie des Sciences, URSS*, 44: 311–314, 1944.

[Lewontin, 2000] R. Lewontin. *It ain't necessarily so: The Dream of the Human Genome and other Illusions*. New York: Basic Books, 2000.

[Machamer *et al.*, 2000] P. Machamer, L. Darden, and C. Craver. Thinking about mechanisms. *Philosophy of Science*, 67: 1–25, 2000.

[Maturana and Varela, 1980] H. R. Maturana and F. J. Varela. Autopoiesis: The organization of the living. In F. J. Varela (ed.), *Autopoiesis and Cognition: The Realization of the Living*. Dordrecht: D. Reidel, pages 59–138, 1980.

[Mayr, 1970] O. Mayr. *The Origins of Feedback Control*. Cambridge, MA: MIT Press, 1970.

[McCauley, 1981] R. N. McCauley. Hypothetical identities and ontological economizing: Comments on Causey's program for the unity of science. *Philosophy of Science*, 48: 218–227, 1981.

[McCauley, 1986] R. N. McCauley. Intertheoretic relations and the future of psychology. *Philosophy of Science*, 53: 179–199, 1986.

[McCauley, 1987] R. N. McCauley. The not so happy story of the marriage of linguistics and psychology: or why linguistics has discouraged psychology's recent advances. *Synthese*, 72: 341–353, 1987.

[McCauley, 1996] R. N. McCauley. Explanatory pluralism and the coevolution of theories in science. In R. N. McCauley (ed.), *The Churchlands and their Critics*. Oxford: Blackwell, pages 17–47, 1996.

[McCauley and Bechtel, 2001] R. N. McCauley and W. Bechtel. Explanatory pluralism and heuristic identity theory. *Theory and Psychology*, 11(6): 736–760, 2001.

[Milgram, 1967] S. Milgram. The small world problem. *Psychology Today*, 2: 60–67, 1967.

[Miller, 1962] G. A. Miller. Some psychological studies of grammar. *American Psychologist*, 17: 748–762, 1962.

[Nagel, 1961] E. Nagel. *The Structure of Science*. New York: Harcourt, Brace, 1961.

[Neurath, 1938] O. Neurath. Unified science as encyclopedic integration. In C. Morris (ed.), *International Encyclopedia of Unified Science*, Vol. I, Chicago: University of Chicago Press, 1938.

[Nickles, 1973] T. Nickles. Two concepts of intertheoretic reduction. *The Journal of Philosophy*, 70: 181–201, 1973.

[Oken, 1809] L. Oken. *Lehrbuch der Naturphilosophie*. Jena: Friedrich Frommann, 1809.

[Oken, 1831] L. Oken. *Lehrbuch der Naturphilosophie* (2nd ed.). Jena: Friedrich Frommann, 1831.

[Oppenheim and Putnam, 1958] P. Oppenheim and H. Putnam. The unity of science as a working hypothesis. In G. Maxwell (ed.), *Concepts, Theories, and the Mind-Body Problem*. Minneapolis: University of Minnesota Press, pages 3–36, 1958.

[Oyama *et al.*, 2001] S. Oyama, P. E. Griffiths, and R. Gray. What is developmental systems theory? In R. Gray (ed.), *Cycles of Contingency*. Cambridge, MA: MIT Press, 2001.

[Paul, 1880] H. Paul. *Principien der Sprachgeschichte*. Halle: Niemeyer, 1880.

[Place, 1956] U. T. Place. Is consciousness a brain process. *British Journal of Psychology*, 47: 44–50, 1956.

[Polger, 2004] T. Polger. *Natural Minds*. Cambridge, MA: MIT Press, 2004.

[Port and van Gelder, 1995] R. Port and T. van Gelder. *It's about Time*. Cambridge, MA: MIT Press, 1995.

[Pourquié, 1998] O. Pourquié. Clocks regulating developmental processes. *Current Opinion in Neurobiology*, 8: 665–670, 1998.

[Price, 1961] D. J. D. S. Price. *Science since Babylon*. New Haven: Yale University Press, 1961.

[Putnam, 1978] H. Putnam. *Meaning and the Moral Sciences*. London: Routledge and Kegan Paul, 1978.

[Pylyshyn, 1984] Z. W. Pylyshyn. *Computation and Cognition: Toward a Foundation for Cognitive Science*. Cambridge, MA: MIT Press, 1984.

[Quine, 1964] W. v. O. Quine. Ontological reduction and the world of numbers. *Journal of Philosophy*, 61: 209–216, 1964.

[Reber, 1987] A. S. Reber. The rise and (surprisingly rapid) fall of psycholinguistics. *Synthese*, 72(3): 325–339, 1987.

[Richardson, 1979] R. C. Richardson. Functionalism and reductionism. *Philosophy of Science*, 46: 533–558, 1979.

[Rorty, 1970] R. Rorty. In defense of eliminative materialism. *The Review of Metaphysics*, 24: 112–121, 1970.

[Rosenberg, 1994] A. Rosenberg. *Instrumental Biology and the Disunity of Science*. Chicago: University of Chicago Press, 1994.

[Rosenblueth *et al.*, 1943] A. Rosenblueth, N. Wiener, and J. Bigelow. Behavior, purpose, and teleology. *Philosophy of Science*, 10: 18–24, 1943.

[Ruse and Wilson, 1986] M. Ruse and E. O. Wilson. Moral philosophy as applied science. *Philosophy: The Journal of the Royal Institute of Philosophy*, 61: 173–192, 1986.

[Schaffner, 1967] K. Schaffner. Approaches to reduction. *Philosophy of Science*, 34: 137–147, 1967.

[Schaffner, 1969] K. F. Schaffner. The Watson-Crick model and reductionism. *British Journal for the Philosophy of Science*, 20: 325–348, 1969.

[Schuyler, 1978] R. L. Schuyler (ed.). *Historical Archaeology: A Guide to Substantive and Theoretical Contributions*. Farmingdale, NY: Baywood Publishing Company, 1978.

[Shapere, 1974] D. Shapere. Scientific theories and their domains. In F. Suppe (ed.), *The Structure of Scientific Theories*. Urbana: University of Illinois Press, 1974.

[Shapiro, 2004] L. Shapiro. *The Mind Incarnate*. Cambridge, MA: MIT Press, 2004.

[Sklar, 1967] L. Sklar. Types of inter-theoretic reduction. *British Journal for the Philosophy of Science*, 18: 109–124, 1967.

[Sklar, 1974] L. Sklar. Thermodynamics, statistical mechanics, and the complexity of reductions. In J. van Evra (ed.), *PSA 1974*. Dordrecht: Reidel, Vol. 32 of Boston Studies in the Philosophy of Science, pages 15–32, 1974.

[Sklar, 1993] L. Sklar. *Physics and Chance: Philosophical Issues in the Foundations of Statistical Mechanics*. New York: Cambridge, 1993.

[Smart, 1959] J. J. C. Smart. Sensations and brain processes. *Philosophical Review*, 68: 141–156, 1959.

[Smith, 2003] K. K. Smith. Time's arrow: Heterochrony and the evolution of development. *International Journal of Developmental Biology*, 47: 613–621, 2003.

[Stotz and Bostanci, 2005] K. C. Stotz and A. Bostanci. The representing genes project: Tracking the shift to "post-genomics". *New Genetics and Society*, 2005.

[Suppes, 1957] P. Suppes. *Introduction to Logic*. Princeton: van Nostrand, 1957.

[Suppes, 1981] P. Suppes. The plurality of science. In I. Hacking (ed.), *PSA 1978*. East Lansing, MI: Philosophy of Science Association, Vol. 2, pages 2–16, 1981.

[Tautz, 2000] D. Tautz. Evolution of transcriptional regulation. *Current Opinion in Genetic Development*, 10: 575–579, 2000.
[Taylor, 1967] C. Taylor. Mind-body identity, a side issue. *Philosophical Review*, 67: 201–213, 1967.
[Thelen and Smith, 1994] E. Thelen and L. Smith. *A Dynamical Systems Approach to the Development of Cognition and Action*. Cambridge, MA: MIT Press, 1994.
[von Bertalanfy, 1951] L. von Bertalanfy. *General Systems Theory: A new Approach to the Unity of Science*. Baltimore: Johns Hopkins Press, 1951.
[Wasserman and Faust, 1994] S. Wasserman and K. Faust. *Social Network Analysis: Methods and Applications*. New York: Cambridge, 1994.
[Watson and Crick, 1953] J. D. Watson and F. H. C. Crick. Molecular structure of nucleic acids. *Nature*, 171: 737–738, 1953.
[Watts and Strogratz, 1998] D. Watts and S. Strogratz. Collective dynamics of small worlds. *Nature*, 393: 440–442, 1998.
[Whaley, 1975] W. G. Whaley. *The Golgi Apparatus*, Vol. 2. New York: Springer-Verlag, 1975.
[Whewell, 1840] W. Whewell. *The Philosophy of the Inductive Sciences, founded upon their History*. London: J. W. Parker, 1840.
[Wiener, 1948] N. Wiener. *Cybernetics: Or, Control and Communication in the Animal Machine*. New York: Wiley, 1948.
[Wilson, 1985] M. Wilson. What is this thing called 'pain'? *Pacific Philosophical Quarterly*, 66: 227–267, 1985.
[Wimsatt, 1975] W. C. Wimsatt. Reductionism, levels of organization, and the mind-body problem. In I. Savodnik (ed.), *Brain and Consciousness*. New York: Plenum, pages 205–267, 1975.
[Wimsatt, 1976a] W. C. Wimsatt. Reductionism, levels of organization, and the mind-body problem. In I. Savodnik (ed.), *Consciousness and the Brain: A Scientific and Philosophical Inquiry*. New York: Plenum Press, pages 202–267, 1976.
[Wimsatt, 1976b] W. C. Wimsatt. Reductive explanation: A functional account. In J. van Evra (ed.), *PSA-1974*. Dordrecht: Reidel, pages 671–710, 1976.
[Wimsatt, 1994] W. C. Wimsatt. The ontology of complex systems: Levels, perspectives, and causal thickets. *Canadian Journal of Philosophy*, Supplemental Volume 20: 207–274, 1994.
[Woodger, 1952] J. H. Woodger. *Biology and Language*. Cambridge: Cambridge University Press, 1952.
[Wray and Love, 2000] G. Wray and C. Love. Developmental regulatory genes and echinoderm evolution. *Systematic Biology*, 49: 28–51, 2000.
[Wundt, 1900] W. Wundt. *Die Sprache*. Leipzig: Englemann, 1900.
[Wylie, 1999] A. Wylie. Rethinking unity as a "Working Hypothesis" for philosophy of science: How archaeologists exploit the disunities of science. *Perspectives on Science*, 7 (3): 293–317, 1999.

LOGICAL, HISTORICAL AND COMPUTATIONAL APPROACHES

Atocha Aliseda and Donald Gillies

1 INTRODUCTION

This chapter is concerned with logical, historical and computational approaches to the philosophy of science. We will deal with these various approaches in the historical order in which they were developed starting around 1920. In section 2 we will discuss how the logical approach to philosophy of science was introduced by the Vienna Circle, and developed by them and their followers and associates. The logical approach to philosophy of science remained dominant in the subject throughout the 1950s; but, from the early 1960's, it was challenged by a striking development of the historical approach. The historical approach was not introduced for the first time in the 1960s. On the contrary, it had been developed by Mach and Duhem much earlier, but, although Mach and Duhem are cited by the Vienna Circle as important influences on their philosophy, the Vienna Circle did not adopt the historical features of these two thinkers. In the excitement generated by the new logic of Frege and Russell, history of science seems to have been temporarily forgotten. But although the general idea of the historical approach is not new in the 1960s, that decade saw striking developments in this approach. After Kuhn, the analysis of scientific revolutions became a major problem for philosophy of science, while Lakatos applied the historical approach to mathematics for the first time. In section 3 we will give an account of the development of the historical approach from the early 1960s to the mid-1970s. After this time, a new factor enters the picture which seems to be changing society in many profound ways, and, in particular, is bringing about striking new developments in philosophy of science. This new factor is of course the development of computers. In section 4, we will show how various factors, including research in artificial intelligence (AI) led to the conception of science as problem solving in the 1970s. Then in section 5 we will trace the emergence in the 1980s and 1990s of logical and computational models for scientific inference and discovery. During this period, work in computer science brought about considerable developments in logic, leading to the introduction of new systems of logic, such as non-monotonic logic and abduction, which were unknown to the Vienna Circle. The new results have brought in to question earlier positions in the philosophy of science. For example some successes in the field of machine learning have provided a strong argument against Popper's claim

Handbook of the Philosophy of Science: General Philosophy of Science - Focal Issues, pp. 431–513.
Volume editor: Theo Kuipers. General editors: Dov M. Gabbay, Paul Thagard and John Woods.

that 'induction is a myth'. In general terms the new computer-based approaches allow the investigation in a formal way of the problem of discovery in science and so perhaps can be considered as closing the gap between the logical and historical approaches, which, up to the mid-1970s, tended to be seen as antagonistic.

2 THE LOGICAL APPROACH OF THE VIENNA CIRCLE AND THEIR FOLLOWERS FROM THE 1920S TO THE 1950S

In 1922 Moritz Schlick was appointed to the Mach-Boltzmann professorship of the inductive sciences at the University of Vienna. His arrival in Vienna that year marks the beginning of the Vienna Circle which was indeed known initially as the Schlick Circle.[1] Schlick organised a seminar which met once a fortnight on Thursday evenings in a room of the university building that housed the mathematics and physics institutes. Attendance at this seminar was by invitation only, and those who attended were the members of the Vienna Circle. They included Rudolf Carnap, Kurt Gödel, Hans Hahn, Otto Neurath, and Moritz Schlick himself. The building where they met has subsequently been restructured, but a room similar to and near their original meeting place has been turned into a kind of museum, and its walls are hung with photographs of the famous circle and their associates.

The views of the Vienna Circle in 1929 are set out in a pamphlet written mainly by Otto Neurath but with the help of some other members of the circle. In this work Neurath et al state very clearly the philosophical methodology which the circle employed. They write:

> The task of philosophical work lies in this clarification of problems and assertions, ... The method of this clarification is that of *logical analysis*; of it, Russell says (*Our Knowledge of the External World*, p. 4) that it "has gradually crept into philosophy through the critical scrutiny of mathematics ..."
>
> It is *the method of logical analysis* that essentially distinguishes recent empiricism and positivism from the earlier version that was more biological-psychological in its orientation. [1929, 8]

Here Neurath et al not only indicate their method (logical analysis), but also one of the main sources of their approach (Russell). Russell was certainly a major influence on the Vienna circle. In his memoir of Hahn, Menger, another member of the Vienna circle, writes: 'During the early 1920s he developed a great admiration for the works of Bertrand Russell. He reviewed some of them in the *Monatshefte für Mathematik und Physik*. In one of these reviews Hahn suggested that one day Russell might well be regarded as the most important philosopher of his time'

[1]There is now in Vienna a Vienna Circle Institute, directed by Friedrich Stadler, which publishes important works on the history of the Vienna Circle. Its website is `www.univie.ac.at/ivc`. Stadler [2001] gives an excellent scholarly and detailed account of the Vienna Circle. In this chapter we have also used the memoirs contained in Frank [1941] and Gadol [1982].

[Menger, 1980, xi]. Hahn also conducted a seminar on Russell and Whitehead's *Principia Mathematica* in the academic year 1924-5 during which the participants went through that work chapter by chapter.

It was perhaps mainly through Russell that the Vienna circle acquired its knowledge and love of logic, but there were other influences as well. Carnap records in his autobiography [1963, 5-6] that he studied under Frege. Carnap's involvement with Frege appears to have been rather by chance. His family lived in Jena and Carnap went to the University of Jena where Frege, although past 60, was only an associate professor of mathematics. Carnap writes:

> In the fall of 1910, I attended Frege's course "Begriffsschrift" (conceptual notation, ideography), out of curiosity, not knowing anything either of the man or the subject except for a friend's remark that somebody had found it interesting. [1963, 5]

Carnap himself was sufficiently interested to attend Frege's advanced course "Begriffsschrift II" in 1913 and his course Logik in der Mathematik in 1914. Carnap records that "Begriffsschrift II" was attended by 3 students: Carnap, a friend of Carnap's and a retired army major. One feels that in a modern 'efficiency'-minded university, Frege would have been forced into early retirement long before Carnap attended his courses. Who could have guessed at that time that Frege would later be lauded as the most important philosopher of his time?

Another important influence on the Vienna circle was Wittgenstein. The circle devoted itself to reading Wittgenstein's *Tractatus Logico-Philosophicus,* which had been published in 1921, 'paragraph by paragraph' during the academic year 1926-7 [Menger, 1980, xii]. Wittgenstein's *Tractatus* is of course a notable example of the application of the logic of Frege and Russell to philosophy.

Frege and Russell themselves had devoted their time much more to philosophy of mathematics than to philosophy of science. Frege had formulated the logicist thesis that mathematics could be reduced to logic, and had tried to demonstrate the correctness of this thesis — inventing a new system of logic in the process. Unfortunately Frege's system proved to be contradictory, but Russell tried to overcome this difficulty and to establish a consistent form of logicism. As Russell saw it, he had successfully applied the method of logical analysis and the new logic to the philosophy of mathematics, and he came to think that logic and logical analysis might be the essence of all philosophy. He expressed this point of view in a series of lectures which he gave at Harvard in March and April 1914, and which were published later that year with the title: *Our Knowledge of the External World.* In the preface, Russell wrote:

> The following lectures are an attempt to show, by means of examples, the nature, capacity, and limitations of the logical-analytic method in philosophy. This method, of which the first complete example is to be found in the writings of Frege, has gradually, in the course of actual research, increasingly forced itself upon me as something perfectly definite, capable of embodiment in maxims, and adequate, in all branches

> of philosophy, to yield whatever objective scientific knowledge it is possible to obtain. [1914, 7]

Lecture II was entitled: 'Logic as the Essence of Philosophy', and in it Russell said:

> The topics we discussed in our first lecture, and the topics we shall discuss later, all reduce themselves, in so far as they are genuinely philosophical, to problems of logic. This is not due to any accident, but to the fact that every philosophical problem, when it is subjected to the necessary analysis and purification, is found either to be not really philosophical at all, or else to be, in the sense in which we are using the word, logical. [1914, 42]

Russell's words had a profound impact on Carnap who records in his autobiography [1963, 13] that he read Russell's book, *Our Knowledge of the External World* in the winter of 1921. Carnap quotes a passage from this book which begins:

> The study of logic becomes the central study in philosophy: it gives the method of research in philosophy, just as mathematics gives the method in physics. [Russell, 1914, 243]

On this Carnap himself comments:

> I felt as if this appeal had been directed to me personally. To work in this spirit would be my task from now on! And indeed henceforth the application of the new logical instrument for the purposes of analyzing scientific concepts and of clarifying philosophical problems has been the essential aim of my philosophical activity. [1963, 13]

The Vienna Circle did continue the investigations of Frege and Russell into the philosophy of mathematics, but their main originality lay in the application of the method of logical analysis to the philosophy of science. This is the origin of the logical approach to the philosophy of science, which can be seen as an application to the philosophy of science of ideas originating in the philosophy of mathematics.[2] Neurath *et al.* formulate this aspect of the Vienna Circle's approach very clearly as follows:

> As we have specially considered with respect to physics and mathematics, every branch of science is led to recognise that, sooner or later in its development, it must conduct an epistemological examination of its foundations, a logical analysis of its concepts. [1929, 17]

[2]More details about this are to be found in Gillies and Zheng [2001], which also discusses the converse influence of the philosophy of science on the philosophy of mathematics. See particularly section 3, pp. 439-445.

Thus philosophers of science have the task of giving a logical analysis of the concepts of science. This essentially is what the logical approach to the philosophy of science amounts to.

The interest of the Vienna Circle in science is easily explained by the historical period in which they were living. The years 1900-30 were those of a great revolution in physics, which called into question the Newtonian mechanics which had been accepted for nearly two centuries, and gave birth to the new theories of relativity and quantum mechanics. There were personal contacts between members of the Vienna Circle and the leading physicists of the time. Schlick was a good friend of Einstein's, and the two of them engaged in a considerable correspondence concerning the philosophical interpretation of relativity. Neurath *et al.* list [1929, 20] three '*Leading representatives of the scientific world-conception*'. These are 'Albert Einstein, Bertrand Russell, Ludwig Wittgenstein.' This choice gives an excellent insight into the attitudes and interests of the Circle.

Let us now look at the logic which was used by the Vienna Circle for their logical analysis. Most of the Circle were inductivists and accepted inductive as well as deductive logic. Deductive logic was for nearly all of them the formal logic of Frege and Russell, together with the additions of later logicians such as Tarski. This was recently invented and had clearly superseded Aristotelian logic. It seemed at the time a powerful new tool for carrying out philosophical investigations. Fregean logic, or classical logic as it is now usually called, had at the time of the Vienna Circle only one opponent — the intuitionistic logic of Brouwer. Although several of the Vienna Circle did take an interest in Brouwer and intuitionism, it is fair to say that for them classical logic remained fundamental.

Turning now to inductive logic, a major problem for the circle was how evidence supported or confirmed scientific hypotheses. Thus investigations of confirmation and connected questions to do with probability were an important area of the Circle's activity. Confirmation theory was generally regarded as constituting an inductive logic which complemented deductive logic.

Nearly all the Vienna Circle accepted the distinction clearly formulated by Reichenbach between the *discovery* of scientific hypotheses and their *justification.* Moreover it was generally felt that discovery involved a creative act of a psychological and perhaps irrational nature. Thus the investigation of discovery fell outside the remit of philosophy of science, since philosophy of science consisted of logical analysis and no logical analysis could be given of discovery. Philosophers of science should therefore concentrate on giving a logical analysis of how scientific hypotheses are justified by evidence. Such a logical analysis constitutes confirmation theory or inductive logic. It is worth noting that this attitude to induction constitutes a considerable divergence from that of Bacon who regarded induction as principally a method of discovery.

The Vienna Circle's existence in Vienna was in fact quite short. Fascists seized power in Austria in 1934, and in 1936 Schlick was shot dead by a deranged student in the University of Vienna. After these events nearly all the other members of the Vienna Circle fled from Vienna — for the most part settling in the English-speaking world. This diaspora of the Vienna Circle had the effect of increasing rather than decreasing the influence of their ideas. In the period after the Second World War, the English-speaking world became the main centre for philosophy of science and the logical approach of the Vienna Circle came to dominate philosophy of science and perhaps the whole of philosophy as well. The high point of the influence of the logical approach may be the 1950s.

Two books which give an excellent illustration of the logical approach to the philosophy of science in this period are (1) Carnap's *Logical Foundations of Probability* which was published in 1950, but with a second edition in 1963, and (2) Hempel's *Aspects of Scientific Explanation and other Essays in the Philosophy of Science*, which was published in 1965, but is a collection of Hempel's essays written between 1942 and 1964. Carnap's book is a lengthy investigation of the central Vienna Circle topic of confirmation theory and inductive logic. Hempel was a younger follower of the Vienna Circle rather than an original member, but he became one of the leading advocates of their approach in the United States. His book contains an investigation of the logic of confirmation, but also of the logic of explanation, the logic of functional analysis, a logical appraisal of operationism, and so on.

Let us now consider two philosophers who might be considered as associates of the Vienna Circle because they had close links with the Circle without actually being members or followers. The first of these is Popper. Popper would certainly not have considered himself a supporter of the Vienna Circle since he criticized many of their ideas in very harsh terms. For example he was opposed to inductivism and even claimed that induction was a myth, while, conversely, he defended metaphysics against the Vienna Circle's claim that metaphysical statements were always meaningless. However, despite these differences, it remains true that Popper in his book of 1934 does adopt the Vienna Circle's logical approach to philosophy of science. He even, despite his dislike of inductive logic, develops a theory of corroboration in which corroboration, though not a probability function, can be defined in terms of probability. Moreover he supports the view of the Circle that discovery lies outside the sphere of philosophy of science. In fact Popper writes:

> ... the work of the scientist consists in putting forward and testing theories.
>
> The initial stage, the act of conceiving or inventing a theory, seems to me neither to call for logical analysis nor to be susceptible of it. The question how it happens that a new idea occurs to a man — whether it be a musical theme, a dramatic conflict, or a scientific theory — may be of great interest to empirical psychology; but it is irrelevant to the logical analysis of scientific knowledge. This latter is concerned not with *questions of fact* (Kant's *quid facti*?), but only with questions of *justification or validity* (Kant's *quid juris?*). Its questions are of the following kind. Can a statement be justified? And if so, how? Is it

> testable? Is it logically dependent on certain other statements? Or does it perhaps contradict them? [1934, 31]

There is something odd here since Popper in this passage explicitly denies that there can be a logic of scientific discovery. However, his book is actually entitled: *The Logic of Scientific Discovery.* Aliseda has thrown some light on this question in her [2004]. In fact the original German title of Popper's book was *Logik der Forschung*, whose literal translation into English would be ' Logic of Research'. But why did Popper, who was a master of the English language in his later period and who carefully and meticulously supervised the publication of his books in English, allow a mistranslation of his original title? The explanation we would like to propose is that Popper's approach to philosophy of science had changed between 1934 and 1959. As we will argue in more detail in section 3, he had moved, by 1959, towards a more historical approach to philosophy of science, and this made him more sympathetic to including the question of discovery within philosophy of science.

Another associate of the Vienna Circle was Quine who, as a young man, went from Harvard to Vienna for a period to study with the Circle. Like Popper, Quine criticized many of the key ideas of the Vienna Circle. In particular in his famous article: 'Two Dogmas of Empiricism', originally published in 1951 and reprinted in [Quine, 1953, 20-46], he criticizes the analytic-synthetic distinction and argues for a holistic view of science. However, like Popper in 1934, Quine, despite his criticisms of the Vienna Circle, accepts their general approach to philosophy. In fact this is shown very clearly in the title of his 1953 book: *From a Logical Point of View.* Quine's position as professor at Harvard and leading figure of the American philosophical establishment is indicative of the dominance in the United States at that time of the view of philosophy as logical analysis. Quine, however, unlike Popper, never moved towards the historical approach to philosophy, but remained faithful to the logical approach to the end of his life. Despite his admiration for logic, however, he did not take any notice of the new developments in logic which were brought about by the rise of computing and which we will discuss in section 5. Moreover Quine rarely if ever raised any questions about inductive logic and probability. Logic for him was always classical deductive logic with the occasional mention of intuitionistic logic.

3 THE CHALLENGE OF THE HISTORICAL APPROACH (C. 1960 TO THE MID 1970S)

The last hundred years have been characterised by very rapid change and this applies to philosophy of science as much as to anything else. We have seen that in the 1920s and 1930s, the logical approach to philosophy of science was a radical new idea and its proponents were liable to political persecution and even, in one case, assassination. By the 1950s the picture had completed changed. The logical approach was dominant in the dominant English-speaking world. Its advocates had become part of the establishment and many of them occupied prestigious chairs in the leading universities. The 1960s proved to be a turbulent decade in which establishment ideas were attacked throughout the world. Philosophy of science did not escape this trend, and in fact the dominant logical approach to the subject came to be challenged by the historical approach. The principal figures of this new movement in the subject which lasted roughly from the early 1960s to the mid-1970s were the later Popper, Kuhn, Lakatos and Feyerabend. Let us start by examining the case of Popper.

3.1 Popper's Historical Turn

We have already seen that Popper in his [1934] adopted the logical approach to the philosophy of science, but in his [1963] and [1972] there is a distinct shift to the historical approach. A study of the titles give an indication this change. The original title of Popper's [1934] was 'Logic of Research'. However the full title of his [1963] is 'Conjectures and Refutations. The Growth of Scientific Knowledge' and of his [1972] 'Objective Knowledge. An Evolutionary Approach.' Logic has been replaced by Growth and Evolution. It is not that Popper changed his basic circle of ideas, namely falsifiability, rejection of induction, partial rehabilitation of metaphysics, etc. However these ideas are presented in the later Popper in a context which is more historical than logical. Thus 1934, Section 21, pp. 84-6, is entitled: 'Logical Investigation of Falsifiability', while in Popper's 1963 falsification or refutation becomes part of the model of conjectures and refutations which is supposed to give an account of how scientific knowledge grows. Moreover Popper's later works contain many more allusions to episodes in the history of science, and sometime quite detailed analyses of historical case histories as well. An example of the latter is Popper's discussion of Galileo's erroneous theory of the tides in chapter 4 of 1972.

These rather qualitative observations can be supplemented by counting the number of references to history of science in Popper's works of various periods. We included in the count any scientist or mathematician whose work was carried out more than a third of a century before the publication of the volume in question. Using this criterion, it turns out that Popper's 1934 contains 0.33 references to history of science per page, his 1963 contains 0.71 such references per page, and his 1972 1.08 references per page. Thus the number of references which Popper makes to history of science more than doubles between 1934 and 1963 and increases by more than 50% between 1963 and 1972. The trend is unmistakeable. It is interesting to observe by way of comparison that Hempel's 1965 contains 0.075 references to history of science per page. This is not much more than a tenth of the figure for Popper's 1963. The distinction between the logical and historical approach to the subject shows up very clearly using this simple statistical criterion.[3]

[3]Hempel became a colleague of Kuhn's and a study of Kuhn's work influenced Hempel in the direction of the historical approach after 1965. An interesting account of this phase of Hempel's thought is to be found in Wolters [2003]. On p. 115 of this article, Wolters quotes the following

Let us now examine some of the differences between the logical and historical approaches to the philosophy of science. To begin with, the historical approach broadens the scope of philosophy of science to include questions about how scientific knowledge grows and develops and about how new scientific theories and entities are discovered. As we saw the logical approach to the subject, at least up to the 1950s, excluded questions of discovery as belonging to 'empirical psychology' rather than 'the logical analysis of scientific knowledge' [Popper, 1934, 31]. Some problems within philosophy of science are, however, common to the logical and historical approaches, but, generally speaking they are tackled in a different way in the two approaches. We can illustrate this by considering the problem of the way in which evidence confirms or disconfirms scientific hypotheses. As we have seen, this was a standard problem within the logical approach to the philosophy of science in the 1950s and early 1960s and is tackled by Carnap in his 1950 and Hempel in his 1965. Carnap begins by setting up a formal logical language. If a hypothesis h and evidence e are expressed in this language, then we can consider the degree of confirmation of h given e [$C(h, e)$]. Carnap proposes various c-functions as measures of $C(h, e)$ and tries to evaluate their merits. Hempel is not quite as formalistic as Carnap, but he uses a lot of logic. His most famous result in his 'Studies in the Logic of Confirmation' is of course the paradox of the ravens. It would generally be agreed that the hypothesis that all ravens are black is confirmed by observing a black raven. However, 'all ravens are black' is logically equivalent to 'all non-black things are non-ravens'. The latter should it seems, applying the same principle, be confirmed by observing any non-black thing which is a non-raven. So it looks as if 'all ravens are black' should also be confirmed by observing any non-black thing which is a non-raven. As Hempel says [1965, 15]: 'Consequently, any red pencil, any green leaf, any yellow cow etc., becomes confirming evidence for the hypothesis that all ravens are black.' I think it would be fair to say that literally hundreds of papers have been written trying to solve this paradox which remains to this day a favourite topic of philosophers of science.

We can contrast these logical approaches to the problem of confirmation with a historical approach to the same problem. Central to the historical approach to the philosophy of science is the use of case histories from the history of science. To tackle the question of confirmation such a case history would be selected. It could, for example, be Harvey's discovery of the circulation of the blood. Before Harvey's work, the Galenic theory was believed by the medical community. However there was shift in opinion and Harvey's new theory came to be accepted. The details of this change could be studied to see what bits of evidence Harvey brought forward in favour of his new theory and how convincing these various bits of evidence

---

passage from an interview with Hempel carried out in 1982. The following words of Hempel refer to Kuhn whom he first met in 1963: 'I was very much struck by his ideas. At first I found them strange and I had very great resistance to these ideas, his historicist, pragmatist approach to problems in the methodology of science, but I have changed my mind considerably about this since then.' Despite this change of mind, Hempel never became an important figure in the historical approach to the philosophy of science in the way that he certainly was in the logical approach to the subject.

were considered to be by his peers. Account could also be taken of what evidence was cited against Harvey's new theory by his opponents. In this way it could be seen what evidence was seen as confirming and to what extent, and what evidence was seen as disconfirming and to what extent. It would be hoped that, from such historical studies, general principles governing the confirmation of scientific hypotheses by evidence might be gleaned. Such a historical investigation obviously has a completely different character from the investigations of Carnap and Hempel, and this shows clearly some of the differences between the logical and the historical approaches to the philosophy of science.

Popper is interesting because of the shift in his ideas from a more logical approach in the 1930s to a more historical approach in the 1960s, and we will revisit his ideas in section 4.4 where his approach will be compared to that of Simon. However the person who did the most to promote the historical approach to philosophy of science in the 1960s was not Popper, but his younger contemporary Kuhn. We will now examine some of Kuhn's ideas and also how they were received by other philosophers of science — particularly by members of Popper's school.

3.2 Kuhn and his Critics

Kuhn begins his 1962 *The Structure of Scientific Revolutions* with the following rousing call for a historical approach to the philosophy of science:

> History ... could produce a decisive transformation in the image of science by which we are now possessed. That image has previously been drawn, even by scientists themselves, mainly from the study of finished scientific achievements as these are recorded in the classics and, more recently, in the textbooks from which each new scientific generation learns to practice its trade. ... This essay attempts to show that we have been misled by them in fundamental ways. Its aim is a sketch of the quite different concept of science that can emerge from the historical record of the research activity itself. [1962, 1]

We will now sketch the main ideas of Kuhn's essay. Kuhn's view is that science develops through periods of *normal science* which are characterised by the dominance of a *paradigm*, but which are interrupted by occasional revolutions during which the old paradigm is replaced by a new one. We will illustrate this theory by considering in turn the three scientific revolutions which constitute Kuhn's main examples. These are (i) the Copernican Revolution, (ii) the Chemical Revolution, and (iii) the Einsteinian Revolution.

(i) *The Copernican Revolution.* Kuhn's first book, published in 1957 was entitled: *The Copernican Revolution*, and it was probably this example more than any other which led him to his general model of scientific revolutions. From late Greek times until Copernicus, astronomy was dominated by the Aristotelian-Ptolemaic paradigm. The earth was considered to be stationary at the centre of the universe.

The different movements of sublunary and heavenly bodies were described by Aristotelian mechanics. The astronomer had to describe and predict the movements of the Sun, Moon and planets as accurately as possible, using the Ptolemaic scheme of epicycles. This was the normal science of the time.

Copernicus, however, challenged the dominant paradigm by suggesting that the Earth spun on its axis and moved round the Sun. He worked out this alternative theory in as detailed a mathematical fashion as Ptolemy's. His results were published in his book *De Revolutionibus Orbium Caelestium* in 1543, and this publication inaugurated a revolutionary period during which the old Aristotelian-Ptolemaic paradigm was overthrown and replaced by a new paradigm — first formulated in detail by Newton in *Philosophiae Naturalis Principia Mathematica* [1687].

(ii) *The Chemical Revolution.* The main theme of the chemical revolution was the replacement of the *phlogiston* theory by the *oxygen* theory, though there were many other important changes as well. According to the phlogiston theory, bodies are inflammable if they contain a substance called phlogiston, and this is released when the body burns. It was known that air was needed to support combustion, and the phlogiston theorists explained this by claiming that the air absorbed the phlogiston given off in combustion until it was saturated when combustion ceased. The phlogiston theory was also used to explain the calcination of metals. When a metal is heated in air, in many cases it turns into a powder known as the *calx*, e.g. iron -> rust. Conversely the calx is usually found in ores of the metal, and the metal itself could often be obtained by heating with charcoal. These transformations were explained by postulating that

calx + phlogiston = metal

When we heat a metal, phlogiston is given off, and the calx remains. Conversely when we heat the calx with charcoal, since charcoal is very rich in phlogiston because it burns easily, the phlogiston from the charcoal combines with the calx to give the metal. In the oxygen theory, burning is explained as the combination of the substance with oxygen. On this theory air is needed for combustion because it contains oxygen and combustion ceases when the oxygen is used up. The calx is identified with the oxide of the metal. So turning a metal into its calx by heating in air is explained by the equation

metal + oxygen = metal oxide

Similarly obtaining the metal by heating the calx with charcoal is explained by the equation

metal oxide + carbon = metal + carbon dioxide

The oxygen theory was developed by Lavoisier. At the beginning of his researches in 1772, he was already sceptical of the then dominant phlogiston theory. In the next decade or so, many experimental discoveries concerning gases were made. These discoveries were mainly owing to the English experimental chemists — particularly Priestley and Cavendish. However, these English chemists remained faithful to the phlogiston theory. For example Priestley prepared the gas which we now call: 'oxygen'. He observed that it supported combustion better than ordinary air, and concluded that it must be dephlogisticated air. On the phlogiston theory, air supports combustion by absorbing phlogiston. So air with the phlogiston removed (dephlogisticated air) will absorb more phlogiston and support combustion better. Lavoisier, on the other hand, reinterpreted the results of Priestley and the other English chemists in terms of his new and developing oxygen theory. Lavoisier's new paradigm for chemistry was set out in his *Traité élémentaire de chimie* of 1789, and within a few years it was adopted by the majority of chemists. Priestley, however, who lived until 1804, never gave up the phlogiston theory.

(iii) *The Einsteinian Revolution.* The triumph of the Newtonian paradigm initiated a new period of normal science for astronomy (c. 1700–c. 1900). The dominant paradigm consisted of Newtonian mechanics including the law of gravity, and the normal scientist had to use this tool to explain the motions of the heavenly bodies in detail — comets, perturbations of the planets and the moon, etc. In the Einsteinian revolution (c. 1900–c. 1920), however, the Newtonian paradigm was replaced by the special and general theories of relativity.

Kuhn introduced one further very important idea: *incommensurablity.* In many scientific revolutions, so Kuhn claimed, the new paradigm is incommensurable with the old paradigm. As he says:

> The normal-scientific tradition that emerges from a scientific revolution is not only incompatible but often actually incommensurable with that which has gone before. [1962, 102]

Incommensurablility was introduced at the same time by Feyerabend, and indeed the concept may have originated in conversations between Kuhn and Feyerabend in the years 1960 and 1961. Feyerabend's account of incommensurability, however, differs significantly from Kuhn's.

Kuhn's book on scientific revolutions was read with great interest by Popper and his school in London. Lakatos invited Kuhn to give a paper at an International Colloquium in the Philosophy of Science held at Bedford College, Regent's Park, London from 11 to 17 July 1965. A collection of papers which developed from the discussions at this conference was published 5 years later (Imre Lakatos and Alan Musgrave (eds.) *Criticism and the Growth of Knowledge*, Cambridge University Press, 1970). This volume contains an essay by Kuhn himself, a series of essays criticizing and developing some of Kuhn's ideas, and Kuhn's replies to his critics. Altogether it is a most important collection for understanding the development of the historical approach to the philosophy of science at that time. Many of Kuhn's critics in the volume came from Popper's school. There is an essay by Popper himself, and essays by Watkins and Lakatos. Feyerabend too could be considered as associated with the Popper school since he had studied with Popper and even spent a year as Popper's research assistant. By the mid-1960s, his thinking had diverged very considerably from Popper's, but he was still on very friendly terms with Lakatos. The volume, however, also includes critical essays by people who were definitely not members of the Popper school — notably Margaret Masterman. We will now turn to considering criticisms and developments of Kuhn's ideas which are to be found in this 1970 volume and also in subsequent discussions. Kuhn's views involve the three key concepts of (i) *paradigm*, (ii) *normal science*, and (iii) *incommensurability*, and we will consider these three concepts in turn.

Let us start then with 'paradigm', a term which was introduced by Kuhn into the philosophy of science. Many authors have criticized it for being too vague and ambiguous. Shapere, for example, in his review of Kuhn's *The Structure of Scientific Revolutions*, goes so far as to claim that Kuhn's relativism is 'a logical outgrowth of conceptual confusions ... owing primarily to the use of a blanket term [paradigm]' [1964, 393]. In her 1970 article 'The Nature of a Paradigm', Masterman is quite sympathetic to Kuhn, yet she says 'On my counting, he uses 'paradigm' in not less than twenty-one different senses in [Kuhn, 1962], possibly more, not less' [1970, 61]. She then proceeds to list the 21 senses. Kuhn took these criticisms somewhat to heart, and in his article 'Second Thoughts on Paradigms' [1974], he suggested replacing 'paradigm' by two new concepts, namely 'disciplinary matrix' and 'exemplar'. However, these terms have never proved as popular as the original term 'paradigm'.

There is a certain irony in Kuhn's doubts about the term 'paradigm' since no other philosophical term coined in the twentieth century has proved so popular among general writers. Typing 'Paradigm' into Google produced 15,700,000 pages in 0.24 seconds. The term is constantly recurring in newspapers and magazines in contexts ranging from politics to fashion. Just one example will serve to illustrate the ubiquity of the expression. In 1998 Levi Strauss's new brand developer proclaimed that 'loose jeans is not a fad, it's a paradigm shift.'[4]

Of course the fact that a term has become so popular with the general public by no means shows that it is suitable for use in philosophy of science which perhaps demands higher standards of rigour. However, despite Kuhn's own doubts, we will argue that the term is a useful one for philosophers of science. Our defence is based on the following famous passage from Aristotle:

> Our discussion will be adequate if it has as much clearness as the subject-matter admits of, for precision is not to be sought for alike in all discussions, ... it is the mark of an educated man to look for precision in each class of things just so far as the nature of the subject admits; ... [Nicomachean Ethics I iii 1094^b 12f]

Aristotle's point of view was supported in the twentieth century by Ramsey who wrote in his general essay on Philosophy:

[4]This pronouncement is quoted in [Klein, 2000, 70].

> The chief danger to our philosophy, apart from laziness and woolliness is *scholasticism*, the essence of which is treating what is vague as if it were precise and trying to fit it into an exact logical category. [1929, 269]

So although the notion of paradigm is indeed not very precise, this does not prove that it is unsuitable for use in philosophy of science. Indeed the search for more precise notions might lead, as Ramsey suggests, to scholasticism. Our claim is that the notion of paradigm has just the right degree of precision for the subject-matter in hand, that is to say for the analysis of how science develops. We will now try to substantiate this by looking in more detail at Kuhn's discussion of the notion.

In his 1962, Kuhn introduces the notion of paradigm as follows:

> ... achievements that some particular scientific community acknowledges for a time as supplying the foundation for its further practice ... today ... are recounted, though seldom in their original form, by science textbooks, elementary and advanced. These textbooks expound the body of accepted theory, illustrate many or all of its successful applications, and compare these applications with exemplary observations and experiments. Before such books became popular early in the nineteenth century (and until even more recently in the newly matured sciences), many of the famous classics of science fulfilled a similar function. Aristotle's *Physica*, Ptolemy's *Almagest*, Newton's *Principia* and *Opticks*, Franklin's *Electricity*, Lavoisier's *Chemistry*, and Lyell's *Geology* — these and many other works served for a time implicitly to define the legitimate problems and methods of a research field for succeeding generations of practitioners. They were able to do so because they shared two essential characteristics. Their achievement was sufficiently unprecedented to attract an enduring group of adherents away from competing modes of scientific activity. Simultaneously, it was sufficiently open-ended to leave all sorts of problems for the redefined group of practitioners to resolve.
>
> Achievements that share these two characteristics I shall henceforth refer to as "paradigms," ... [1962, 10]

We would like to draw particular attention to the connection which Kuhn makes in this passage between paradigms and textbooks. Since the early nineteenth century, paradigms have, according to Kuhn, been generally taught by means of textbooks. Before the nineteenth century, he thinks that many of the famous classics of science fulfilled a similar function. However, of the classics he mentions, some were not in fact used to teach a paradigm to students, while others were so used, and can to all intents and purposes be regarded as textbooks. Thus Newton's *Principia* was not the canonical text of Newtonian mechanics for the mainstream mathematicians of the 18^{th} century, since these mathematicians preferred an approach more analytical and less geometrical than Newton's. Ptolemy's *Almagest* was certainly a classic of science, but it was also a textbook expounding the fruits of earlier work, though doubtless with many interesting additions by Ptolemy himself. Aristotle's *Physica* was actually used as a textbook in medieval universities.

If, therefore, we include under the term 'textbook' those classics of science which actually were used as textbooks, we can introduce what could be called the *textbook criterion for paradigms*. The suggestion is that, if historians wish to identify the paradigm of a group of scientists at a certain time and place, they should examine the textbooks which were used to teach the novices the knowledge needed to become fully recognised members of the group. The contents of these textbooks will then (more-or-less) define the paradigm accepted by the group.[5] This textbook criterion constitutes, in our view, a sufficient answer to those who complain that the notion of paradigm is too vague. The criterion in fact enables a historian of science to use the term 'paradigm' in quite a concrete and definite fashion.

Let us now turn from the concept of paradigm to the related concept of normal science. Here are a couple of quotations in which Kuhn picks out important features of normal science.

> When examining normal science ... we shall want finally to describe that research as a strenuous and devoted attempt to force nature into the conceptual boxes supplied by professional education. [Kuhn, 1962, 5]

> Normal science does not aim at novelties of fact or theory and, when successful, finds none. [Kuhn, 1962, 52]

Kuhn also describes the activity of the normal scientist as 'puzzle-solving' [1962, 36].

All this makes normal science sound a rather dreary and dogmatic affair. The Popperians and Feyerabend in Lakatos and Musgrave [1970] express strong hostility to the idea. Kuhn in his reply remarks quite wittily [1970, 233] that 'normal science ... calls forth some of the oddest rhetoric: normal science does not exist *and* is uninteresting.' This comment is really quite fair. The feature of normal science which is disliked by the Popperians and Feyerabend is the alleged consensus in commitment to a *single* paradigm. They think that science progresses better if there is competition between different theories, paradigms, or research

[5]The full grasp of a paradigm may also involve the ability to carry out experiments and observations. Thus we should perhaps take textbooks to include laboratory manuals and the like. Moreover, there lies beyond the reach of textbooks a certain amount of knowledge which can only be learnt by a kind of apprenticeship, e.g. practical training in the laboratory or in the field. Kuhn alludes to these matters when he speaks [1962, 41] of 'instrumental commitments that, as much as laws and theory, provide scientists with rules of the game.' In the light of this, the textbook criterion should be regarded as only approximate.

programmes. Moreover they think that competition rather than consensus is in fact the more usual state of science. These criticisms are well expressed by Feyerabend in his contribution to Lakatos and Musgrave [1970]. This has the interesting title 'Consolations for the Specialist' which suggests a way in which Kuhn's philosophy of science can be viewed. In the article, Feyerabend writes:

> More than one social scientist has pointed out to me that now at last he had learned how to turn his field into a 'science' ... . The recipe, according to these people, is to restrict criticism, to reduce the number of comprehensive theories to one, and to create a normal science that has this one theory as its paradigm. Students must be prevented from speculating along different lines and the more restless colleagues must be made to conform and 'to do serious work'. *Is this what Kuhn wants to achieve*? [1970, 198]

Feyerabend suggests that to prevent normal science getting off the ground [1970, 205]: 'we must be prepared to accept a *principle of proliferation*', and expresses [1970, 207]: 'the suspicion that normal or "mature" science, as described by Kuhn, *is not even a historical fact*.' After all, as Feyerabend goes on to say:

> ... why should we not start proliferating *at once* and *never* allow a purely normal science to come into existence? And is it too much to be hoped that scientists thought likewise, and that normal periods, if they ever existed, cannot have lasted very long and cannot have extended over large fields either? [1970, 207]

This is certainly a strong attack on both the existence and desirability of normal science, but Kuhn had already published some interesting remarks in its defence. Thus he says:

> ... history strongly suggests that, though one can practice science — as one does philosophy or art or political science — without a firm consensus, this more flexible practice will not produce the pattern of rapid consequential scientific advance to which recent centuries have accustomed us. [1959, 232]

Kuhn stresses that commitment to a paradigm may force scientists to investigate the natural world in a detail and depth which would not otherwise be achieved. This is one of the secrets of the success of normal science:

> By focusing attention upon a small range of relatively esoteric problems, the paradigm forces scientists to investigate some part of nature in a detail and depth that would otherwise be unimaginable. ... during the period when the paradigm is successful, the profession will have solved problems that its members could scarcely have imagined and would never have undertaken without commitment to the paradigm. And at least part of that achievement always proves to be permanent. [1962, 24–25]

This controversy between Feyerabend and Kuhn regarding the value of normal science is of very great interest and importance, and we will now try to assess the merits of the two sides in the debate and to suggest a possible compromise. Our discussion will hinge on the distinction between a paradigm becoming established in a community *for empirical reasons*, and a paradigm becoming established *by political methods*. Essentially we are going to argue that if a paradigm is established for empirical reasons, then normal science is likely to be fruitful and Kuhn is correct; whereas, if a paradigm is established by political methods when there are no good empirical reasons in its favour, then normal science is likely to be harmful and Feyerabend is right.

Let us start by considering the concept of empirical reasons. Suppose we have two competing theories A and B. Suppose that A explains a wide range of observations and known experimental results, while B explains much fewer of these, or does not explain them so well. Suppose that both A and B have been subjected to a number of severe experimental tests, and that A has passed them all with flying colours, while B has not done so well. We would then say that the evidence confirms (or corroborates)[6] A much more than B. If we write the degree of confirmation (or degree of corroboration) of A given the evidence (e) as $C(A, e)$, we could then say that $C(A, e)$ is much greater than $C(B, e)$. If under these circumstances A is preferred to B, we will say that A is preferred to B *for empirical reasons*. It is worth noting that this formulation assumes that the notions of confirmation or corroboration are coherent and that it is possible to develop a viable confirmation or corroboration theory. These assumptions have been questioned and we will consider some objections to confirmation theory later one. For the time being, however, we will assume that the notion of confirmation is coherent.

Perhaps the most standard example of normal science is constituted by the dominance of the Newtonian paradigm in astronomy and mechanics from the beginning of the 18^{th} century to the beginning of the 20^{th} century. With one qualification which will be made later, this can be regarded as a genuine example of normal science. So we can say that Feyerabend is wrong and that normal science has existed — at least on some occasions. Moreover it is also clear that the Newtonian paradigm was established for empirical reasons. When Newton's theory came to be accepted it had a very high degree of empirical confirmation and one that was much higher than any rival theory.[7] Now did this Newtonian normal science bring about progress in the period from c.1700 to c. 1900? It seems to us that it certainly did. To begin with there was a great deal of mathematical progress. The development of infinitesimal calculus and then of ε, δ analysis was mainly stimulated by the need to tackle ever more complicated problems in mechanics and astronomy, and this development brought a great advance to mathematics, enormously strength-

[6]We will use confirms and corroborates, confirmation and corroboration, etc. as synonyms.

[7]This is argued in detail in Gillies [1993] (see particularly pp. 218-20). However, we think it would be accepted by nearly all philosophers of science, excepting only the few who deny that the possibility of a viable concept of empirical confirmation.

ening the power of mathematical tools. Then in mechanics itself, we have a long list of developments: hydrodynamics, elasticity, Coriolis forces and the Foucault pendulum, the gryroscope and the wobbling and rotating of the Earth and other celestial bodies, etc. In astronomy, the paths of planets and comets were traced with ever more accuracy, and these calculations were extended to the stars. All these advances took place within the Newtonian paradigm and would have been impossible had Newton's theory not been generally accepted and taught. Kuhn's analysis of why normal science can succeed applies particularly well to one famous advance of this period — the discovery of Neptune. Kuhn emphasizes that normal science focuses [1962, 24]: 'attention upon a small range of relatively esoteric problems'. The esoteric problem which led to the discovery of Neptune arose because of small perturbations in the orbit of Uranus. Without the detailed development of the Newtonian mathematical apparatus, these perturbations would never have been detected. Nor would it have been possible to calculate that they could be caused by a hitherto unknown planet located in a specified position. The preceding developments of normal science were a precondition for the discovery of Neptune, and yet that discovery was a startling and dramatic one.

The example of the discovery of Neptune and others like it show that Kuhn was right to suggest that normal science could be fruitful and lead to advances. However, they also show that Kuhn was wrong in his depictions of normal science as inevitably a dreary and dogmatic affair. In particular, Kuhn gives a misleading picture in his well-known claim [1962, 52] that 'normal science does not aim at novelties of fact or theory and, when successful, finds none.' In reality many interesting novelties of fact and theory can appear within normal science. We will come back to this point in a different context in section 5.2, where the concept of abductive novelty is introduced.

The fact that normal science can be a more lively affair than Kuhn's account would lead us to believe, might help to promote a reconciliation between the Kuhnian and Popperian traditions. However this reconciliation should perhaps take a form somewhat different from the one that is usually suggested. In his contribution to the 1970 volume edited by Lakatos and Musgrave, Kuhn writes:

> I suggest then that Sir Karl has characterized the entire scientific enterprise in terms that apply only to its occasional revolutionary parts. [1970, 6]

As a matter of fact, however, it is questionable whether Popper's theory of conjectures and refutations gives a satisfactory account of scientific revolutions. The problem is that, although paradigms may be confirmed or disconfirmed by evidence, they cannot be directly refuted or falsified by evidence. Only low-level empirical generalisations, or specific, detailed models are subject to falsification by observation or experiment.[8] Thus a revolutionary change from an old to a new

[8]These claims would, we think, be accepted by most philosophers of science. Detailed arguments for them are to be found in [Gillies, 1993, 204-30].

paradigm needs a more complex characterisation than that of conjectures and refutations. On the other hand, patterns of conjectures and refutations do often occur in the context of normal science. So, in the example of the discovery of Neptune, Adams and Leverrier conjectured that the mysterious perturbations in the orbit of Uranus were caused by a hitherto unknown planet. They were able to develop this conjecture into a detailed form, which specified where this hypothetical planet should be, and so rendered the conjecture (in this specific form) refutable. In fact it was not refuted, as we know, but confirmed. However, if Neptune had not after all existed, the conjecture of Adams and Leverrier would have been refuted in due course.

Thus, contrary to Kuhn, we would like to suggest that Popper's methodology of conjectures and refutations applies not to revolutionary science, but to normal science. However, in order to adapt Popper's ideas in this way, we need to make a change to Popper's account of conjectures and refutations. Popper argues that scientists can put forward any arbitrary conjecture which is testable by experience. Indeed Popper urges scientists to put forward bold, sweeping conjectures. However, if we accept Kuhn's concept of normal science, then, during a period of normal science, a scientist cannot put forward any arbitrary testable conjecture (as Popper suggests), but only a conjecture which is compatible with the dominant paradigm. If a scientist puts forward a conjecture which contradicts the dominant paradigm, it is likely to be regarded as inadmissible by the rest of the scientific community. Admittedly in some cases, for example that of Copernicus, such a hypothesis may mark the beginning of a revolution, but, even if the hypothesis is vindicated in the long run, it is likely to be strongly opposed at first by the scientific community. So, to sum up, Popper's methodology of conjectures and refutations can be regarded as one of the principal patterns of development in normal science — provided the conjectures considered are limited to ones which are compatible with the dominant paradigm.

Naturally the formulation just given distinguishes rather too sharply between normal and revolutionary science, and this brings us to a consideration of a qualification which needs to be made to our example of Newtonian normal science in the period c. 1700 to c. 1900. This qualification affects not just this example, but the concept of normal science in general.

The qualification comes from Lakatos's paper: 'Newton's Effect on Scientific Standards' which was written in the years 1963-4, but not published until 1978 after Lakatos's death. This somewhat neglected but highly interesting paper, was written in the years immediately following the publication of *The Structure of Scientific Revolutions*, and contains an interesting criticism of Kuhn's notion of normal science. This criticism is concerned with developments in astronomy in the 18^{th} century. Lakatos begins by saying that in 1746:

> ... Clairaut found that the progress of the Moon's apogee is in reality twice what would follow from Newton's theory, and he proposed an additional term to Newton's formula involving the inverse fourth power of the distance. [1978, 219]

In other words, in the face of an anomaly, Clairaut, one of the leading scientists of the time, suggested a modification of Newton's law of gravity. Now Newton's law of gravity was part of the dominant paradigm of the time, and so Clairaut was not acting as a normal scientist should have done. His suggestion did not prove successful, however, for, as Lakatos goes on to say:

> But as it turned out, Clairaut's mathematics was wrong, and in fact later a correct calculation was found among Newton's unpublished manuscripts. But even so, a small discrepancy remained: a "secular acceleration". In 1770 the Paris Academy put up a prize for the solution of this problem. Euler won this prize with an essay in which he first concluded that "it appears to be established, by indisputable evidence, that the secular inequality of the moon's motion cannot be produced by the [Newtonian] forces of gravitation", and he proposed a rival formula again involving an additional term, which, in a sequel published a year later, he tried to explain from the resistance of Cartesian ether. However, Laplace in 1787 showed that the problem can be solved *better* within the Newtonian research programme. [1978, 219]

This historical example does have some features which Kuhn attributes to normal science since it shows scientists [1962, 24]: 'focusing attention upon a small range of relatively esoteric problems'. However, it does not exhibit the respect which scientists are supposed to show to the dominant paradigm during a period of normal science. Once again a leading scientist (Euler) was prepared to modify Newton's theory of gravity in order to explain a small observational anomaly, although, once again, the suggestion proved to be unsuccessful. Lakatos comments as follows:

> Did Clairaut and Euler make a methodological blunder — as Kuhn would surely say — when they tried alternative research programmes to solve Newtonian puzzles and only wasted time, energy and talent? [1978, 219]

Of course the answer to Lakatos's rhetorical question is obvious. Clairaut and Euler acted very reasonably. As a matter of fact, their suggested modifications of Newtonian theory were not successful, but this could not have been known in advance. Moreover the challenge of Clairaut and Euler led the Newtonians to produce an explanation of the difficulty within their own framework.

But does an example of this sort show that we should condemn normal science and follow Feyerabend's strategy of trying always to proliferate alternative theories. This would, in our view, be too extreme a response. During the long period (c. 1700 to c. 1900) of Newtonian normal science, it would not, in our opinion, have helped scientific progress if scientists had devoted a great deal of time and energy to proliferating alternatives theories of mechanics, and then debating the value of these alternatives as compared to Newtonian mechanics. In fact it was only a long series of mathematical and empirical developments based upon

Newtonian mechanics which created the possibility of creating new systems of mechanics (relativity and quantum mechanics) in the twentieth century. So it does not seem correct to give up normal science in favour of a Feyerabendian 'anything goes' position. However, Lakatos's illuminating historical example does suggest that the dogmatism of normal science should not be too rigid. Scientists should consider the possibility of now and again introducing hypotheses which contradict the some features of the dominant paradigm. Such hypotheses may often prove unsuccessful, but occasionally they may be the beginning of some new and exciting revolutionary development. Moreover, by the same token, the scientific community should allow some dissidents who do not accept the general consensus. Some discipline may be required, but too much discipline can be counter-productive.

So far we have been arguing in favour of a (somewhat qualified) normal science on the assumption that the paradigm underlying that science has come to be accepted by the community for empirical reasons. Now, however, let us consider whether there could be cases of normal science where the paradigm becomes established by political means without their being any very strong empirical reasons in its favour. Feyerabend seems to suggest that this might be possible in the passage [1970, 198] which, because of its importance, we will quote again.

> More than one social scientist has pointed out to me that now at last he had learned how to turn his field into a 'science' ... . The recipe, according to these people, is to restrict criticism, to reduce the number of comprehensive theories to one, and to create a normal science that has this one theory as its paradigm. Students must be prevented from speculating along different lines and the more restless colleagues must be made to conform and 'to do serious work'. *Is this what Kuhn wants to achieve*? [1970, 198]

Now it is worth noting that in this passage Feyerabend speaks of 'more than one social scientist', and it could be claimed that this is unfair to Kuhn, since Kuhn explicitly states [1962, x] that his theory of normal science dominated by a single paradigm applies only to mature natural sciences and *not* to the social sciences. On the contrary, there are, according to Kuhn, competing schools of thought in every branch of the social sciences and we never find a single dominant paradigm. Feyerabend, however, does in effect recognize that Kuhn holds this position, and indeed takes it as a starting point for his criticism of Kuhn. Feyerabend's argument is that social scientists who had read Kuhn could be inspired to turn their field into a science by banning all schools of thought except one. But could such a strategy actually be carried out in the contemporary university system in order to produce a normal science based on a paradigm for which there is little or no empirical confirmation?

To make the question more specific, suppose that we have two general theories A and B which are potential paradigms in some area of research. Suppose further that there is little or no empirical confirmation of A, and that the degree of confirmation of A is less than that of B. Would it nonetheless be possible to establish a normal science based on A by political methods? The answer we suggest is that it would indeed be possible.

It is a characteristic of contemporary universities in the English-speaking world and many other countries that they are arranged in a fairly strict hierarchy with those at the top exercising a great deal of power and influence over the system as a whole. The first step in establishing A would therefore have to be that of the supporters of A gaining a majority of the positions in the top universities, and particularly of the senior positions such as full professorships. Once this is achieved, then establishing A in the system as a whole becomes relatively easy. Most academics aspire to a position in a top university which is not only more prestigious but offers better conditions such as a higher salary and more research time. Once the supporters of A have gained control of the top universities, it will become clear to any aspiring academic in the field (including graduate students just starting research) that they have a much better chance of getting to a good position if they adopt A rather than B. This will be a strong incentive for adopting A. Moreover the top universities exercise a great deal of control over appointments in other universities. For example, most of those who get lectureships will have done their graduate work in one of the top universities and so will have been trained in A rather than B. Of course there will inevitably be a few obstinate characters who adopt B rather than A. However their fate is likely to be an uncomfortable one. To begin with, they may fail to get a university job at all, and, if they do get a job in a university, it is likely to be one low down in the hierarchy. In such lowly universities, the staff have much worse conditions, usually having to teach for much longer hours and having much less research time. Thus the number of research hours available for research on B will be much less than those available for research on A, which makes progress in B less likely even if it is really the better theory.

However this does not end the methods available for ensuring the triumph of A. We have not yet spoken of the control of peer-reviewed journals. Contemporary journals, like contemporary universities, are arranged in a strict hierarchy. Once the supporters of A have established themselves in the top universities, they will find it easy to acquire the editorships of the most prestigious journals. Any papers submitted can then be sent for refereeing to the friends of the editor, i.e. supporters of A. So a paper which is based on theory B is very likely to be rejected. Thus the supporters of theory B will find that they are unable to publish in the most prestigious journals, but only in the less prestigious ones. Confined to low prestige universities and publishing in low prestige journals, it will naturally be concluded that their research is no good. If there is a research assessment exercise (as in Britain), their rating will be low, and therefore they will have their research time cut still further, and might even be sacked for incompetence. Given the grim fate which is likely to hang over the supporters of B, it is only to be expected that the vast majority of researchers will adopt A, and that a normal science based on A will be established. Our conclusion then is that it would be relatively easy in many contemporary universities, to establish a normal science by purely political means.

Our argument was that this is easy to do where there is a hierarchical ranking of the universities. If the universities were more equal with similar conditions of work and levels of prestige throughout, then it would be much harder to use the political methods just described.

So far we have spoken of contemporary universities, but, at other times, other less 'gentle' methods have been available for establishing a normal science by purely political methods. Dissident supporters of a rival approach to the dominant paradigm could have been handed over to the inquisition or sent to a labour camp.

It is, and has been, therefore eminently possible to establish, by purely political means, a dominant paradigm which has little or no empirical confirmation (or at least much less empirical confirmation than some rival), but which would nonetheless become the basis of a normal science research tradition. However, this is to speak hypothetically. We can still ask whether this possibility has ever actually occurred or is actually occurring. Here we enter a speculative and controversial area, about which there is likely to be disagreement. However, three possible examples of a normal science established by political methods do suggest themselves. First of all the Ptolemaic theory was the basic paradigm for astronomy among the Jesuits in the 17^{th} century. Secondly Lysenkoism was the basic paradigm for research in biology in the Soviet Union under Stalin[9]. Thirdly neo-classical economics is the basic paradigm for economics in most contemporary universities. Fullbrook [2004] is a recent collection of 27 essays by different authors who criticize neo-classical economics from many points of view. These criticisms establish that there is indeed little or no empirical confirmation for this theory. Indeed because of the lack of realism of its basic assumptions, one could say that the degree of confirmation of the theory is negative. Yet neo-classical economics is unquestionably the dominant paradigm in most economics departments throughout the world.

These examples show that paradigms with little or no empirical confirmation, or at least with a degree of empirical confirmation much less than some rival, can indeed be established by political methods. We can agree with Feyerabend that normal science research founded on such a paradigm is unlikely to be fruitful. On the other hand if a paradigm comes to be accepted for empirical reasons, and has a degree of corroboration which is not only high but very much higher than any rival, then a normal science research tradition founded on such a paradigm may very well prove very fruitful, as was the case with research founded on the Newtonian paradigm in the period c. 1700 to c. 1900. In such a case Kuhn seems to be right and Feyerabend wrong.

At this point it might be objected that it is a little naïve to distinguish so sharply between paradigms which are accepted for empirical reasons, and those which are established by political methods. Surely, it could be said, that a mixture of the two processes occurs in most cases. Of course there is some truth in this, but it does not require a strong modification of the position here advocated. Strictly we should speak of paradigms which are accepted for predominantly empirical reasons as against those which are established principally by political methods. Our judge-

[9]A good account of Lysenkoism is to be found in Sheehan [1985].

ment is that most paradigms which have been established historically fall into one of these two categories, and there is not in reality a spectrum of intermediate cases. The reason for this is that the scientific community, if left to itself and not influenced by ideological/political factors originating from outside science, will accept paradigms predominantly for empirical reasons. Thus we can speak of science as a rational enterprise whose rationality is occasionally disturbed by powerful ideological/political currents coming from outside science. In the three examples we gave earlier of paradigms which were perhaps established purely by political means, it is clear that ideological/political factors were acting strongly. The Jesuits in the 17^{th} century continued to hold to the Ptolemaic paradigm for astronomy for religious reasons because the Copernican paradigm was held to contradict the teachings of the Catholic Church. Stalin, who of course was quite ignorant of biology, took a liking to Lysenko's ideas, and this was sufficient for these ideas to become the dominant paradigm in the Soviet Union. As for neo-classical economics, its principal function is to justify contemporary neo-conservative economic policies. It is thus not surprising, in view of recent political trends, that it has come to be the dominant paradigm among economists, despite its empirical weaknesses.

Let us now turn to a key question which arises in this connection. We have drawn the distinction between paradigms accepted for empirical reasons, and those established by political methods. However this distinction is only valid if there can indeed be empirical reasons for accepting a paradigm, which, as we argued earlier, is equivalent to saying that a theory of confirmation or corroboration for scientific theories can be developed? But is it really possible to develop such a theory? Both the earlier logical tradition of the Vienna Circle and also Popper did accept the possibility of confirmation theory. Admittedly, as we have seen, Hempel produced some paradoxes of confirmation such as the famous paradox of the ravens. However, this certainly did not lead him, or others of the same way of thinking, to conclude that a confirmation theory could not exist. After all, in deductive logic, paradoxes such as Russell's paradox had arisen, but such paradoxes had not proved fatal to deductive logic. On the contrary, ways round these paradoxes, such as the theory of types or axiomatic set theory had been developed and seemed to function well. In the same way, it was held that the paradoxes of confirmation could also be overcome. Even Popper, despite his criticism of the Vienna Circle, accepted the existence of what he called: 'corroboration'. Popper used a different term from 'confirmation' to distinguish his theory from that the confirmation theory of Carnap. In this paper, we are using 'confirmation' and 'corroboration' as synonyms, and so prefer to speak of Popper having a different theory of confirmation (or corroboration) from Carnap's theory of confirmation (or corroboration). Certainly there were differences between the two theories. Carnap was a Bayesian and held that the confirmation function $C(h, e)$ satisfied the axioms of probability, while Popper held that $C(h, e)$ did not satisfy these axioms. There were other differences besides. Yet the two thinkers did both accept the possibility of developing a confirmation theory of some kind. If we turn to Kuhn, we find that in his [1962] he appears to reject the possibility of a confirmation theory.

Kuhn brings up the question in Chapter XII of his [1962] which is entitled: 'The Resolution of Revolutions'. Here Kuhn considers first Bayesian confirmation theories which he calls: 'probabilistic verification theories'. He has this to say about them:

> Few philosophers of science still seek absolute criteria for the verification of scientific theories. Noting that no theory can ever be exposed to all possible relevant tests, they ask not whether a theory has been verified but rather about its probability in the light of the evidence that actually exists. ... In their most usual forms, however, probabilistic verification theories all have recourse to one or another of the pure or neutral observation-languages ... If, as I have already urged, there can be no scientifically or empirically neutral system of language or concepts, then the proposed construction of alternate tests and theories must proceed from within one or another paradigm-based tradition. Thus restricted it would have no access to all possible experiences or to all possible theories. As a result, probabilistic theories disguise the verification situation as much as they illuminate it. [1962, 144–5]

Kuhn then goes on to consider Popper's views. He first makes the point that what he calls 'anomalies' have some points in common with what Popper calls 'falsifications'. However, Kuhn then continues:

> If any and every failure to fit were ground for theory rejection, all theories ought to be rejected at all times. On the other hand, if only severe failure to fit justifies theory rejection, then the Popperians will require some criterion of "improbability" or of "degree of falsification." In developing one they will almost certainly encounter the same network of difficulties that has haunted the advocates of the various probabilistic verification theories. [1962, 145–6]

These passages show that Kuhn in 1962 was very doubtful about the possibility of a confirmation theory either of the Bayesian or the Popperian kind.

Moreover if a confirmation theory is going to be useful in analysing scientific revolutions, it would be necessary to compare the degrees of confirmation given the evidence of the two competing paradigms. But can paradigms be compared in this way? Kuhn held that two competing paradigms are incommensurable, and this suggests that he would deny that they could be compared as to their respective degrees of confirmation (or corroboration) given the available evidence. In fact the Collins Dictionary of the English Language defines incommensurable as follows: 'incapable of being judged, measured or considered comparatively.' If this is what Kuhn meant by incommensurable, then it would follow that two paradigms could not be compared with regard to their degrees of empirical confirmation, and that consequently there could not be empirical reasons for preferring one to the other.

It would seem to follow from this that a new paradigm could only be established by political methods. Now, as we shall see when we come to discuss the incommensurability problem in section 3.4, Kuhn later denied that he had ever intended to use 'incommensurable' in such an extreme sense. However, established English usage certainly suggested that that was what he meant, and he was interpreted in this way by many of his contemporaries, including notably Lakatos who described Kuhn's position as follows:

> For Kuhn scientific change — from one 'paradigm' to another — is a mystical conversion which is not and cannot be governed by rules of reason and which falls totally within the realm of the *(social) psychology of discovery*. Scientific change is a kind of religious change. [1970, 9]

This was elaborated later in the same paper as follows:

> ... a new 'paradigm' emerges, incommensurable with its predecessor. There are no rational standards for their comparison. Each paradigm contains its own standards. The crisis sweeps away not only the old theories and rules but also the standards which made us respect them. The new paradigm brings a totally new rationality. There are no super-paradigmatic standards. The change is a bandwagon effect. Thus *in Kuhn's view scientific revolution is irrational, a matter for mob psychology*. [1970, 90–1]

Naturally Lakatos does not approve of Kuhn's position as thus interpreted. He (Lakatos) has this to say about it:

> If even in science there is no other way of judging a theory but by assessing the number, faith and vocal energy of its supporters, then this must be even more so in the social sciences: truth lies in power. Thus Kuhn's position vindicates, no doubt, unintentionally, the basic political *credo* of contemporary religious maniacs ('student revolutionaries').[10] [1970, 9–10]

Lakatos was determined to struggle against what he saw as Kuhn's irrationalism by developing:

> [a] position which, I think, may escape Kuhn's strictures and present scientific revolutions not as constituting religious conversions but rather as rational progress. [1970, 10]

In the next section, we will examine how Lakatos carried out this project, and how successful he was in achieving his goal of defending rationality.

[10] It is interesting that in this passage written in the late 1960s, Lakatos regards 'student revolutionaries' as the obvious contemporary examples of 'religious maniacs'. *Tempora mutantur.*

3.3 Lakatos and the Methodology of Scientific Research Programmes

Lakatos based his new approach to methodology on the concept of *research programme*. This had already been used for the purposes of analysing science within the Popperian school. In his [1934], Popper had defended the meaningfulness of metaphysics against the claim by the Vienna Circle that metaphysics is meaningless. One of Popper's most striking examples to support this thesis was that of *atomism*. Atomism was first introduced in the West by the pre-Socratic thinkers Leucippus and Democritus. It continued as a powerful trend in the ancient world with Epicurus in Greece and Lucretius in Rome. This ancient atomism was surely meaningful, but also definitely metaphysical.

Ancient atomism was revived in Western Europe in the seventeenth century, and discussed by the leading scientists of the day, but it was still at that time metaphysical rather than scientific. It was not till the nineteenth century with the work of Dalton, Maxwell and Boltzmann that atomism became scientific. These scientists were however influenced by the earlier metaphysical atomism, which shows that metaphysics can be, not only meaningful, but also helpful to science.

In his [1983] *Realism and the Aim of Science*, Popper develops his views on metaphysics by introducing the concept of a *metaphysical research programme* for science. Thus he says:

> ... atomism is an excellent example of a non-testable metaphysical theory whose influence upon science has exceeded that of many testable theories. [1983, 192]

And, after giving some further examples of metaphysical theories which have influenced science, he continues:

> Each of these metaphysical theories served, before it became testable, as a research programme for science. It indicated the direction of our search, and the kind of explanation that might satisfy us; and it made possible something like an appraisal of the depth of a theory. [1983, 192–3]

Although not published until 1983, this was written in 1956, and undoubtedly influenced Lakatos in the development of his new ideas on methodology. Lakatos, however, changed Popper's concept of metaphysical research programmes to that of *scientific research programmes*.

Lakatos uses two notions to characterise a scientific research programme. These are the *hard core* or *negative heuristic*; and the *positive heuristic*. We will deal with them in turn.

Lakatos explains the notion of *hard core* as follows:

> All scientific research programmes may be characterized by their '*hard core*'. The negative heuristic of the programme forbids us to direct the *modus tollens* at this 'hard core'. Instead, we must use our ingenuity to

> articulate or even invent 'auxiliary hypotheses', which form a *protective belt* around this core, and we must redirect the *modus tollens* to *these*. [1970, 48]

He then goes on to give the following example:

> In Newton's programme the negative heuristic bids us to divert the *modus tollens* from Newton's three laws of dynamics and his law of gravitation. This 'core' is 'irrefutable' by the methodological decision of its proponents: anomalies must lead to changes only in the 'protective' belt of auxiliary, 'observational' hypotheses and initial conditions. [1970, 48]

Here Lakatos is influenced by the Duhem thesis. Let T stand for the conjunction of Newton's three laws of motion and the law of gravitation, and A for some auxiliary assumptions. Duhem pointed out that we cannot derive observational results from T alone, but only from the conjunction of T and A. If O is derived, and not-O is observed, then we have the choice of changing T or A. Lakatos suggests that those working on the Newtonian programme should decide in advance to change A, and not T.

Let us now turn to the second of Lakatos' characterising notions — that of *positive heuristic*. Here are two passages in which he describes this concept.

> ... the positive heuristic consists of a partially articulated set of suggestions or hints on how to ... develop ... the research-programme ...' [1970, 50]

> Positive heuristic is thus in general more flexible than negative heuristic. ... It is better therefore to separate the 'hard core' from the more flexible metaphysical principles expressing the positive heuristic. [1970, 51]

We see here the influence on Lakatos of Popper's ideas about the possibility of metaphysical ideas helping science forward.

This then is a brief outline of Lakatos' concept of scientific research programme. Let us now turn to considering some criticisms of the notion. The first question which could be raised is whether the concept of scientific research programme really differs from that of paradigm. After all, Lakatos' Newtonian research programme with its hard core looks very like Kuhn's normal science based on the Newtonian paradigm. Has Lakatos done no more than express Kuhn's ideas in a different terminology? Our earlier discussion and attempted clarification of the concept of paradigm shows that this is not the case and that the two notions really are different. Moreover it makes clear how they differ.

A paradigm, we argued, consists of the assumptions shared by all those working in a given branch of science at a particular time. Historians can reconstruct the paradigm of a specific group at a particular time by studying the text-books used to instruct those wishing to become experts in the field in question. Thus a paradigm is what is common to a whole community of experts in a particular field at a particular time. By contrast only a few of these experts (or in the limit only one) may be working on a particular research programme. Characteristically only a handful of vanguard researchers are working on a specific research programme at a particular time. Historians who wish to reconstruct a research programme will look, not at textbooks in wide circulation, but at the writings of a few key figures. They will examine the notebooks, the correspondence, and the research publications of these leading figures, and, in this way, reconstruct the programme on which they were working. This shows clearly how research programmes differ from paradigms.

In a moment we would like to defend the concept of scientific research programme still further by arguing that it is not only different from the concept of paradigm, but is needed in addition to the concept of paradigm in order to give an adequate analysis of scientific revolutions. Before doing so, however, we would like to make a couple of further criticisms of the notion of scientific research programme which, we think, can result in some modifications and improvements of the concept.

The first of these criticisms is directed against the notion of hard core, and, oddly enough, is based on the example given earlier from Lakatos' paper: 'Newton's Effect on Scientific Standards.' It will be remembered that this paper, though first published in 1978, was largely written in the years 1963–4 when Lakatos' ideas were closer to Popper's than they became later on. In this paper Lakatos discusses the work of Clairaut and Euler in the 18^{th} century. These two scientists would presumably have been working on what in Lakatos' terminology could be characterised as the Newtonian scientific research programme. Yet Clairaut in 1746 tried to explain the motion of the Moon's apogee by changing Newton's law of gravity, and a similar strategy was followed by Euler in 1770. So both Clairaut and Euler suggested changing the hard core of the programme rather than the auxiliary assumptions — contrary to the methodology of scientific research programmes. Admittedly these suggestions did not prove successful in the long run, but there is no *a priori* reason why they should not have succeeded.

Although the Clairaut and Euler example contradicts Lakatos' methodology of scientific research programmes, it does not perhaps constitute a very severe counter-example. Kvasz in his discussion of Lakatos' methodology [2002, 236] divides the hard core into a series of layers like an onion. The changes of Clairaut and Euler were, in Kavasz's terminology, *re-formulations* only affecting the outer layer of the onion. This suggests that some modification of Lakatos' approach is needed, but not too drastic a one.

Another general argument in the same direction is that the methodological decision which Lakatos recommends of rendering the hard core 'irrefutable' contradicts the open-mindedness, and lack of dogmatism, which should characterise the good researcher.

For these reasons, we suggest replacing the notion of the *hard core* of a programme by that of the *aim* or *goal* of the research programme. After all, scientific research is a conscious human activity, and so has a goal. Thus we can say that Clairaut and Euler were working on a research programme whose aim was to explain the motion of heavenly bodies using mechanical laws. The concept of the aim of a research programme is related to that of hard core, but is less dogmatic. It is possible to have an aim or goal without being certain that one can attain it, and it is moreover always possible to change the aim of an activity when the original aim is shown to be impossible.

This suggested change from 'hard core' to 'aim' is further supported by the fact that many notable scientific research programmes do not appear to have had anything resembling a hard core. A good example of this is the research programme which Lavoisier began around 1772. We are fortunate in having Lavoisier's own description of his research programme in a memorandum which he wrote probably on 20 February 1773.[11] Here he speaks of '... the long series of experiments that I intend to make on the elastic fluid that is set free from substances, either by fermentation or distillation or in every kind of chemical change, and also on the air absorbed in the combustion of a great many substances ... ...' There is nothing here at all like a 'hard core' for the research programme, but the programme certainly had an aim, because Lavoisier writes that these experimental investigations are 'in order to link our knowledge of the air that goes into combination or that is liberated from substances, with other acquired knowledge, and to form a theory.' Moreover he also says: 'The importance of the end in view prompted me to undertake all this work, which seemed to me destined to bring about a revolution in physics and chemistry.'

Another feature of Lavoisier's research programme is that he seems to have been largely uninfluenced by metaphysical considerations — unlike other scientists such as Kepler. Thus we could identify the positive heuristic of his programme as consisting of a range of experimental techniques and apparatus such as the pneumatic trough, burning glasses, furnaces, electric sparks, balances, etc. This example leads to our second criticism of Lakatos which is that, in his notion of positive heuristic, he is perhaps over-influenced by Popper's emphasis on metaphysics. As well as the 'metaphysical principles' mentioned by Lakatos, the positive heuristic could contain other things such as mathematical and experimental techniques.

Having suggested a few modifications in the concept of scientific research programme, we will now argue that this concept is needed *in addition to* that of paradigm in order to explain how paradigms come into existence. Kuhn has a rather romantic theory that a new paradigm is born in a flash of intuition. As he puts it:

> ... normal science ultimately leads only to the recognition of anoma-

[11]The quotations from Lavoisier's memorandum are taken from the English translation in [McKie, 1935, 120–3]. There has been some scholarly discussion about the date of this memorandum because it is clearly dated February 20, 1772, but is written on the opening pages of a laboratory note-book, dated from February 20 to August 28, 1773. We have adopted McKie's date of 1773 in view of the convincing arguments he presents for it in his [1935, 123–4].

> lies and to crises. And these are terminated, not by deliberation and interpretation, but by a relatively sudden and unstructured event like the gestalt switch. Scientists then often speak of the "scales falling from the eyes" or of the "lightning flash" that "inundates" a previously obscure puzzle, enabling its components to be seen in a new way that for the first time permits its solution. On other occasions the relevant illumination comes in sleep. No ordinary sense of the term "interpretation" fits these flashes of intuition through which a new paradigm is born. [1962, 121–2]

Now, there may indeed be a few cases in which paradigms are born in something like this fashion. The most convincing example is one suggested by Arthur Miller. If we regard the Bohr atom as a paradigm and quantum mechanics as the new paradigm which replaced it, then it does indeed seem that the basic ideas of the new quantum mechanics came to Heisenberg, if not in a 'lightning flash', then at least in a few months of feverish inspiration.[12] In general, however, a new paradigm is fashioned over a much longer period of time, and by a process which may involve flashes of inspiration, but which may also involve long periods of systematic and painstaking research. It is usually, in fact, work on research programmes by small groups, or, in the limit a single individual, which gives rise to a new paradigm.

Consider the case of Copernicus. He introduced a new research programme whose aim was to explain the motion of the heavenly bodies on the assumption that the Earth rotated on its axis, and moved once a year round a stationary Sun. Copernicus was indeed influenced by metaphysical ideas, more specifically by Pythagoreanism and Neo-Platonism which were both popular during the Renaissance period in which he lived. However, the positive heuristic of his programme contained some technical considerations. Copernicus used epicycles but deliberately eschewed the equants which had been used by Ptolemy. Copernicus' research programme was certainly not a paradigm, i.e. part of the preliminary 'text-book' instruction received by scientists training in the field. Indeed he was the only scientist working on his research programme. The theory which resulted from his long years of research did not become a paradigm either, though it was taken up by a few of the scientists working in the field. After Prolemy, the next paradigm to become generally accepted was the Newtonian, and this new paradigm, although it did contain Copernicus's heliocentric assumption, was in other ways, quite different from anything that Copernicus could have imagined. Moreover considerable work on further research programmes — those of Kepler, Descartes, Galileo, and Newton himself — were necessary before Copernicus' theory could be transformed into the new Newtonian paradigm. This example shows clearly that the concept of scientific research programme differs from that of paradigm, and that we need the concept of scientific research programme in order to explain how paradigms come into existence.

Thus far we have defended Lakatos' concept of scientific research programme —

[12]Some details about this example are to be found in [Miller, 1986, 127–43 & 248–54].

albeit in a somewhat modified form. However, we have defended this concept as filling a gap in Kuhn's account rather than as a replacement for Kuhn's account. It was of course the latter which Lakatos himself intended. As we explained earlier Lakatos held that [1970, 91] '... *in Kuhn's view scientific revolution is irrational, a matter for mob psychology.*', and that for Kuhn, scientific change occurs according to the 'political *credo*' that 'truth lies in power' [1970, 10]. Lakatos saw himself as defending the rationality of science against such doctrines. But did he succeed in this defence? This is the question which we must next consider.

In order to defend the rationality of science, Lakatos has to formulate some rational criteria for preferring one scientific research programme to another. This he does by distinguishing between *progressive* and *degenerating* research programmes. Progressiveness has both a theoretical and empirical character, and, as regards the empirical side, Lakatos stressed the exclusive importance of the production of *novel facts*. Thus he writes:

> The time-honoured empirical criterion for a satisfactory theory was agreement with the observed facts. Our empirical criterion for a series of theories is that is should produce new facts. *The idea of growth and the concept of empirical character are soldered into one.* [1970, 35]

and again [1970, 38]: 'the only relevant evidence is the evidence anticipated by a theory'.

Lakatos' views on novel facts were, however, criticized in a decisive fashion by his former pupil and then colleague at the London School of Economics — Zahar. Zahar wrote:

> Lakatos mentions the return of Halley's comet as a new fact anticipated by the Newtonian programme and, of course, I agree with him that the discovery of any new type of fact is the discovery of a novel fact. But, if we equate novelty simply with *temporal* novelty, we are driven into a paradoxical situation. We should, for example, have to give Einstein no credit for explaining the anomalous precession of Mercury's perihelion, because it had been recorded long before General Relativity was proposed. Similarly, we should have to say, contrary to informed opinion, that Michelson's experiment did not confirm Special Relativity and Galileo's experiments on free fall did not confirm Newton's theory of gravitation. Lakatos, who does not easily dismiss the judgements of physicists, is aware of this difficulty and tries to avert it by shifting his original view and saying that, in the light of a new theory, some known facts may 'turn into' novel ones. For example, whereas Balmer merely 'observed' that the hydrogen lines obey a certain formula, Bohr connected these lines with the energy levels of the electron in the hydrogen atom.
>
> However, Lakatos's modified notion of 'novel fact' is open to the following fatal objection. Any theory is a set of propositions connecting

> different terms and relations. We can always define the properties of a physical entity like mass through the relations which 'mass' bears to other concepts and notions within the theory. Consequently a new hypothesis will generally ascribe new meanings to old terms. For instance, any experimental consequence of relativity theory involving say mass, would trivially become the expression of a novel fact. Thus the fact that a steel ball rolling down a slope takes a certain time to reach the bottom, could become a novel fact when the steel ball is considered as having relativistic mass. This is obviously absurd. Therefore Lakatos's 1970 criterion for novelty is too liberal, while his 1968 criterion is too stringent. [1973, 101–2]

Zahar, however, is not just critical but makes a positive suggestion as to how the difficulty could be overcome. He proposes a redefinition of 'novel fact' which he states as follows:

> A fact will be considered novel with respect to a given hypothesis if it did not belong to the problem-situation which governed the construction of the hypothesis. ... Temporal novelty in a research programme is then a sufficient but not a necessary condition for novelty. [1973, 103]

The advantage of this definition of novel fact is that it allows us to say that the anomalous precession of Mercury's perihelion was a novel fact for Einstein's General Relativity because it was not part of the problem-situation which led Einstein to construct his new theory, or, in other words, it was not part of the heuristic of Einstein's programme. Lakatos accepted Zahar's modification of the concept of novel fact, but there do appear to be difficulties with Zahar's concept of novel fact as well. Zahar points out some of the consequence of his new concept as follows:

> This new criterion for novelty of facts also implies that the traditional methods of historical research are even more vital for evaluating experimental support than Lakatos had already suggested. The historian has to read the private correspondence of the scientist whose ideas he is studying; his purpose will not be to delve into the psyche of the scientist, but to disentangle the heuristic reasoning which the latter used in order to arrive at a new theory. Let us give an example. In Newton's time there was a well-known inverse square law for the intensity of light, Newton might have used some reasoning by analogy in order to propose that the gravitational 'intensity' is also distributed over the surface of a sphere and hence obeys an inverse square law; in this case Kepler's laws would support gravitational theory more strongly than if Newton had used them as his heuristic starting point. [1973, 103–4]

These points of Zahar's have rather counter-intuitive consequences. Suppose there had been another scientist (Dupont say) who was a contemporary of Newton and working like Newton on gravitational theory. We know that Newton used Kepler's laws as part of his heuristic, but let us suppose that Dupont was familiar with large parts of Descartes and Galileo, but, for some curious reason, quite ignorant of Kepler's work. Dupont used the analogy with light suggested by Zahar to develop the mechanics of Descartes and Galileo, and, in this way arrived at exactly the same theory as Newton. Only after formulating the theory did Dupont discover Kepler's work, and, being a man of great genius, he quickly showed that Kepler's laws in an approximate form followed from his new mechanics. According to Zahar, Dupont's theory would be much better supported by the evidence than Newton's theory — even though the two theories are identical. This seems to constitute a rather severe difficulty.

Moreover, in general terms, it seems rather questionable whether the scientific community needs to investigate the private correspondence of scientists in order to decide whether their theories are well-supported by experiment and observation. Suppose, for example, that a historian discovers a hitherto unknown notebook of Einstein's which reveals that Einstein spent several months considering possible explanations of the anomalous precession of Mercury's perihelion, and that this research was actually crucial for his later development of General Relativity. According to Zahar, this historical discovery should lead the scientific community to lower the empirical confirmation they have hitherto accorded to General Relativity. Surely, however, this would not be the case.

In the light of all this, there still seem to be unresolved problems concerning the notion of novel fact used in Lakatos' methodology of scientific research programmes. Let us leave these problems aside for the moment, however, as we have to consider some further difficulties. In Lakatos' framework, scientists have to choose on rational grounds between competing research programmes — R_1 and R_2 say. Now suppose that R_1 is degenerating and R_2 progressing, then it would seem rational for scientists to choose R_2 in preference to R_1. But now comes the difficulty. It could happen that R_1 having degenerated for a while, suddenly turns the corner and starts progressing again, while the opposite occurs with R_2. R_2 having been progressing, suddenly loses its momentum and begins to degenerate. Then the choice of R_2 rather than R_1 would turn out to have been the wrong one. Lakatos was aware of this difficulty and responded by proclaiming the end of instant rationality. He writes:

> It is very difficult to decide ... when a research programme has degenerated hopelessly; or when one of two rival programmes has achieved a decisive advantage over the other. There can be no "instant rationality". [1974, 149]

During the period 1968-74, Lakatos discussed the question of the rationality of science and other issues in the philosophy of science continually with Feyerabend with whom he was very friendly. Luckily their correspondence in these years has survived and has been published by Motterlini in his [1999], together with Lakatos' last lectures on scientific method. To simplify one could say that in the discussions between Lakatos and Feyerabend in this period, Lakatos attempted to defend the rationality of science whereas Feyerabend argued for its irrationality. In his 1975 book: *Against Method*, Feyerabend argued for the principle (p. 28): '*anything goes*', and the title of his 1987 book was: *Farewell to Reason*. This gives a good general idea of Feyerabend's position, while we have already seen that Lakatos was concerned to defend the rationality of scientific change against a threat thereto which he saw as coming from Kuhn. Those who are interested in details of the discussions between Feyerabend and Lakatos should consult the material in Motterlini's volume, including Motterlini's own admirable account of the controversy. Here we want to pick out just one point — that regarding the issue of the end of instant rationality. Feyerabend took up this issue in his criticism of Lakatos in 1975, where he writes:

> Considering a research programme in an advanced state of degeneration one will feel the urge to abandon it, and to replace it by a more progressive rival. This is an entirely legitimate move. *But it is also legitimate to do the opposite* and to retain the programme. ... it is ... unwise to reject research programmes on a downward trend because they might recover and might attain unforeseen splendour ... Hence, one cannot *rationally* criticize a scientist who sticks to a degenerating programme and there is no *rational* way of showing that his actions are unreasonable. ... the arguments that established the need for more liberal standards make it impossible to specify conditions in which a research programme *must* be abandoned, or when it becomes *irrational* to continue supporting it. *Any* choice of the scientist is rational, because it is compatible with the standards. 'Reason' no longer influences the actions of the scientist. [1975, 185–6]

Somewhat reluctantly perhaps one has to admit that Feyerabend gets the better of the argument here. His criticism, and the difficulties connected with the concept of novel fact which were described earlier, together show that Lakatos did not, with his methodology of scientific research programmes, succeed in what he had hoped to do, namely to construct a defence of the rationality of scientific change. Lakatos' failure in this respect is connected, in our view, with a feature of his thinking which we will now discuss.

One of the basic distinctions frequently made in the philosophy of science is between the *discovery* of scientific hypotheses and their *justification*. Most philosophers of science (including the present authors) accept this distinction, but a minority do challenge it, and Lakatos was among that minority. This is shown in the fact that Lakatos tried to reduce the problem of the *appraisal* of knowledge (a matter of justification) to that of the *growth* of knowledge (a matter of discovery). Lakatos puts forward this position in the following passage:

> But then two new problems arose. The *first* problem was the *appraisal of conjectural knowledge. ...* The *second* problem was *the growth of conjectural knowledge. ...*
>
> In this situation *two schools of thought emerged.* One school — *neoclassical empiricism* — started with the first problem and never arrived at the second. The other school — *critical empiricism* — started by solving the second problem and went on to show that this solution solves the most important aspects of the first too. [1968, 132–3]

Our own position in contrast to Lakatos' is that we must solve both problems — that of the *appraisal* of knowledge, and that of the *growth* of knowledge. Although the two problems are connected, they are distinct nonetheless, and a solution to the second problem does not solve the most important aspects of the first too. It was the erroneous assumption that it does, which, in our view, is the root cause of Lakatos' failure to defend successfully the rationality of scientific change. The same cause is also responsible for the difficulties in Zahar's development of Lakatos' position.

Our claim then is that philosophers of science have to develop not only a theory of the growth of science, but also a theory of the appraisal of scientific hypotheses. Lakatos' theory of scientific research programmes constitutes, in a modified form, an important contribution to the theory of the growth of science. However, it does not contribute to the theory of appraisal of scientific hypotheses. For that one needs to develop a theory of the confirmation or corroboration of scientific hypotheses by evidence — a theory which cannot be based on the methodology of scientific research programmes. In fact the appraisal of a scientific research programme as either progressing or degenerating may give little indication as to whether the theories of which it is composed are well-confirmed or strongly disconfirmed. To see why this is so, let us define two scientific research programmes:

$$\mathbf{R}_1 = (T_1, T_2, \dots\ T_n)$$
$$\mathbf{R}_2 = (S_1, S_2, \dots\ S_m)$$

The following situation is possible:

1. $\mathbf{R}_1$ makes very good progress, but the theory T_n is not very well confirmed. This case occurs (for n small) in the initial stages of many programmes, for example Bohr's programme.

2. $\mathbf{R}_2$ degenerates, but S_m has a very high degree of confirmation. An example of this case is given by the hidden variables programme in quantum mechanics. The attempts to replace standard quantum mechanics by a new and better theory have hitherto failed. So $S_1 = S_m$ = standard quantum mechanics, and this shows a total stagnation of the research programme. But S_m has a very high degree of confirmation, partly because of those experiments, for example Aspect's experiment, which were carried out in the context of work on $\mathbf{R}_2$.

Examples of this sort show that it is not possible to base a theory of the confirmation of scientific theories by evidence on the methodology of scientific research programmes. Lakatos' mistaken attempt to reduce the appraisal of scientific theories to the theory of the growth of science is, we think, the consequence of a notable aspect of his writings. Lakatos gives many examples of scientific and mathematical discoveries, but he never mentions the practical applications of science and mathematics. Reading only Lakatos, a Martian would have the impression that mathematics and science are intellectual amusements for humans similar to novels. He, she (or it!) would not be in the position to guess that mathematics and science are used in industry and commerce. Modern science, however, depends for its existence on the continual application of mathematics and science.

For the satisfactory application of science, we need a theory of the appraisal of scientific hypotheses which does not involve detailed considerations of how those hypotheses are discovered. To give just one obvious example, it is not permitted to sell a new medicine until the hypothesis that it is effective but has no harmful side effects has been very well confirmed. Indeed most governments specify the tests which must be carried out with satisfactory results before a company is allowed to put a new medicine on the market. Pharmaceutical companies are not, however, required to make public the heuristic research strategies which led them to their discoveries, and indeed they make every effort to keep these strategies secret.

Turning now to Zahar, we can agree that reading the private correspondence of Dupont and Newton may be highly relevant for uncovering the heuristic strategies which led them to their theories. However, given that they discovered the same theory (as in our hypothetical example), then these heuristic strategies would not be relevant to evaluating the experimental support for that theory. The scientific community would have to consider the degree to which this new theory is confirmed in the light of the observations and experimental results in the public domain, and to consider what further observations might be made and experiments carried out in order to test the new theory severely.

In fact the methodology of scientific research programmes, contrary to Lakatos' intention, would make science very vulnerable to manipulation by politics and ideology coming from outside the scientific community. Let us consider again two scientific research programmes $\mathbf{R}_1$ and $\mathbf{R}_2$ as defined above. This time, however, let us suppose that $\mathbf{R}_2$ has been making much better progress than $\mathbf{R}_1$, and that progress and confirmation are in agreement so that S_m is much better confirmed empirically than T_n. Suppose, however, that some powerful group favours the approach of $\mathbf{R}_1$ and dislikes that of $\mathbf{R}_2$ for political and ideological reasons. This group might use political methods, such as we described earlier, to ensure that almost all further research is done in $\mathbf{R}_1$ rather than $\mathbf{R}_2$. As a result the researchers in $\mathbf{R}_1$ produce a series of new theories $T_{n+1}, \ldots, T_{n+r}$. There is some slight improvement here so that by the criteria of the methodology of scientific research programmes, $\mathbf{R}_1$ can be said to be progressing slightly. Meanwhile, because almost no researchers are working on $\mathbf{R}_2$, no new theories are produced in that programme and it remains stuck at S_m. $\mathbf{R}_2$ would therefore be stagnant and $\mathbf{R}_1$ progressing.

So Lakatos and his followers would have to judge that $\mathbf{R}_1$was superior to $\mathbf{R}_2$ even though this result had clearly been brought about by political manipulation, and even though it might still be the case that S_m was much better corroborated by the evidence than T_{n+r}.

If however the scientific community has well established criteria for when and to what extent a theory is confirmed or disconfirmed by evidence, it could resist such political manipulation. Scientists could object that it is obviously unfair to give all the research funding to one research programme when another research programme has produced theories which are better confirmed empirically. This example shows once again that the defence of scientific rationality requires the development of a theory of empirical confirmation or corroboration, and that questions about confirmation are distinct from those concerning the progress, stagnation or degeneration or research programmes.

Thus our conclusion is that Lakatos did not solve the problem he set out to solve, that is the problem of whether scientific revolutions can be rational. However, he did create the tools for solving another problem namely the problem of how new paradigms are created, since new paradigms almost always come into existence as the result of work on one or more research programmes. This may seem a strange claim to make about Lakatos' ideas, but it is surprisingly in accordance with something Lakatos himself wrote, namely:

> After Columbus one should not be surprised if *one does not solve the problem one has set out to solve.* [1963-4, 90]

Our discussion has also led to the conclusion that the solution to the problem of the rationality of scientific revolutions (if there is one) must lie in the development of a theory of empirical confirmation or corroboration. This is a very major task and one which lies outside the scope of the present article. However, we can now consider one issue which arose in the 1960 and early 1970s and which we have ignored until now. That is the question of incommensurability. If 'incommensurable' is taken in the strong sense in which it is defined in the Collins English Dictionary, and if a new paradigm is incommensurable with an old one, then it would not be possible to compare the old and new paradigms as to their respective degrees of empirical confirmation. The whole project of defending the rationality of scientific revolutions in terms of confirmation theory would collapse. But is it really the case that in some or all scientific revolutions, the new paradigm is incommensurable with the old one in such a strong sense? This is something which we will investigate in the next sub-section.

3.4 The Incommensurability Problem

The concept of incommensurability (as applied to scientific theories) was introduced and developed by Feyerabend and Kuhn. This introduction was not independent, and indeed the concept emerged from discussions between them when they were both in the philosophy department in Berkeley in the years 1960 and

1961. Feyerabend in his 1970 has some interesting reminiscences of this period. He recalls that they had long discussions:

> Some of which were carried out in the now defunct *Café Old Europe* on Telegraph Avenue and greatly amused the other customers by their friendly vehemence. [1970, 198]

Later in the same article, Feyerabend says:

> I do not know who of us was the first to use the term 'incommensurable' in the sense that is at issue here. It occurs in Kuhn's '*Structure of Scientific Revolutions*' and in my essay 'Explanation, Reduction, and Empiricism' both of which appeared in 1962. [1970, 219]

Despite this joint origin of the concept of incommensurability, it was, as we shall see, developed in rather different ways by Feyerabend and Kuhn.

The issue of incommensurability has given rise to a great deal of often heated discussion among philosophers of science, and this discussion continues to the present day. So far in this section we have focussed on the historical approach to philosophy of science in the period form c. 1960 to the mid 1970s. This is a natural period to choose because interest in the historical approach to philosophy of science declined sharply after about 1975. There were a number of reasons for this, among which we can mention the sudden and unexpected death of Imre Lakatos from a heart attack on 2 February 1974, and the publication of Feyerabend's *Against Method* in 1975. Many philosophers of science saw this book of Feyerabend's as a *reductio ad absurdum* of the whole historical approach to philosophy of science. As a result of these and other factors, many of the discussions which we have described in this section petered out in the late 1970s. Discussions about incommensurability were an exception to this, because they continued unabated into the 1980s and 1990s. The reason for this is probably that the problem of incommensurability involved questions of language and meaning, and so fitted in with the linguistic philosophy which dominated the English speaking world. Ideas about language, meaning, translation, and reference which had been developed by Davidson, Kripke, Putnam and Quine were applied to the problem. As we shall see, neither Feyerabend nor Kuhn made much reference to language in their initial discussions of incommensurability, but Kuhn in his later period took a distinctly linguistic turn in accordance with the prevailing climate in philosophy at the time. Significantly in his 1983 Kuhn wrote:

> If I were now rewriting *The Structure of Scientific Revolutions*, I would emphasize language change more and the normal/revolutionary distinction less. [1983, 57]

The continuing interest in the problem of incommensurability is shown by a recent publication [Massimi, 2005] in which incommensurability is discussed in connection with the introduction of Pauli's Exclusion Principle.

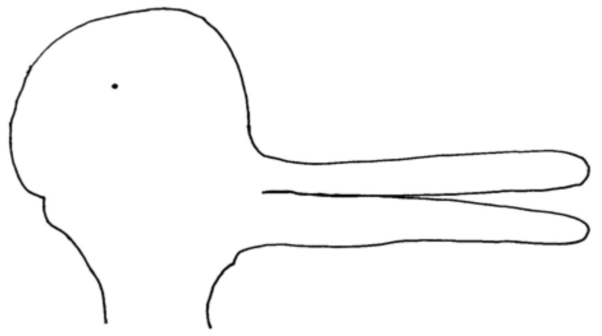


Figure 1. The Duck-Rabbit

Because of this situation, we will here not limit ourselves to discussions of incommensurability up to the mid-1970s, but take some account of how the debate has continued since then. The literature on this question is extensive and complicated, however, and we will only be able to discuss some of the arguments, choosing particularly those which relate to the question of the rationality of scientific change. The fundamental work for understanding the controversy as a whole is Sankey [1994] which gives a fine critical discussion of the most important works up to that date as well as Sankey's own contribution. Kuhn's contributions to the debate in the period from 1970 until his death in 1996 are conveniently collected in the 2000 volume: *The Road since Structure.*

Let us however begin by returning to the period c. 1960 to the mid 1970s and examine what Feyerabend and Kuhn said then about incommensurability. We will start with Kuhn. In his discussion of scientific revolutions in 1962, Kuhn claims that the new paradigm produced by such a revolution is incommensurable with the old paradigm. As he puts it:

> The normal-scientific tradition that emerges from a scientific revolution is not only incompatible but often actually incommensurable with that which has gone before. [1962, 102]

Kuhn's further discussions of incommensurability in his 1962 are not in terms of a change of language and meaning. He prefers to use the metaphor of the gestalt switch. Let us consider the duck-rabbit which is perhaps the most famous example of a gestalt switch because it was discussed by Wittgenstein in his [1953, 194]. Wittgenstein of course influenced Kuhn.

Figure 1 shows the duck-rabbit. The drawing can either be seen as a duck looking right or as a rabbit looking left. With a little practice, the viewer can make the 'gestalt switch' from seeing the drawing as a duck to seeing it as a rabbit at will. Kuhn compares the change of paradigm in a scientific revolution to a gestalt switch:

> ... at times of revolution, when the normal-scientific tradition changes, the scientist's perception of his environment must be re-educated — in some familiar situations he must learn to see a new gestalt. After he has done so the world of his research will seem, here and there, incommensurable with the one he had inhabited before. [1962, 111]

Kuhn does however make some qualifying remarks about the gestalt switch metaphor for he writes:

> That parallel can be misleading. Scientists do not see something *as* something else; instead they simply see it. ... In addition, the scientist does not preserve the gestalt subject's freedom to switch back and forth between ways of seeing. Nevertheless, the switch of gestalt, particularly because it is today so familiar, is a useful elementary prototype for what occurs in full-scale paradigm shift. [1962, 85]

Kuhn's reasons for regarding the parallel as misleading are not entirely convincing. The sentence S_1: 'The physicist sees the object as a tangent galvanometer' does not seem to differ greatly in meaning from the sentence S_2: 'The physicist sees a tangent galvanometer'. Moreover a scientist might preserve the gestalt subject's freedom to switch back and forth. Suppose a scientist first learns Newtonian mechanics and then Einsteinian mechanics. He or she might retain the capacity to switch back and forth between seeing the world as a Newtonian world and as an Einsteinian world.

We earlier criticized one use which Kuhn makes of the gestalt switch metaphor, namely his claim [1962, 121–2] that 'a new paradigm is born' through 'a relatively sudden and unstructured event like the gestalt switch.' We argued instead that new paradigms are born as the result of often long and painstaking work on a series of research programmes. However, once a new paradigm has been created, the process of switching from the old to the new paradigm can be compared quite accurately to a gestalt switch.

In his 1975, Feyerabend, like the early Kuhn, does not adopt a linguistic approach to the question of incommensurability. Indeed Feyerabend claims that it is not possible to give an explicit definition of incommensurability, and that the concept must be introduced by giving a number of different examples:

> As incommensurability depends on covert classifications and involves major conceptual changes it is hardly ever possible to give an explicit definition of it. ... The phenomenon must be shown, the reader must be led up to it by being confronted with a great variety of instances, and he must then judge for himself. [1975, 225]

Feyerabend follows the method by giving an interesting, and stimulating series of examples which are sometimes elaborated in considerable detail. These include the following: (1) the successive stages which, according to Piaget, children go through in their conception of the world [1975, 227–8], (2) the conceptual schemes of primitive tribes as compared with those of modern Westerners [1975, 249–51] — Feyerabend mentions particularly the studies of the Nuer by Evans-Pritchard, and (3) the change from the Homeric world-view to that of the Pre-Socratics [1975, 229–49 and 260–71]. In addition in his 1978, Feyerabend gives the striking example of the change brought about by waking up:

> ... waking up brings new principles of order into play and thereby causes us to perceive a waking world instead of a dream world ... [1978, 70]

We can see from this that incommensurability has for Feyerabend a general import, and is not restricted to changes brought about by scientific revolutions. However, he goes on to give examples of incommensurability in the scientific case as well. It is interesting to note, however, that Feyerabend is here somewhat more restrictive than Kuhn and denies that incommensurability was involved in the change from Ptolemy to Copernicus. He says:

> ... I never assumed that Ptolemy and Copernicus are incommensurable. They are not. [1975, 114]

Feyerabend's standard examples of incommensurability in science are the changes from Aristotelian mechanics to Newtonian mechanics, from Newtonian mechanics to relativistic mechanics, and from Newtonian mechanics to quantum mechanics [1975, 224–5 and 275–7].

Despite his reluctance to give a definition of incommensurability, Feyerabend comes close to doing so in his 1978. Significantly the passage occurs in a footnote. It is the following:

> ... mere *difference* of concepts does not suffice to make theories incommensurable in my sense. The situation must be rigged in such a way that the conditions of concept formation in one theory forbid the formation of the basic concepts of the other ... [1978, 68]

We can illustrate Feyerabend's idea here by considering the example of the transition from Aristotelian to Newtonian mechanics.[13] In Aristotelian mechanics every body in motion requires a force to move it along. If, for example, someone throws a stone, the thrower imparts to the stone a special force called 'impetus'. Impetus is then the force which continues to move the stone. The quantity of impetus gradually declines and the stone correspondingly ceases to move forward. Now at first sight we might think we could identify Aristotelian 'impetus' with Newtonian

[13]There are good discussions of this example in [Sankey, 1994, 88–89 and 109].

'momentum', but this would be a mistake. In Newtonian mechanics a body does not require a force to continue moving with uniform motion in a straight line. Momentum is a property of such a body but it is not a force moving it along. In fact there is no force in the Newtonian system moving the body along. The conditions of concepts formation in Newtonian mechanics forbid the formation of the Aristotelian concept of impetus.

Kuhn makes the point (e.g. in [1979, 205]) that the meaning of planet changed from Ptolemy to Copernicus. In Ptolemy's theory, the Sun and Moon were planets, but the Earth was not a planet. In Copernicus' theory, the Earth became a planet, but the Sun and Moon ceased to be planets. This is a significant change of meaning, but, as we have seen, it does not in Feyerabend's view lead to incommensurability. This is actually supported by Feyerabend's definition of incommensurability. We can define planet in the sense of Copernicus (or P_C) in terms of planet in the sense of Ptolemy (or P_P) as follows:

$$P_C(x) =_{def} [P_P(x) \;\&\; \neg \{(x = \text{Sun}) \vee (x = \text{Moon})\}] \vee (x = \text{Earth})$$

It follows therefore that the conditions of concept formation in Ptolemy's theory do not forbid the formation of Copernicus' concept of planet. As the converse also holds, this shows that according to Feyerabend's definition of incommensurabililty, the change in the concept of planet does not lead to an incommensurability between Ptolemy's theory and Copernicus'. It may have been generalising from examples such as this that led Feyerabend to the conclusion that Ptolemy's theory is not incommensurable with Copernicus'. This is a significant conclusion because it shows that dramatic revolutionary changes of theory can occur without incommensurability, and that consequently incommensurability is not an essential feature of scientific revolution.

As the whole idea of incommensurability has been severely attacked by many philosophers of science, let us now say something in its favour. The discussions of the notion by Kuhn and Feyerabend which we have just described do bring out the fact that scientific revolutions can cause profound conceptual changes, and alter very significantly the way in which scientists see the world. The nature of these changes are well-illustrated by the series of interesting analogies which Kuhn and Feyerabend provide: Gestalt switches, the change from sleeping to waking, Piaget's stages in child development, primitive tribes compared to modern Western societies, Homer's world-view compared to that of the Pre-Socratics. These analogies do, in our opinion, illuminate the nature of the profound changes in world-view which can be brought about by scientific revolutions. They are in danger of being forgotten if incommensurability is analysed purely in terms of language and meaning in the manner which we will describe in a moment.

So far we have talked of scientific revolutions bringing about a change in the way scientists see the world, but both Kuhn and Feyerabend sometimes make the stronger claim that in a scientific revolution the world itself changes. Thus Kuhn says:

> Nevertheless, paradigm changes do cause scientists to see the world of their research-engagement differently. In so far as their only recourse to that world is through what they see and do, we may want to say that after a revolution scientists are responding to a different world. [1962, 110]

This view is supported with greater zeal by Feyerabend, who writes:

> ... we certainly cannot assume that two incommensurable theories deal with one and the same objective state of affairs (to make the assumption we would have to assume that both at least *refer* to the same objective situation. But how can we assert that 'they both' refer to the same situation when 'they both' never make sense together? ... Hence, unless we want to assume that they deal with nothing at all we must admit that they deal with different worlds and that the change (from one world to another) has been brought about by a switch from one theory to another. [1978, 70]

This view has some points in common with Kant's. Kant thought that the intersubjective world of human beings is partly the result of the way things are in themselves, and partly the result of the conceptual schemes used to process sensory input. Kant, however, thought that these conceptual schemes are the same for all times and all human beings (Euclidean geometry and the twelve categories). Kuhn and Feyerabend allow the possibility that different communities, or the same community at different times, can have different fundamental conceptual schemes, and conclude that the members of different communities may inhabit different worlds.

Kuhn in his later period quite explicitly adopted such a modified Kantian position. He writes:

> By now it may be clear that the position I'm developing is a sort of post-Darwinian Kantianism. Like the Kantian categories, the lexicon supplies preconditions of possible experience. But lexical categories, unlike their Kantian forebears, can and do change, both with time and with the passage from one community to another. [1990, 104]

This could be called 'Kant on Wheels'[14]. Kuhn, however, was undecided as to whether he should include Kant's concept of the thing in itself in his post-Darwinian Kantianism. In an earlier formulation of his new version of Kantianism, he definitely repudiates things in themselves, writing:

> The view toward which I grope would also be Kantian, but without "things in themselves" and with categories of the mind which could change with time as the accommodation of language and experience proceeded. A view of that sort need not, I think, make the world less real. [1979, 207]

[14]This is the title of Lipton's 2003 article which discusses Kuhn's later views.

However, in 1990, the thing in itself is reinstated:

> Underlying all these processes of differentiation and change, there must, of course, be something permanent, fixed, and stable. But, like Kant's *Ding an sich*, it is ineffable, undescribable, undiscussible. Located outside of space and time, this Kantian source of stability is the whole from which have been fabricated both creatures and their niches, both the "internal" and the "external" worlds. [1990, 104]

The general question of Kantianism and realism is a fascinating one, but its study requires consideration of some complicated issues in the theory of reference, and, in particular, an evaluation of different versions of the causal theory of reference. We will therefore not discuss this question further here[15], but rather return to what is the central theme of this section namely the problem of the rationality of scientific revolutions.

As we have seen, Lakatos took Kuhn's view to be that the change of paradigm in a scientific revolution was irrational and analogous to a religious conversion. We have already given some quotations from Kuhn [1962] which support this interpretation, and here are a few more. Kuhn writes:

> The man who embraces a new paradigm at an early stage must often do so in defiance of the evidence provided by problem-solving. . . . A decision of that kind can only be made on faith. [1962, 157]

Moreover Kuhn hints that a scientific revolution may only seem to be an advance, because the victorious revolutionaries assert that an advance has been made:

> Revolutions close with a total victory for one of the two opposing camps. Will that group ever say that the result of its victory has been something less than progress? That would be rather like admitting that they had been wrong and their opponents right. To them, at least, the outcome of revolution must be progress, and they are in an excellent position to make certain that future members of their community will see past history in the same way. [1962, 165]

This looks very like the kind of 'might is right' doctrine which Lakatos was concerned to oppose.

After about 1970, the views of Kuhn and Feyerabend on this question begin to diverge. Feyerabend, as we have seen, took great pleasure in defending the thesis that science is irrational, and in proclaiming a 'farewell to reason'. Kuhn, on the contrary, drew back from the seemingly irrationalist implications of his earlier views, and seemed genuinely upset that his philosophy should have been taken up and developed in this sense. This divergence is already noticeable in the 1970 collection edited by Lakatos and Musgrave. As we have seen, Feyerabend in his contribution to this volume attacks Kuhn's views on normal science, but he stresses his agreement with Kuhn's views on incommensurability, writing:

[15]For the interested reader, there is an excellent discussion of these issues in Sankey [1994].

> With the discussion of incommensurability, I come to a point of Kuhn's philosophy which I wholeheartedly accept. [1970, 219]

Kuhn, however, in his 'Reflections on my Critics' in the same volume seems distinctly less enthusiastic about incommensurability. Perhaps he had already become worried by the interpretation which Lakatos had given to his views. At all events he writes:

> Such communication breakdown is important and needs much study. Unlike Paul Feyerabend (at least as I and others are reading him), I do not believe that it is ever total or beyond recourse. Where he talks of incommensurability *tout court*, I have regularly spoken also of partial communication, and I believe it can be improved upon to whatever extent circumstances may demand and patience permit, ... [1970, 232]

In 1976, Kuhn explicitly states that his earlier views on incommensurability had been misunderstood.

> Most readers of my text have supposed that when I spoke of theories as incommensurable, I meant that they could not be compared. But 'incommensurability' is a term borrowed from mathematics, and it there has no such implication. The hypotenuse of an isosceles right triangle is incommensurable with its side, but the two can be compared to any required degree of precision. What is lacking is not comparability, but a unit of length in terms of which both can be measured directly and exactly. In applying the term 'incommensurability' to theories, I had intended only to insist that there was no common language within which both could be fully expressed and which could therefore be used in a point-by-point comparison between them. [1976, 189]

Note that here Kuhn relates incommensurability to questions of language — something which he did not do in his 1962, and which is indicative of his linguistic turn. This linguistic approach is elaborated in his 1983 where he defines a 'modest version of incommensurability' in terms of the impossibility of translating some of the terms of one theory into the language of the other theory. This is how he puts it:

> Only for a small subgroup of (usually interdefined) terms and for sentences containing them do problems of translatability arise. The claim that two theories are incommensurable is more modest than many of its critics have supposed.
>
> I shall call this modest version of incommensurability 'local incommensurability'. ... The terms that preserve their meanings across a theory change provide a sufficient basis for the discussion of differences and for comparisons relevant to theory choice. [1983, 36]

Kuhn illustrates this by arguing that some of the terms of Newtonian mechanics cannot be translated into Aristotelian or Einsteinian mechanics (p. 44).

> ... Newtonian 'force' and 'mass' are not translatable into the language of a physical theory (Aristotelian or Einsteinian, for example) in which Newton's version of the second law does not apply. To learn any one of these three ways of doing mechanics, the interrelated terms in some local part of the web of language must be learned or relearned together and then laid down on nature whole. They cannot simply be rendered individually by translation.

This approach to incommensurability is not dissimilar to Feyerabend's 1978 definition which we described earlier.

Davidson and Putnam objected to the 'untranslatability' criterion on the grounds that we would not be able to understand an older theory unless we could translate its terms into our current language. However, Kuhn replied very reasonably to this objection that [1983, 39]: 'acquiring a new language is not the same as translating it into one's own.' Feyerabend made a similar reply to this objection pointing that we can learn a new language in the same way that children learn their first language.[16]

After these further clarifications of the concept of incommensurability, we will return to the key question of whether rational change from an old paradigm to a new incommensurable one is possible. Before doing so, however, it will be useful to look at one of the most interesting and original contributions to the discussion about incommensurability — Jane English's paper of 1978.

Many would think that there is no greater contrast among philosophers of science than that between Carnap on the one hand and Kuhn and Feyerabend on the other. Carnap is the extreme representative of the logical approach to philosophy of science. Nearly everything he considers is formalised in first order logic and a set of elaborate logical techniques is brought to bear upon it. Kuhn and Feyerabend on the other hand adopt the historical approach, basing their analysis on the history of science and proceeding informally without any use of formal logic. Despite these enormous differences, however, English argues that Carnap's partial-interpretation account of the meaning of theoretical terms has a very great deal in common with the views on meaning of Kuhn and Feyerabend and, in particular, gives rise to the same difficulties and counter-intuitive consequences. This is how she puts her thesis:

> Among the current views of the meaning of theoretical terms, Carnap's partial interpretation account and the meaning-change account of Kuhn and Feyerabend are usually thought of as antithetical. On the contrary, I will argue that the two have much in common. In particular, I will show that some of the major objections brought against

[16]For a fuller discussion, see [Sankey, 1994, 102–37].

> the meaning-change position apply equally to partial interpretation. [English, 1978, 57]

Let us therefore examine Carnap's partial interpretation account. Carnap of course begins his analysis of a scientific theory by formalising it in first order logic. He then decomposes the theory (TC say) into two parts. The first part contains the factual content of the theory, while the second part contains the meaning postulates. Each sentence of this second part is analytic, or true in virtue of meaning, and, in effect, the sentences of the second part, taken together, give an implicit definition of the theoretical terms of the theory. Carnap tried various ways of dividing a theory into these two parts, but finally decided on a method which uses the technical device of the Ramsey sentence. Give our theory TC, we form its Ramsey sentence R, and this represents the factual content of TC. The meaning postulates of TC are then given by $R \rightarrow TC$.

One interesting thing to note here is that although Carnap is employing the standard syntax of first order logic, he does not use the standard Tarskian semantics of first order logic. Let us illustrate Tarskian semantics by considering a mathematical example. Suppose we are dealing with a first order formalisation of Peano arithmetic. To give the formal symbols meaning using Tarskian semantics, we would first select a domain, which in this case would be the set N of natural numbers $\{1, 2, \ldots, n, \ldots\}$. Then to each of the individual constants of the theory we would assign a member of N. For example, there might be just one individual constant in this formalisation (a say) and we would assign to a the number 1. To each function letter in the formal theory we would assign a function over N. For example if there is a formal symbol $s()$, we might assign to $s()$ the function $+1$, so that $ss(a)$ would then stand for the number 3, and so on. To the 1-place predicate letters of the formal theory we would assign subsets of N. So to a predicate letter $O()$, we might assign the set of odd numbers, to a predicate letter $P()$, we might assign the set of prime numbers, and so on. In this way all the expressions of the formal theory are given meaning.

It is clear that Carnap does not use Tarskian semantics of this kind. He does not, for example, give 1-place predicate letters meaning by assigning to them subsets of a given domain, but rather by setting meaning postulates which implicitly define these 1-place predicates. His approach to meaning is in fact closer to that of Wittgenstein than to that of Tarski, for Carnap is in effect saying that the meaning of a predicate is given by the rules governing its use. It must be used in accordance with the meaning postulates. The fact that Carnap uses a Wittgensteinian approach may partly explain why his approach exhibits some strong resemblances to those of Kuhn and Feyerabend.

But why does Carnap, who was such a strong advocate of standard logic and of Tarski's ideas not use the standard Tarskian semantics? The answer is not far to seek, because a little reflection shows that it would be very difficult to apply Tarskian semantics to give a convincing account of meaning for formalised scientific theories. Such an application would result in a very artificial construction. Let us suppose, for example, that we have formalised Newtonian mechanics and are considering how to give meaning to the term $m(x)$ which is the formal equivalent of the mass of x. On a Tarskian approach, we might identify $m(x)$ with a function whose domain is the set of bodies and whose range is the set of positive real numbers. However, this definition of '$m(x)$' diverges completely form the way the term is in practice given meaning by physicists. Physicists explain the meaning of $m(x)$ to beginners by giving the laws governing masses, and the experimental and observational procedures used for determining the mass of a body. Without knowledge of these laws and procedures, the meaning of mass could not be grasped. Moreover it is not clear that the formal Tarskian approach is even coherent. It involves considering the set of bodies, for example, but is the concept of body clearly defined? It certainly is not. Let us consider an electromagnetic field, for example. A section of such a field would not normally be considered a body, but yet it could have a mass associated with it.

Our conclusion is that, while Tarskian semantics does indeed appear quite natural when dealing with mathematical examples, it seems, on the contrary, strained, artificial and inappropriate for handling the semantics of scientific theories. It is likely that this is why Carnap took a different approach when considering the question of the meaning of theoretical terms in a scientific theory. However, the result is interesting in a more general way. As we have seen, modern formal logic was developed by its pioneers (Frege, Peano, Russell, etc.) to handle mathematics. It was only later applied by the Vienna Circle and their followers to science. Now it may well be that many of the techniques of formal logic, while quite reasonable in the mathematical context, are inappropriate in a scientific context. This could be the reason for the appearance of some of the paradoxes which arose when applying formal logic to science, such as, for example, Hempel's paradox of the ravens.

Let us now, however, return to English's treatment of Carnap. Having explained Carnap's method of partial-interpretation of theoretical terms, she goes on to point out that it leads to some 'Kuhnian' consequences.

> Carnap's account here nicely supports Kuhn's meaning-change view. For instance, Kuhn relates in detail the history of the term 'compound'.[17] He claims that Dalton's assertion, "Compounds can be formed only in fixed proportions," and the pre-Daltonians' assertion, "Compounds can be formed in any ratios," did not contradict, because they meant different things by 'compound'. Dalton's predecessors included some of what we now call alloys, solutions, and suspensions under that term, whereas Dalton reserved it for things that follow his law. If we apply Carnap's method, Kuhn's interpretation results. Dalton is construed as saying, "If there is anything that ... and obeys the law of fixed proportions and ... then let us call it 'compound' ..." and his rivals are taken to say, "If there is anything that ... and combines in any ratios and ... then let us call *that* 'compound' ..." But then their theoretical statements "Compounds are formed only in

[17][Kuhn, 1962, 130–5]. This reference is given by English.

> fixed proportions" and "Compounds can be formed in any ratios" are both true; so they fail to contradict. [English, 1978, 70–71]

Indeed it would appear that Carnap's theory is more radical than Kuhn's, for as English says:

> This holism leads Carnap to an account of meaning change more extreme than Kuhn's. Since every postulate of the theory is represented in TC, any theoretical disagreement — not only disagreements in the most central assumptions — indicates a difference in meaning conventions. Although Kuhn has failed to specify how large a change must be to constitute a scientific revolution, he does hold that meanings are fixed despite small changes within "normal science." For Carnap, small changes as well as large are reflected in a change in the theory's Ramsey sentence, and thus in its meaning conventions. [1978, 71]

So on Carnap's account if a scientific theory T is changed even in a very slight way to produce a new theory T', then the terms of T' have different meanings form those of T. So no sentence of T' can contradict one of T. Thus the change from T to T' would appear to be an irrational leap of faith since T cannot be compared with T' to see which one is better confirmed by the evidence available. The problems of incommensurability seem to arise in a more extreme form in Carnap's account. Let us now see if they can be resolved.

Let us suppose then that we have two scientific theories T and T' — say Newtonian theory and Einsteinian theory. Since the theories are scientific, they will each contain a set of observation statements $\{O\}$ and $\{O'\}$. An observation statement is one whose truth-value, whether true or false, can in practice be decided by the scientific community on the basis of observation and experiment. Some philosophers of science maintain that there is a neutral observation language, but we will not make this assumption, which is anyway challenged by Kuhn and Feyerabend. We will assume to the contrary that the observations statements of T are made in the language of T, and those of T' in the language of T'. Thus if a particular observation statement is 'The mass of this body is 2.5 grams', we will assume that, within T, mass will be understood in a Newtonian sense yielding the observation statement O, while within T', mass will be understood in an Einsteinian sense yielding the observation statement O'. Now O and O' have different meanings, but, nonetheless, if we are dealing with an ordinary medium sized body moving with a low velocity, then the adherents of T' would certainly agree to give the same truth-value to O' as the adherents of T give to O on the basis of making the same observations and experiments. Thus these two observation statements would be ascribed the same truth-value by the two camps, a situation which we could describe by writing $O \sim O'$. Generalising we could establish a sequence of observation statements of T, $O_1, O_2, \ldots, O_n, \ldots$ say, and a corresponding sequence of observation statement of T', $O'_1, O'_2, \ldots, O'_n, \ldots$ say, such that $O_n \sim O'_n$. It now becomes easy to compare T and T' empirically. We work out how well T is confirmed (or disconfirmed) by the sequence $O_1, O_2, \ldots, O_n, \ldots$, and then how well T' is confirmed (or disconfirmed) by the sequence $O'_1, O'_2, \ldots, O'_n, \ldots$. If one of the two theories has a very much higher degree of confirmation than the other it becomes rational to accept it in preference to the other. No religious conversion, leap of faith, or political manoeuvring is needed here!

The same technique enables us to establish logical relations in an informal sense between T and T'. Suppose for example that T logically entails O_n and T' logically entails $\neg O'_n$, we can then say, speaking informally, that T contradicts T'. This can apply even if T and T' are formalised in two different systems S and S' within which the predicates of the two theories have different meanings. Of course if we work exclusively within the formal system S or within the formal system S', we cannot say that T and T' contradict each other. However the example of Gödel's incompleteness theorems surely shows that it is perfectly reasonable to extend reasoning outside a given formal system or formal systems, and to apply logic informally in this extension. The present examples shows that this technique should be used when applying logic to scientific theories, and that an exclusive reliance on say formal first order classical logic is not adequate for the philosophy of science.

Our conclusion then is that incommensurability is not such a monster threatening the rationality of scientific change as it might at first have appeared. On the contrary, it is quite easy to compare incommensurable theories both logically and empirically — provided one uses logic in a judicious fashion. But does this show that the question of incommensurability is not, after all, such an important one? Very different opinions have been expressed on this issue. Kuhn continued in his later period to believe in the importance of incommensurability. He wrote in 1990:

> No other aspect of *Structure* has concerned me so deeply in the thirty years since the book was written, and I emerge from those years feeling more strongly than ever that incommensurability has to be an essential component of any historical, developmental, or evolutionary view of scientific knowledge. [1990, 91]

Sankey on the other hand writes in the last few pages of his 1994 book: *The Incommensurability Thesis.*

> The overall thrust of my argument in this book is deflationary. Incommensurability is less of a problem than has generally been thought. The conceptual and semantical variance which initially gave rise to the idea of incommensurability do not threaten an unmitigated relativism of radically incompatible conceptual schemes. Nor do they force any concession upon an essentially realist view of the relation between scientific theory and extra-theoretic reality. ... there seems little point in saying that theories are incommensurable. [1994, 219 and 221]

On the whole we here side more with Sankey than with Kuhn, but would nonetheless like to make some observations in favour of the importance of incommensurability. There does seem some point in saying that two theories are incommensurable. This indicates that there is a radical conceptual shift in moving from one to the other. Moreover the study of such conceptual shifts has brought to light some interesting features of scientific change. It has shown that the logical analysis of science requires something more than the use of standard first order logic. The meaning of the theoretical terms in a scientific theory cannot be plausibly given using Tarskian semantics. If an alternative approach is adopted, such as Carnap's partial interpretation involving implicit definitions, then the study of the logical relations between two different scientific theories may well require considering both the formalisation of the theories in two different formal systems and then an examination of the relations between these formal systems. However the question of incommensurability is not simply one of logic and semantics. One of the most interesting features of the treatment of the question by Kuhn and Feyerabend was their stress on how incommensurable theories lead to different world views, and their attempts to illustrate the nature of such a change by a whole series of striking metaphors — gestalt switches, the change from sleeping to waking, the change from the Homeric world view to that of the pre-Socratics, and so on. These passages in Kuhn and Feyerabend are very insightful and illuminating, but unfortunately the dominance of the linguistic approach to philosophy often means that they are lost sight of. It is significant, for example, that English in her 1978 paper speaks of Kuhn's meaning-change view, and, in a passage already quoted cites [Kuhn, 1962, 130–5] as given an instance of this view (see our footnote 17). In this passage, however, Kuhn nowhere speaks of meaning. Instead he says things like the following:

> As a result, chemists came to live in a world where reactions behaved quite differently from the way they had before. [1962, 133]

Of course later in his life Kuhn did begin to speak of language, meaning, translation, etc. but this was because he too had fallen under the influence of the dominant linguistic paradigm in philosophy.

Another point in favour of the importance of the incommensurability problem is that many of the issues to which it has given rise could fruitfully be investigated further. The work of Kvasz (see his [1998; 1999; 2000]) is important here. We have argued that it is still worth speaking of theories or paradigms being incommensurable, because this indicates that there is a considerable conceptual change in passing from one theory or paradigm to the other. However the phrase 'a considerable conceptual change' is rather vague. Might there be conceptual changes of different magnitudes, and could we give some kind of classification of the size of these magnitudes? Kvasz takes up this problem in his 1999, which is concerned with the classification of scientific revolutions. He ingeniously suggests using perturbation theory as a device to measure the magnitude of the epistemic ruptures [1999, 219], and, as a result, comes up with three different kinds of epistemic rupture. Looking at the problem from a more linguistic point of view, we can say that a new theory or paradigm may well have a different language from the old theory or paradigm, but then the question arises of how the new language differs from the old, and how starting from the old language, a new language can be created. Kvasz tackles these problems in his 1998 and 2000. His 1998 is concerned with the history of geometry, and he shows that successive geometrical theories were expressed in different languages. He uncovers a mechanism by which a new language to express a new geometrical theory could be created. Following Wittgenstein in the *Tractatus*, Kvasz regards any language (L say) as having a form which is not expressible in the language. We can however incorporate the form of the language L into L thereby creating a new language L' say. Kvasz shows that this is precisely the way in which new languages for new geometrical theories were created, and in his 2000, he extends his results to mathematics as a whole. These investigations of Kvasz are closely connected to the issues which arose from the incommensurability problem, and his work shows that these issues can fruitfully be investigated further.

4 THE 1970S: SCIENCE AS PROBLEM SOLVING

4.1 Cognitive Science: The Emergence of a New Discipline

The Study of the Mind

The invention of computers brought forward a new dimension for the study of the mind. The study of cognition called for an integrated approach of theoretical as well as empirical disciplines, notably philosophy, psychology, linguistics, anthropology, neurosciences, and computer science.

The study of the mind was for a long time an exclusive topic for philosophy, going back to the Greeks and continuing into the 19^{th} century, when the beginnings of experimental psychology emerged. In the 20^{th} century, the gradual decline of the influence of behaviourism resulted in the preference of psychological theories taking into account mental representations and memory aspects. During the fifties, empirical results showed a limited capacity of the human mind, creating a point of contact between psychology and philosophy, for it introduced similar challenges to them. On the one hand, it opened new avenues to applied research, much of it focused on short-memory problems, and on the other, it introduced new epistemological questions, notably those having to do with the modelling of the generation and development of scientific knowledge.

Also during the fifties, but in this case between philosophy and computer science, representational theories of the mind emerged from the analogy of the mind as a computer [Turing, 1950; Fodor, 1975]. This idea served as a bridge between philosophy of mind and artificial intelligence (AI), the former providing the conceptual basis, the latter the tools to represent and manipulate knowledge.

The challenge of knowledge representation was at the core of all these disciplines and had logic as its main tool from the 50's to the 70's and 80's[18], when new logics emerged and proliferated in artificial intelligence. Moreover, the task of creating computational models of human intelligence put forward proposals such as the GPS (General Problem Solver), a program aiming at the mimicking of human problem solving. A task of such dimensions involved philosophy, psychology, and computer science, the constituent disciplines of cognitive science. A society of the cognitive sciences and a new journal emerged in the 70's and this kind of interdisciplinary research began to evolve. Pioneers such as John McCarthy, Marvin Minsky, Allen Newell, and Herbert Simon are the founding fathers of artificial intelligence. In addition, Noam Chomsky rejected behaviourist assumptions about language as a learned habit and proposed instead to explain language comprehension in terms of mental grammars consisting of rules [Chomsky, 1972; 1976]. All these researchers are to be regarded as the key founders of the field of Cognitive Science.

As far as the impact of computers on mainstream philosophy of science, although they had been around for some time, they are hardly mentioned before the seventies, as judged by the writings of Popper, Kuhn, Lakatos and Feyerabend. However, scientific knowledge and discovery constituted a substantial object of research for computational models of intelligent behaviour. Cognitive Scientists imported some of the problems of philosophy of science. An example of an interaction between computer science, artificial intelligence, logic and philosophy of science, is the development of machine learning, which casts doubt on Popper's claim that "induction is a myth", since computer programs began to be able to carry out induction successfully in some cases. Abduction was also studied as a form of explanatory reasoning, and new forms of computational representations and processes were devised for such an inference. More generally, there were considerable developments in logic with new logical systems such as non-monotonic logics (several of these will be analyzed in the next section). This meant that the logical approach to philosophy of science could be revived with a more powerful set of tools. Moreover, the scope of the logical approach could be extended to include discovery, though perhaps even at present this should be confined to discovery within a normal science context.

4.2 Philosophy of Science: Background Issues

This section aims firstly to set the scene for a refreshed view of the task of philosophy of science, that of analyzing science as a problem solving process, the guiding idea for a renewed enterprise in the seventies. Under this view, the analysis of scientific knowledge is addressed by questions having to do with the growth and evolution of knowledge, with the progress of science and the discovery and de-

[18]In the 1980's, a new paradigm emerged, namely connectionism and its companion neural networks. There was a change in the logical approach as well about that time. This was the appearance of probability (especially with the development of Bayesian networks). Most advocates of the logical approach these days would include probability, and interestingly the opposing connectionist approach also uses probability. The decade of 1990's is marked as the so called "antirepresentationalism" age. However, logic remained extensively used.

velopment of new theories, rather than centred in the fundamental concepts by themselves. The guiding principle of science as problem solving is however by no means genuinely new, for it pertains to marginal views in the philosophy of science, which up to the seventies, had no privileged place in the received view.

This view constitutes the challenge to broaden the scope of philosophy of science, but it is not until the 90's that issues like scientific discovery are decidedly in its research agenda. As is well-known, great philosophers and mathematicians have been brilliant exceptions in the study of discovery and development in science, and that their non conventional contributions to this field, although great inspirations, have not set new paradigms in the methodology of science. Therefore, before describing the proposals which shape the move to problem solving in the 70's, let us review the work of some of the main predecessors.

As far as logic goes in the seventies, the formal advances which shape the task of philosophy of science up to the 50's, that of giving a logical analysis of the concepts of science, were for some pretty much logically exhausted, and were for all often obscured by the historical analysis in vogue from the sixties. The logic used up to that date was fundamentally classical logic, which cannot help to account for the infallibility and the dynamics of scientific knowledge. It was not until the developments of logics in artificial intelligence in the eighties, that these new tools were exploited. However, there are important antecedents of work in logic which are more in accord with the new task of science, which is definitively marked by a new conception of logic altogether. Two figures from the turn of the 19^{th} century are worthy of mention here, namely Bernard Bolzano (1781-1848) and Charles S. Peirce (1839-1914). As for authors in the 20^{th} century, we will mention Polya, Hanson and Popper. We will illustrate some aspects of their proposals which will be relevant for our later discussion.

Bernard Bolzano

In his "Wissenschaftslehre" [1837], Bolzano engaged (among other things) in the study of different varieties of inference. One of Bolzano's goals was to show why the claims of science form a theory as opposed to an arbitrary set of propositions. For this purpose, he defines his notion of deducibility as a logical relationship extracting conclusions from premises forming *compatible propositions*, those for which some set of ideas make all propositions true when uniformly substituted throughout. In addition, compatible propositions must share *common ideas*. Restated in model-theoretic terms, Bolzano's notion of deducibility reads as follows (cf. [van Benthem, 1984]):

$T, C \Rightarrow E$ if

(1) The conjunction of T and C is consistent.

(2) Every model for T plus C verifies E.

Therefore, Bolzano's notion may be seen (anachronistically) as Tarski's consequence plus the additional condition of consistency. Bolzano does not stop here. A finer grain to deducibility occurs in his notion of *exact deducibility* which imposes greater requirements of 'relevance'. A modern version, may be transcribed (again, with some historical injustice) as:

$T, C \models^+ E$ if

(1) $T, C \models E$

(2) There is no proper subset of T, T', or of C, C', such that $T', C' \models E$.

That is, the premise set (composed by T, C) must be 'fully explanatory' in that no subpart of it would do the derivation[19]. Bolzano's agenda for logic is relevant to the study of general non-monotonic consequence relations for several reasons. It suggests the methodological point that what we need is not so much proliferation of different logics as a better grasp of different styles of consequence.

Charles S. Peirce

Now let's turn to Charles S. Peirce. He proposed abduction to be the logic for synthetic reasoning, that is, a method to acquire new ideas. He was indeed the first philosopher to give to abduction a logical form, on a pair with deduction and induction. His formulation is reproduced as follows [Peirce 1931-1935, 5.189]:

The surprising fact, C, is observed.

But if A were true, C would be a matter of course.

Hence, There is reason to suspect that A is true.

For Peirce, three aspects determine whether a hypothesis is promising: it must be *explanatory*, *testable*, and *economic*. A hypothesis is an explanation if it accounts for the facts. Its status is that of a suggestion until it is verified, which explains the need for the testability criterion. Finally, the motivation for the economic criterion is twofold: a response to the practical problem of having innumerable explanatory hypotheses to test, as well as the need for a criterion to select the best explanation amongst the testable ones. The above formulation accounts for the explanatory aspect.

George Polya

The next reference, already within the 20^{th} century, is G. Polya [1962], regarded as the modern founder of heuristics [Hintikka and Remes, 1974; 1976]. He analyzed mathematical problems and their relation to discovery. In the context of number theory, for example, a general property may be guessed by observing some relation as in:

[19]Notice that this leads to non-monotonicity (cf. 5.1). A consequence $\models$ is non-monotonic whenever $T \models b$ does not ensure $T, a \models b$. That is, the addition of new premises is no warrant for validity reservation. Here is an example: $T, a \rightarrow b,\ a \models^+ b$, but it is not the case that T, $a \rightarrow b,\ a,\ b \rightarrow c \models^+ b$.

$$3 + 7 = 10 \qquad 3 + 17 = 20 \qquad 13 + 17 = 30$$

Notice that the numbers 3,7,13,17 are all odd primes and that the sum of any of two of them is an even number. An initial observation of this kind eventually led Goldbach (with the help of Euler) to formulate his famous conjecture: '*Every even number greater than two is the sum of two primes*'. Moreover, Polya contrasts two types of arguments. A demonstrative syllogism in which from $A \Rightarrow B$, and B false, $\neg A$ is concluded, and a heuristic syllogism in which from $A \Rightarrow B$, and B true, it follows that A is more credible. The latter, of course, recalls Peirce's abductive formulation.

Russell Hanson and Karl Popper

Already in the 60's, an author emphasizing explanation as a process of discovery is Hanson [1961], who gave an account of patterns of discovery, recognizing a central role for retroduction (another name for abduction). Another intellectual inheritance from the past decade is the work of Popper in Conjectures and Refutations in 1963. In this work, the growth of scientific knowledge is the most important of the traditional problems of epistemology [Popper, 1934, 22]. His fallibilist position provided him with the key to reformulate the traditional problem in epistemology, which was focused on the reflection on the sources of our knowledge. Rather, for him, the focus should be on the advancement of knowledge. This concern is intimately related to his view of science as a problem solving activity: '*Science should be visualized as progressing from problems to problems — to problems of increasing depth. For a scientific theory — an explanatory theory — is, if anything, an attempt to solve a scientific problem, that is to say, a problem concerned with the discovery of an explanation*' [Popper, 1960, 179]. As we shall see, this view is in accord with Simon's famous slogan that *scientific reasoning is problem solving* made in research in cognitive psychology and artificial intelligence (to be later introduced). However, in regard to giving a logical account for discovery processes, Popper's position is broadly recognized as neglecting this kind of scientific practice as part of the methodology of science agenda, and rather regarding its study a business of psychology. (But as we shall see, under a broad view of discovery[20], Simon's and Popper's positions are not so far apart.).

4.3 Philosophy of Science: Discovery as Problem Solving

The work in the sixties most relevant to our discussion is the work by Lakatos (1963-4), namely *Proofs and Refutations*, a critical response to Popper's logic of scientific discovery:

[20]Here is a useful distinction between a narrow and a broad view of discovery. While the former view regards issues of discovery as those dealing exclusively with the initial conception of an idea, the latter view is that which deals with the overall process going from the conception of a new idea to its settlement as an idea subject for ultimate justification [Laudan, 1980].

> *There is no **infallibilist** logic of scientific discovery leading infallibly to results, but there is a fallibilistic logic of discovery which is the logic of scientific progress. But Popper, who has laid the basis for this logic of discovery was not interested in the meta-question of what is the nature of this investigation, so he did not realize that it is neither psychology nor logic, but an independent field, the logic of discovery, **heuristics*** [Lakatos, 1963-4, 167, our emphasis].

It is interesting to note that Lakatos was greatly inspired by the history of mathematics, paying particular attention to processes that created new concepts — often referring to G. Polya, one of his predecesors. Another key reference for the view of science as problem solving is the work of Laudan [1977], namely "Progress and its Problems", in which scientific progress is analyzed as a case of naturalization of science and in its relation to history as well as with problems having to do within the rationalist view of science. Still another important reference is the work of Rescher [1978], which introduces a *direction of thought*. Interestingly, this establishes a temporal distinction between 'prediction' and 'retroduction', by marking the precedence of the explanandum over the hypothesis in the latter case.

Does Scientific Discovery Have a Logic?

In principle, the pioneering work of Herbert Simon and his team shares the ideal on which the whole enterprise of artificial intelligence was initially grounded, namely that of constructing intelligent computers behaving like rational beings. In his essay *Does scientific discovery have a logic*? Simon sets himself the challenge to refute Popper's general argument, reconstructed for his purposes as follows: *'If "There is no such thing as a logical method of having new ideas", then there is no such thing as a logical method of having **small** new ideas'* [Simon, 1973, 327, my emphasis], and his strategy is precisely to show that an antecedent in the affirmative does not commit to an assessment of the consequent, as Popper seems to suggest. Thus, Simon converts the ambitious aim of searching for a logic of discovery revealing the process of discovery at large, into an unpretentious goal: *'Their modesty [of the examples dealt with] as instances of discovery will be compensated by their transparency in revealing the underlying process'* [Simon, 1973, 327].

This humble and brilliant move allows Simon to further draw distinctions on the type of problems to be analysed and on methods to be used. For Simon and his followers, scientific discovery is a problem-solving activity. To this end, a characterization of problems into those that are well structured versus those that are ill structured is provided, and the claim for a logic of discovery focuses mainly on the well-structured ones[21].

[21] A well structured problem is one for which there is a definite criterion for testing, and for which there is at least one problem space in which the initial and the goal state can be represented and all other intermediate states may be reached with appropriate transitions between them. An

Although there is no precise methodology by which scientific discovery is achieved, as a form of problem solving, it can be pursued via several methodologies. The key concept in all this is that of *heuristics*, the guide in scientific discovery which is neither totally rational nor absolutely blind. Heuristic methods for discovery are characterized by the use of selective search with fallible results. That is to say, while they provide no complete guarantee to reach a solution, the search in the problem space is not blind, but it is selective according to a predefined strategy. The authors further distinguish between 'weak' and 'strong' methods of discovery[22].

4.4 Simon and Popper Revisited

When confronting the work of two philosophical giants, we compare their views on inquiry in science, in order to explore to what extent their stances are close together. On the one hand, although Popper [1934] was genuinely interested in an analysis of new ideas in science, he rendered the very first process of the conception of an idea to be outside the boundaries of the methodology of science, and centered his efforts in giving an account of an ensuing process, concerned with the methods of analyzing new ideas logically, and accordingly produced his method of conjectures and refutations. Simon's [1973] aim was to simulate scientific discovery at large, giving an account both for the generation and evaluation of scientific ideas, convinced that the way to go was to give both an empirical and a normative account of discovery. The former to describe and then represent computationally the intellectual development of human discoveries made in science. The latter to provide prescriptive rules, mainly in the form of heuristic strategies to perform scientific discoveries.

Both authors hold a fallibilist position, one in which there is no certainty of attaining results and where it is possible to refute already assessed knowledge, in favour of new one that better explains the world. However, while for Popper there is one single method for scientific inquiry, the method of conjectures and refutations, for Simon there are several methods for scientific inquiry, for the discovery and justification processes correspond to several heuristic strategies, largely based on pattern seeking, the logic of scientific discovery. A further difference between these approaches is found in the method itself for the advancement of science, in what they regard to be the 'logic' for discovery. While for Popper ideas are generated by the method of blind search, Simon and his team develop a full theory to support the view that ideas are generated by the method of 'selective search'. Clearly the latter account allows for a better understanding of how theories and

ill-structured problem lacks at least one of the former conditions.

[22]The former is the type of problem solving used in novel domains. It is characterized by its generality, since it does not require in-depth knowledge of its particular domain. In contrast, strong methods are used for cases in which our domain knowledge is rich, and are specially designed for one specific structure. Weak methods include generation, testing, heuristic methods, and means-ends analysis, to build explanations and solutions for given problems.

ideas may be generated. Whether any of these methods corresponds to natural phenomena or rather belongs to the province of the artificial, is another question.

Popper's and Simon's approaches are close together, at least in so far as the following basic ideas are concerned: they both hold a fallibilist stance in regard to the well-foundedness of knowledge and view science as a dynamic activity of problem solving, in which the growth of knowledge is the main aspect to characterize, as opposed to the view of science as an static enterprise in search of the assessment of theories as true. But Popper failed to appreciate the philosophical potential of a normative theory of discovery, for he was blinded to the possibility of devising a logic for the development of knowledge. His view of logic remained static: *'I am quite ready to admit that there is a need for a purely logical analysis of theories, for an analysis which takes no account of how they change and develop. But this kind of analysis does not elucidate those aspects of the empirical science which I, for one, so highly prize'* [Popper, 1934, 50].

One reason that allows for the convergence of these two accounts, perhaps obvious by now, is that neither the "Friends of discovery" really account for the epistemics of creativity at large nor Popper neglects its study entirely. Both accounts fall naturally under the study of discovery — when a broad view is endorsed (cf. footnote 20) — and neither of them rejects the context of justification, or any other context for that matter. Therefore, it seems that when the focus is on the processes of inquiry in science, rather than on the products themselves, any possible division of contexts of research is doomed to fail sooner or later.

4.5 Logic: Logic for Problem Solving

The introduction of computers had a profound impact on logical research, to the extent of providing a new paradigm, that of viewing logic in a goal directed way. This idea led Bob Kowalski to propose 'logic programming' (together with Alain Colmerauer) in the early seventies. In his own words: "*The fundamental thesis of LP is that appropriate forms of logic can serve as a high level programming language*" [Kowalski, 1994, 38].

Logic programming is inspired by first-order logic, and it consists of logic programs, queries, and an underlying inferential mechanism known as resolution. It is implemented in (amongst others) the programming language Prolog. Roughly speaking, a Prolog program P is an ordered set of rules and facts. Rules are restricted to clause form:

$$A \Leftarrow L_1, \ldots, L_n$$

which contains one atom (A) in its consequent and a set of literals in its antecedent[23]. A is called the head and $L_1, \ldots, L_n$ is called the body of the program clause. A query q (theorem) is posed to program P to be solved (proved). If the query follows from the program, a positive answer is produced, and so the query

[23]An atom is an atomic formula. A literal is an atom or the negation of an atom.

is said to be successful. Otherwise, a negative answer is produced, indicating that the query has failed. However, the interpretation of negation is 'by failure'. That is, 'no' means 'it is not derivable from the available information in P — without implying that the negation of the query $\neg q$ is derivable instead. Resolution is an inferential mechanism based on refutation working backwards: from the negation of the query to the data in the program. In the course of this process, valuable by-products appear: the so-called 'computed answer substitutions', which give more detailed information on the objects satisfying given queries.

Kowalski's 1979 book "Logic for Problem Solving", was a key source which made logicians and theoretically-oriented computer scientists start to talk to one another. While logicians were at first shocked by a formal language sensitive to rule order and in demand of strange things like the "occur check" to warrant metalogical properties, computer scientists were having a hard time to digest a declarative programming language, one in which there was no distinction between data and program, as well as with the highly demanding logical rigour of a program specification. But gradually the idea of logic programming set in the curricula and soon thereafter, new courses were given and new research was carried out.

5 THE 1980S AND 1990S: LOGICAL AND COMPUTATIONAL MODELS FOR SCIENTIFIC INFERENCE AND DISCOVERY

5.1 Artificial Intelligence and Logic: Non-monotonic Logics

The invention of computers naturally brought forward the challenge to represent knowledge in a systematic manner, and in turn, confronted logicians and computer scientists with the problem of the formalization of non-monotonic reasoning, broadly conceived as the reasoning to conclusions on the basis of incomplete information. Just as in most of scientific reasoning, given more information, one must be capable to retract previously drawn inferences.

An early attempt to formalize this type of reasoning as part of a computer's reasoning mechanism, was proposed by John McCarthy in the 1970's; but it was until the 1980's when a proliferation of non-monotonic logics was at the core of logical and computational research. A multitude of logical systems was proposed, varying both in their logical approach (syntactic, semantic) as well as in their computational particular application. These applications concern three types of problems, namely puzzles and deductive databases (DB), default reasoning, and explanation-based reasoning.

The first of these applications points to one major problem of knowledge representation in a database, that is, the way to treat the ontological status of existing information, something which led to the assumption that it contains all relevant and true information needed to reason about. A prominent proposal was the *closed-world assumption* (CWA) (cf. [Brewka *et al.*, 1997]), aiming to capture that all of the non-given information is taken to be false[24]. Still, implicit information

[24]CWA (DB) = DB $\cup\{\neg P(t)|$ DB $\not\models P(t)\}$, where $P(t)$ is a ground predicate instance (a

found in almost all commonsense reasoning puzzles (such as "the only way across the river is by the boat" in the famous missionaries and cannibals puzzle), was in need of explication. In the framework of second order logic, McCarthy [1980] proposed a solution known as *circumscription*. These two solutions however, lay outside the realm of classical logic and are rather committed to one general class of non-monotonic formalisms, namely model preference logics. These systems have the property of giving a characterization of logical consequence based on a (frequently defined in advance) class of preferred models, in which positive facts are minimized.[25]

The second of these applications shaped a prominent proposal for non-monotonic logics, namely *default logic*, put forward by Reiter [1980]. This formalism is based on the notion of a default, a *prima facie* justification of a conclusion, meaning that the inference is drawn on the basis of available information, likely to be defeasible in the presence of later conflicting information. More precisely, there is an initial set of defaults, validating all new consequences consistently generated. This idea was commonly formalized as a fixed-point equation, and accordingly these logics were referred to as *fixed-point logics* or *consistently-based logics*. It is worth mentioning that a special issue of the journal *Artificial Intelligence* (vol. 13, numbers 1 and 2, 1980) was devoted to these and other new formalisms[26].

As for the third computer application, devoted to explanation-based reasoning in problems involving diagnosis, it directly points to abductive reasoning, roughly defined as the reasoning from an observation to its possible explanations. As stated in the previous section, this type of reasoning was first prompted as such by Charles Peirce (1839-1914). His logical formulation (cf. 4.2), has played a key role in Peirce scholarship, and it has been the point of departure of many classic studies on abductive reasoning in artificial intelligence [Reggia *et al.*, 1985], such as in logic programming [Kakas *et al.*, 1995], knowledge acquisition [Kakas and Mancarella, 1990], and natural language processing [Hobbs *et al.*, 1990]. Nevertheless, these approaches have paid little attention to the elements of this formulation and none to what Peirce said elsewhere in his writings. In this field, the formulation has been generally interpreted as the following logical argument-schema:

$$\begin{array}{l} C \\ A \rightarrow C \\ \hline A \end{array}$$

where the status of A is tentative (it does not follow as a logical consequence from

ground term or predicate is that containing no variables). That is, if a ground term cannot be inferred from the database, its negation is added to the closure. Cf. [Reiter, 1987] and [Brewka *et al.*, 1997].

[25] A later proposal in this direction was Shoham's [1988] notion of causal and default reasoning, which introduces a preference order on models, requiring that only the most preferred models of the premises be included in the models of the conclusion. And this again contrasts with Tarskian classical consequence, in which it is required that all models of the premises are included in the models for the conclusion.

[26] Cf. This subsection was largely based upon [Brewka *et al.*, 1997], in which a much more in-depth analysis is to be found.

the premises). However intuitive, this interpretation certainly captures neither the fact that C is surprising nor the additional criteria Peirce proposed (cf. 4.2). The additional Peircean requirements of testability and economy are not recognized as such in AI, but are nevertheless to some extent incorporated. Economy is carried out as a further selection process to produce the best explanation, since there might be several formulae that satisfy the above formulation but are not appropriate as explanations. As for the testability requirement, when the second premise is interpreted as logical entailment this requirement is trivialized, since given that C is true, in the simplest sense of 'testable', A will always be testable.

Axiomatic Theory of Consequence Relations

The proliferation of non-monotonic systems brought forward still another challenge to logicians, this time pointing to a methodological as well as to a demarcation question. On the one hand, there was a need for a common framework in order to analyze and compare all these new systems; but at the same time, there was an urgency to put some order, and establish the limits to the kind of systems accepted as logical. Logic had gone out of its mathematical domain with apparently no rules and clear cut ends:

> "*In an attempt to put some order in what was then a chaotic field, Gabbay asked himself what minimal properties do we require of a consequence relation $A_1, \ldots, A_n \vdash B$ in order for it to be considered as a logic. In his seminal paper [1985] he proposed the following:*
>
> *Reflexivity*: $\Delta, A \vdash A$
>
> *Restricted Monotonicity:* $\dfrac{\Delta \vdash A \qquad \Delta \vdash B}{\Delta, A \vdash B}$
>
> *Cut:* $\dfrac{\Delta, A \vdash B \quad \Delta \vdash A}{\Delta \vdash B}$
>
> *The idea is to classify non-monotonic systems by properties of their consequence relations. Kraus–Lehman–Magidor [1990] developed preferential semantics corresponding to various additional conditions on $\vdash$ and this has started the area now known as the axiomatic approach to non-monotonic logics".* [Ohlbach and Reyle, 1999]

This type of analysis started with Dana Scott [1971], and was inspired in the early works of logical consequence by Tarski and those of natural deduction by Gerard Gentzen [1934]. It describes a style of inference at a very abstract structural level, giving its pure combinatorics. The basic idea of an structural analysis (as it is also known) is the following: *A notion of logical inference can be completely characterized by its basic combinatorial properties, expressed by structural rules.* Structural rules are instructions which tell us, e.g., that a valid inference remains valid when we insert additional premisses ('monotonicity'), or that we may safely chain valid inferences ('transitivity' or 'cut'). The general format is that of logical sequents, with a finite sequence of premisses to the left, and one conclusion to the right of the sequent arrow ($\Delta \Rightarrow B$).

As already mentioned, this type of analysis has proved very successful in artificial intelligence for studying different types of plausible reasoning [Kraus *et al.*, 1990], and indeed as a general framework for inference, including non-monotonic consequence relations [Gabbay, 1985]. Another area where it has proved itself is dynamic semantics, where not one but many new notions of dynamic consequences are to be analyzed [van Benthem, 1994; 1996]. This new framework served to analyze and compare many proposed logical systems. One important contribution is that it goes beyond the view of classifying a set of logical systems for what they fail to validate — not surprisingly were labelled *non-monotonic logic* — and rather looks in a positive way for the properties that they do observe. However, the claim that these three specific rules proposed by Gabbay are valid in every system was refuted.[27]

Theory Change

When talking about theory change, an obvious related territory is found in theories of belief change in AI, mostly inspired by the work of Gärdenfors [1988], a work whose roots lie in the philosophy of science. These theories describe how to incorporate a new piece of information into a database, a scientific theory, or a set of common sense beliefs.

Given a consistent theory θ, called the belief state, and a sentence φ, the incoming belief, there are three *epistemic attitudes* for θ with respect to φ: either φ is accepted ($\varphi \in \theta$), φ is rejected ($\neg\varphi \in \theta$), or φ is undetermined ($\varphi \notin \theta$, $\neg\varphi \notin \theta$). Given these attitudes, three main operations may incorporate φ into θ, thereby effecting an epistemic change in our currently held beliefs:

> Expansion ($\theta + \varphi$)
> An accepted or undetermined sentence φ is added to θ.

[27]Ten years or so later on, Gabbay himself [1994] acknowledged the following: "*although some classification was obtained and semantical results were proved, the approach does not seem to be strong enough. Many systems do not satisfy restricted monotonicty. Other systems such as relevance logic, do not satisfy even reflexivity. Others have richness of their own which is lost in a simple presentation as an axiomatic consequence relation. Obviously, a different approach is needed, one which would be more sensitive to the variety of features of the systems in the field*". As is well-known in this field, Gabbay then moved to propose his *Labelled Deductive Systems*, certainly a much more robust framework for logical systems.

Still, the question remains in regard to what extent we may use the axiomatic theory of consequence relations as an attempt to provide a logical criterion of demarcation, and this seems to be a fertile area to explore. In [Aliseda, 2005b], the suggestion is that rather than aiming at an specific set of minimal rules, we should be asking for a *minimal schema set of structural properties* a system should satisfy to be considered a logical one. The particular proposal is that this schema set must consist of some forms of monotonicity, transivity or cut, and of reflexivity. And these forms, of course, need not be the same ones as those for classical logic.

Contraction ($\theta - \sigma$):
Some sentence σ is retracted from θ, together with enough sentences implying it.

Revision ($\theta^* \varphi$):
In order to incorporate a rejected φ into θ and maintain consistency in the resulting belief system, enough sentences in conflict with φ are deleted from θ (in some suitable manner) and only then φ added.

Of these operations, revision is the most complex one. It may indeed be defined as a composition of the other two. First contract those beliefs of θ that are in conflict with φ, and then expand the modified theory with sentence φ. While expansion can be uniquely defined, this is not so with contraction or revision, as several formulas may be retracted to achieve the desired effect. These operations are intuitively non-deterministic. The contraction operation per se cannot state in purely logical or set-theoretical terms which of the available formulae should be chosen. Therefore, an additional criterion must be incorporated in order to fix which formula to retract. Here, the general intuition is that changes on the theory should be kept 'minimal', in some sense of informational economy. Various ways of dealing with the latter issue occur in the literature[28].

In practice, however, full-fledged AI systems of belief revision can be quite diverse. Here are some aspects that help to classify them:

Representation of Belief States
Operations for Belief Revision
Epistemological Stance

Regarding the first, we find there are essentially three ways in which the background knowledge θ is represented: (i) belief sets, (ii) belief bases, or (iii) possible world models[29]. As for the second aspect, operations of belief revision can be given either constructively or via 'postulates'[30].

[28]We mention only that in [Gärdenfors, 1988]. It is based on the notion of *entrenchment*, a preferential ordering which lines up the formulas in a belief state according to their importance. Thus, we can retract those formulas which are 'least entrenched' first.

[29]A belief set (i) is a set of sentences from a logical language L closed under logical consequence. In this classical approach, expanding or contracting a sentence in a theory is not just a matter of addition and deletion, as the logical consequences of the sentence in question should also be taken into account. The second approach (ii) emerged in reaction to the first. It represents the theory θ as a *base for a belief set* B_θ, where B_θ is a finite subset of θ satisfying $Cons(B_\theta) = \theta$. (That is, the set of logical consequences of B_θ is the classical belief state). The intuition behind this is that some of the agent's beliefs have no independent status, but arise only as inferences from more basic beliefs. Finally, the more semantic approach (iii) moves away from syntactic structure, and represents theories as sets W_θ of possible worlds (i.e., their models). Various equivalences between these approaches have been established in the literature (cf. [Gärdenfors and Rott, 1995]).

[30]The former approach is more appropriate for algorithmic models of belief revision, the latter serves as a logical description of the properties that any such operations should satisfy. The two can also be combined. An algorithmic contraction procedure may be checked for correctness according to given postulates. [Say, one which states that the result of contracting θ with φ should be included in the original state ($\theta - \varphi \subseteq \theta$.)].

Finally, each approach takes an 'epistemological stance' with respect to justification of the incoming beliefs. Here are two major paradigms. A 'foundationalist' approach argues one should keep track of the justification for one's beliefs, whereas a 'coherentist' perspective sees no need for this, as long as the changing theory stays consistent and keeps its overall coherence.

Therefore, each theory of epistemic change may be characterized by its representation of belief states, its description of belief revision operations, and its stand on the main properties of belief one should be looking for[31]. In particular, the theory proposed in Gärdenfors [1988], known as the AGM paradigm after its original authors [Alchourrón, Gärdenfors and Makinson, 1985], represents belief states as theories closed under logical consequence, while providing 'rationality postulates' to characterize the belief revision operations, and finally, it advocates a coherentist view. The latter is based on the empirical claim that people do not keep track of justifications for their beliefs, as some psychological experiments seem to indicate (cf. [Harman, 1986]).

5.2 Cognitive Science and Philosophy of Science: Computational Philosophy of Science

Concrete and quite articulated computer programs of scientific discovery are found in the late eighties in the work of Simon and his team [Langley *et al.*, 1987]. These are the BACON system, one which simulates the discovery of quantitative laws in Physics (such as Kepler's laws and Ohm's law) and the GLAUBER program, which simulates the discovery of qualitative laws in Chemistry. In the same spirit, Paul Thagard proposes a new field of research, namely *Computational Philosophy of Science* [Thagard, 1988], an integrated approach of psychology, history and philosophy of science, all of it directed to questions of scientific discovery, having to do with its cognitive patterns, its place and time in the history of science and with core notions in the philosophy of science (such as explanation, confirmation, falsification, evaluation, induction, abduction and theory revision). Thagard's proposal, for example, puts forward the computational program PI (Processes of Induction) to model some aspects of scientific practice, such as concept formation and theory building. The general idea consists of the solution of a problem as a "match" between an initial and a final state. And when there is no match, several kinds of induction may be performed (generalization, abduction, concept formation, etc...).

We may identify at least two principles (1 and 2 below) and three claims (given 3) which characterize research and the computer programs found in the area of

[31]These choices may be interdependent. Say, a constructive approach might favor a representation by belief bases, and hence define belief revision operations on some finite base, rather than the whole background theory. Moreover, the epistemological stance determines what constitutes *rational epistemic change*. The foundationalist accepts only those beliefs which are justified, thus having an additional challenge of computing the reasons for an incoming belief. On the other hand, the coherentist must maintain coherence, and hence make only those minimal changes which do not endanger (at least) consistency.

computational philosophy of science, namely:

1. Scientific discovery is problem solving
2. The study of discovery is part of the methodological agenda of philosophy of science.
3. The computer programs are to be historically, psychologically and philosophically adequate.

The first principle is in accordance with the paradigm already identified as emerging in the 70's, that of science as problem solving, but in this case applied to discovery alone. Thus, the second principle states that problem solving is a notion to be handled within the methodology of science. In turn, this conceptual move suggests that existing computational and logical tools devised for other disciplines, like those existing in cognitive science and in artificial intelligence, may be imported and thus help bring some order to represent and model the aspects and machinery of scientific knowledge, its birth and development as well. Heuristic strategies are immersed in BACON's computer discovery simulation, machine learning is performing Popper's neglected induction, and abductive inference is modelling the epistemics of explanation generation.

The three claims given in (3) (identified in [Kuipers, 2001]), point to adequacy conditions for discovery computer programs. Ideally, a computer program that simulates discoveries should capture some aspects of its history, at least as far as a credible description of its development is told. Moreover, the design of this kind of computer programs should not overlook that it is, to some extent, a simulation of the real way a human would proceed. This implies some kind of cognitive commitment with its machinery, which gives sense to the psychological adequacy requirement. Finally, the computer program must be philosophically adequate in that there must exist some philosophical theory as a base for its epistemics. However, as noted by Kuipers, these claims "*may come into conflict with one another*" [Kuipers, 2001, 290].

The resulting enterprise is impressive in the unification of the disciplines involved, all for a common goal, that of giving an account of discovery and development of a privileged type of human knowledge, scientific knowledge. The methodological point is that the methods and heuristic strategies existing in computer science, have proved useful in artificial intelligence and cognitive simulation, and are used by several computer programs. All these tools have therefore been imported to philosophy of science to give a computer modeling account of processes such as explanation, confirmation, falsification, evaluation and discovery, and in general, of the modeling of the dynamics of scientific theories.

A major criticism to this whole enterprise however, is reflected in the debate of whether these computer programs really make new discoveries, for they seem to produce theories new to the program but not new to the world, and its discoveries seem spoon-fed rather than created. However, more recent research reveals that indeed the computer has been able to help produce important new research. One prominent example is reported in [Gillies, 1996, 50–55], and concerns the discovery of new laws about protein secondary structure. Further cases concern taxonomic discoveries in astrophysics, as well as qualitative laws in biochemical cancer research [Langley, 2000].

The Case of Abduction

Here we will deal with a particular case of scientific inference, that of inference to explanatory hypotheses, namely abduction. Research on abduction in artificial intelligence dates back to the seventies [Pople, 1973], but it is only fairly recently that it has attracted great interest, in areas like logic programming [Kakas *et al.*, 1995], knowledge assimilation [Kakas and Mancarella, 1990], and diagnosis [Poole *et al.*, 1987], to name just a few. It has been a topic of several workshops in artificial intelligence conferences (ECAI96, IJCAI97, ECAI98, ECAIOO) and model-based reasoning ones (MBR'98, MBR'01). It has also been at the center of some computer applied publications [Josephson and Josephson, 1994], and also present when compared with induction [Flach and Kakas, 2000]. Moreover, explanation based systems for computer applications were at the core of research in non-monotonic logics (cf. section 5.1). In all these places, the discussion about the different aspects of abduction has been conceptually challenging but also shows a (terminological) confusion with its close neighbour, induction.

We will now present the standard logical format for this inference, followed by the implementation given by the logic programming community, to continue by a proposal for a general taxonomy of this kind of reasoning. We then propose a particular interpretation, which conceives abduction as a process of epistemic change, a conception which goes beyond the interpretation of abduction as logical inference. We will finish by a brief coverage of the place of abduction in cognitive science, broadly conceived.

Abduction as Inference

The general trend in logic based approaches to abduction in AI interprets abduction as *backwards deduction plus additional conditions.* This brings it very close to deductive-nomological explanation in the Hempel style, witness the following format. What follows is the standard version of abduction as deduction via some consistent additional assumption, satisfying certain extra conditions. It combines some common requirements from the literature (cf. [Konolige, 1990; Kakas *et al.*, 1995; Mayer and Pirri, 1993; Aliseda, 1997] for further motivation):

Given a theory θ (a set of formulae) and a formula φ (an atomic formula), α is an explanation if

$\theta, \alpha \models \varphi$
α is consistent with θ

α is minimal[32]
α has some restricted syntactical form (usually an atomic formula or a conjunction of them).

An additional condition not always made explicit is that φ is not a logical consequence of θ. This says that the fact to be explained should not already follow from the background theory alone. Sometimes, the latter condition figures as a precondition for an *abductive problem* (cf. [Kakas *et al.*, 1995]).

Abduction as Computation in Logic Programming

Abduction emerges naturally in logic programming (cf. section 4.5) as a 'repair mechanism', completing a program with the facts needed for a query to succeed. This may be illustrated by the famous abductive rain example in Prolog:

Program P:
lawn-wet $\leftarrow$ rain.
lawn-wet $\leftarrow$ sprinklers-on.

Query q: lawn-wet.

Given program P, query q does not succeed because it is not derivable from the program. For q to succeed, either one (or all) of the facts 'rain', 'sprinklers-on', 'lawn-wet' would have to be added to the program. Abduction is the process by which these additional facts are produced. This is done via an extension of the resolution mechanism that comes into play when the backtracking mechanism fails. In our example above, instead of declaring failure when either of the above facts is not found in the program, they are marked as 'hypothesis', and proposed as those formulas which, if added to the program, would make the query succeed.

In actual Prolog abduction, for these facts to be counted as abductions, they have to belong to a pre-defined set of 'abducibles', and to be verified by additional conditions (so-called 'integrity constraints'), in order to prevent a combinatorial explosion of possible explanations[33].

[32]There are several ways to characterize minimality, cf. [Aliseda, 2006].

[33]In logic programming, the procedure for constructing explanations is left entirely to the resolution mechanism, which affects not only the order in which the possible explanations are produced, but also restricts the form of explanations, for rules cannot occur as abducibles, since explanations are produced out of sub-goal literals that fail during the backtracking mechanism. The additional restrictions select the best hypothesis. Thus, processes of both construction and selection of explanations are clearly marked in logic programming. (Another relevant connection here is to research in 'inductive logic programming' [Michalski, 1994], which integrates abduction and induction.). Logic programming does not use blind deduction. Different control mechanisms for proof search determine how queries are processed. This additional degree of freedom is crucial to the efficiency of the enterprise. Hence, different control policies will vary in the abductions produced, their form and the order in which they appear. To us, this variety suggests that the procedural notion of abduction is intensional, and must be identified with different practices, rather than with one deterministic fixed procedure.

A Taxonomy for Abduction

What we have seen so far may be summarized as follows. Abduction is a general process for producing explanations, with a certain inferential structure. We consider these two aspects to be of equal importance. Moreover, on the process side, we may distinguish between constructing possible explanations and selecting the best one amongst these. As for the logical form of abduction, we have found that it may be viewed as a threefold relation:

$$T, C \Rightarrow E$$

between an observation E, an abduced item C, and a background theory T. (Other parameters are possible here, such as a preference ranking — but these would rather concern the further selection process.) Against this background, we propose three main parameters that determine types of abduction. (i) An 'inferential parameter' ($\Rightarrow$) sets some suitable logical relationship among explananda, background theory, and explanandum. (ii) Next, 'triggers' determine what kind of abduction is to be performed: E may be a novel phenomenon, or it may be in conflict with the theory T. (iii) Finally, 'outcomes' (C) are the various products of an abductive process: facts, rules, or even new theories.

In the above schema, the notion of explanatory inference $\Rightarrow$ is not fixed. It can be classical derivability $\vdash$ or semantic entailment $\models$, but it does not have to be. Instead, we regard it as a parameter which can be set independently. It ranges over such diverse values as probable inference ($T, C \Rightarrow_{probable} E$), in which the explanans renders the explanandum only highly probable, or as the inferential mechanism of logic programming ($T, C \Rightarrow_{prolog} E$). Further interpretations include dynamic inference ($T, C \Rightarrow_{dynamic} E$, cf. [van Benthem, 1996]), replacing truth by information change potential along the lines of belief update or revision. Our point here is that abduction is not one specific non-standard logical inference mechanism, but rather a way of using any one of these.

As previously stated, for Peirce, abductive reasoning is triggered by a *surprising phenomenon*. The notion of surprise, however, is a relative one, for a fact E is surprising only with respect to some background theory T providing 'expectations'. What is surprising to me (e.g. that the canal bridge floor goes up from time to time) might not be surprising to a Dutch person. We interpret a surprising fact as one which needs an explanation. From a logical point of view, this assumes that the fact is not already explained by the background theory $T : T \not\Rightarrow E$.

Moreover, our claim is that one also needs to consider the status of the negation of E. Does the theory explain the negation of observation instead ($T \Rightarrow \neg E$)? Thus, we identify at least two triggers for abduction: *novelty* and *anomaly*:

Abductive Novelty: $T \not\Rightarrow E, T \not\Rightarrow \neg E$

E is novel. It cannot be explained ($T \not\Rightarrow E$), but it is consistent with the theory ($T \not\Rightarrow \neg E$).

Abductive Anomaly: $T \not\Rightarrow E, T \Rightarrow \neg E$

E is anomalous. The theory explains rather its negation ($T \Rightarrow \neg E$).

As already stated, novelty is the condition for an abductive problem. The suggestion in [Aliseda, 1997; 2006] is to incorporate anomaly as a second basic type[34]. Abducibles themselves come in various forms: facts, rules, or even theories. Sometimes one simple fact suffices to explain a surprising phenomenon, such as rain explaining why the lawn is wet. In other cases, a rule establishing a causal connection might serve as an explanation, as in our case connecting cloud types with rainfall. And many cases of abduction in science provide new theories to explain surprising facts. These different options may sometimes exist for the same observation, depending on how seriously we want to take it[35].

Once the above parameters get set, several kinds of abductive processes arise. For example, abduction triggered by novelty with an underlying deductive inference, calls for a process by which the theory is expanded with an explanation. The fact to be explained is consistent with the theory, so an explanation added to the theory accounts deductively for the fact. However, when the underlying inference is statistical, in a case of novelty, theory expansion might not be enough. The added statement might lead to a 'marginally consistent' theory with low probability, which would not yield a strong explanation for the observed fact. In such a case, theory revision is needed (i.e. removing some data from the theory) to account for the observed fact with high probability.

Our aim is to point out that several kinds of abductive processes are used for different combinations of the above parameters. (In Aliseda [1997] some procedures for computing different types of outcomes in a deductive format are explored in detail).

This taxonomy gives us the big picture of abductive reasoning. We can now see the patterns in a clearer focus. Varying the inferential parameter, we cover not only cases of deduction (plus additional conditions) but also statistical inferences. Different forms of outcomes will play a role in different types of procedures for producing explanations. In computer science jargon, triggers and outcomes are, respectively, preconditions and outputs of abductive devices, whether these be computational procedures or inferential ones.

Abduction as Belief Revision in Theory Change

Abductive reasoning may be seen as an epistemic process for belief revision. In this context an incoming sentence φ is not necessarily an observation, but rather a belief for which an explanation is sought. Existing approaches to abduction

[34]Of course, non-surprising facts (where $T \Rightarrow E$) should not be candidates for explanation. Even so, one might speculate if facts which are merely probable on the basis of T might still need explanation of some sort to further cement their status.

[35]Moreover, we are aware of the fact that genuine explanations sometimes introduce new concepts, over and above the given vocabulary. (For instance, the eventual explanation of planetary motion was not Kepler's, but Newton's, who introduced a new notion of 'force' — and then derived elliptic motion via the Law of Gravity.) Abduction via new concepts is outside the scope of our analysis. (Cf. [Thagard, 1992]) for an account of new concepts via conceptual combination).

usually do not deal with the issue of incorporating φ into the set of beliefs. Their concern is just how to give an account for φ. If the underlying theory is closed under logical consequence, however, then φ should be automatically added once we have added its explanation (which a foundationalist would then keep tagged as such).

Practical connections of abduction to theories of belief revision have often been noted. Of many references in the literature, we mention Aravindan and Dung [1994] (which uses abductive procedures to realize contractions over theories with 'immutability conditions'), and Williams [1994] (which studies the relationship between explanations based on abduction and 'Spohnian reasons').

Our claim however, is stronger. Abduction can function in a model of theory revision as a means of producing explanations for incoming beliefs. But also more generally, as defined above, it provides a model for epistemic change. Let us discuss some reasons for this. First, what we call the two 'triggers' for abductive reasoning correspond to two of the three epistemic attitudes of a formula introduced by Gärdenfors's (cf. section 5.1), viz., being undetermined or rejected. We did not consider accepted beliefs, since these do not call for explanation.

> φ is a novelty iff neither φ nor $\neg\varphi$ is a logical consequence of θ
>
> (φ is undetermined)
>
> φ is an anomaly iff φ is not a logical consequence of θ and $\neg\varphi$ is indeed a logical consequence of θ
>
> (φ is rejected)
>
> φ is an accepted belief
>
> (φ is a logical consequence of θ)[36].

In our account of abduction, both a novel phenomenon and an anomalous one involve a change in the original theory. The former calls for expansion and the latter for revision, which in turn involves contraction and then expansion. So, the basic operations for abduction are expansion and contraction. Therefore, both epistemic attitudes and changes in them are reflected in the presented abductive model.

However, our main concern is not the incoming belief φ itself. We rather want to compute and add its explanation α. But since φ is a logical consequence of the revised theory, it could easily be added. Here, then, are our abductive operations for epistemic change:

> Abductive Expansion
>
> Given a novel formula φ for θ, a consistent explanation α for φ is computed and then added to θ.

[36]The epistemic attitudes are presented in Gärdenfors [1988] in terms of membership (e.g., a formula φ is accepted if $\varphi \in \theta$). We defined them in terms of entailment, since theories are not necessarily closed under logical consequence.

> Abductive Revision
>
> Given a novel or an anomalous formula φ for θ, a consistent explanation α for φ is computed, which will involve modification of the background theory θ into some suitably new θ'. Again, intuitively, this involves both 'contraction' and 'expansion' (cf. section 5.1).

In its emphasis on explanations, our abductive model for belief revision is richer than many theories of belief revision[37]. Admittedly, though, not all cases of belief revision involve explanation, so our greater richness also reflects our restriction to a special setting.

Abduction in Cognitive Science

In cognitive science, abduction is a crucial ingredient in processes like inference, learning, and discovery, performed by people and computers to build theories of the world surrounding them. There is a growing literature on computer programs modelling these processes, and on abduction in particular.

A noteworthy reference is found in the earlier mentioned field of computational philosophy of science, and in broader computational cognitive studies of inductive reasoning [Thagard, 1988; 1992]. These studies distinguish several relevant mechanisms for hypotheses generation, indeed four kinds of abduction are implemented in the program PI (Processes of Induction): "*simple, existential, rule-forming, and analogical. Simple abduction produces hypotheses about individual objects ... Existential abduction postulates the existence of previously unknown objects... Rule-forming abduction produces rules that explain other rules, and hence is important for generating theories that explain laws. Finally, analogical abduction uses past cases of hypothesis formation to generate hypotheses similar to existing ones.*" [Thagard, 1992, 54]

This particular approach shows, on the one hand, the multiplicity of contexts in which abduction may appear, something which explains the need of a further distinction into abductive kinds, and on the other hand, it shows that it is closely related to other inferential processes, such as induction. In fact, simple abduction[38] seems to be a case of enumerative induction, perhaps one in which the conclusion is not a general or universal one.

5.3 *Philosophy of Science: Logics of Discovery*

The renewed enterprise of logics of discovery is based on two fundamental assumptions. The first of them is that the context of discovery allows, to some extent,

[37] Cf. [Aliseda, 2000] for a discussion of the type of belief revision theory this model of abduction corresponds to.

[38] According to Thagard: "the simplest case of abduction is one in which you want to explain why an object has some characteristic and you know that all objects with a particular property have that characteristic. Hence you conjecture that the object has the property in order to explain why it has the characteristic." [Thagard, 1992, 9]

a precise formal treatment. The second one claims that there is no single logical method in scientific practice in general, and with respect to abduction in particular. By this assumption, however, it is not claimed that it is possible to provide a logical analysis for all and every part of scientific inquiry. In this respect, the enterprise is as modest as the one proposed by Simon (cf. 4.3) and has no pretensions that it can offer either a logical analysis of great scientific discoveries, or put forward a set of logical systems that would provide general norms to make new discoveries. The aim is rather to lay down logical foundations in order to explore some of the formal properties under which new ideas may be generated and evaluated. The compensation we gain from this very modest approach is that we can gain some insight into the logical features of some parts of the scientific discovery and explanation processes. This is in line with a well-known view in the philosophy of science, namely that phenomena take place within traditions, something which echoes Kuhn's distinction between normal and revolutionary science. Hence, a general assumption is that a logical analysis of scientific discovery is for normal science, not denying there may be a place for some other kind of logical analysis of revolutionary science, but clearly leaving it out of the scope of this enterprise, at least for the moment.

The potential for providing a logic for scientific discovery is found in a normative account in the methodology of science, such as the one proposed by Simon (cf. 4.3)[39]. However, in this kind of computational approach, logic is identified with *pattern seeking methods*, a notion which fits very well their algorithmic and empirical approach to the question of a logic of discovery, but has little to do with providing logical foundations for their programs, either as conceived in the logico-mathematical tradition or as in artificial intelligence logical research.

Our claim is that logic, as understood in modern non-standard formal systems, has a place in the study of logics of discovery. By putting forward a logic for scientific discovery we claim no lack of rigour. What is clear is that standard deductive logic cannot account for abductive or inductive types of reasoning, but the present situation in logical research has gone far beyond the formal developments that deductive logic reached last century, and new research includes the formalization of several other types of reasoning, like induction and abduction. And a general goal could be to study the wider field of human reasoning while hanging on to these standards of rigour and clarity.

[39] From antiquity to the mid 19^{th} century, researchers had in mind a logic of discovery of a descriptive nature, one that would capture and describe the way humans reason in science. These 'logics', however had little success, for they failed to provide such an account of discovery. Thereafter, when the search for a logic of discovery was abandoned, a normative account prevailed in favour of proposals of logics of justification. Regarding the approaches of Popper and Simon in this respect, it is clear that while Popper overlooked the possibility of a normative account of logics of discovery, Simon centered his efforts in the development of heuristic procedures, to be implemented computationally, but not on logics — per se — of discovery. In fact, there are proposals strictly normative and formal in nature, such as that found in Kelly [1997], in which one argues for a computational theory as the foundation of a logic of discovery, one which studies algorithmic procedures for the advancement of science. (Cf. [Aliseda, 2006] for an in-depth analysis of the development of logics of discovery).

In connection to philosophy of science, there are already some logical proposals that lay bridges between non-monotonic logics and Hempel's models for explanation. In [Tan, 1992] an inductive statistical model is constructed based on Reiter's default logic and in [Aliseda, 1997; 2006] two models of scientific explanation (deductive and statistical) are presented as cases of abductive logic, thus claiming that these models do not follow the canons of classical logic. These proposals naturally bring up the question of the properties of such enriched notions of consequence, which are in turn studied within the axiomatic theory of consequence relations (Cf. [Aliseda, 1997; 2006] for a structural characterization of abduction).

There are still many challenges ahead for the formal study of reasoning in scientific discovery, such as giving an integrated account of deductive, inductive, abductive and analogical styles of inference, the use of diagrams by logical means, and in general the device of logical operations for theory building and change. Already new logical research is moving into these directions. In Burger and Heidema [2002] degrees of 'abductive boldness' are proposed as spectrum for inferential strength, ranging from cases with poor background information to those with (almost) complete information. Systems dealing with several notions of derivability all at once have also been proposed. A formula may be 'unconditionally derived' or 'conditionally derived', the latter case occurring when a line in a proof asserts a formula which depends on hypotheses which may be later falsified, thus pointing to a notion of proof which allows for addition of lines which are non-deductively derived as well as for deletion of them when falsifying instances occur. This account is found in the framework of (ampliative) adaptive logics, a natural home for abductive inference [Meheus *et al.*, 2002] and for (enumerative) induction inference [Batens, 2005] alike, in which it is possible to combine deduction with ampliative steps [Meheus, 1999].

The formalization of analogical reasoning is still a growing area of research, without a precise idea of what exactly an analogy amounts to. Perhaps research on mathematical analogy [Polya, 1954], work on analogy in cognitive science [Helman, 1988] or investigations into analogical argumentation theory recently proposed for abduction [Gabbay and Woods, 2005], may serve to guide research in this direction. Finally, the study of 'diagrammatic reasoning' is a research field on its own right [Barwise and Etchemendy, 1995], showing that the logical language is not restricted to the two-dimensional left to right syntactic representation, but its agenda still needs to be expanded on research for non-deductive logics.

As for formal approaches for theory building and change applied to philosophy of science, one line of contemporary research in this direction, adopted by Aliseda [2005a], concerns the extension of a classical logical method to model empirical progress in science, as conceived by Theo Kuipers [1999; 2001]. In particular, the goal is to operationalize the task of *instrumentalist abduction*, that is, theory revision aiming at empirical progress. This particular account shows that evaluation and improvement of a theory can be modelled by (an extension of) the framework of Semantic Tableaux [Aliseda, 2005a].

All the above suggests that the use of non-standard logics to model processes in scientific practice, such as confirmation, falsification, explanation building, theory improvement and discovery is, after all, a feasible project. All of these logics are either characterized via structural rules, by an axiomatic or semantic approach as well as within dynamic theory revision systems. Nevertheless, this claim requires a broad conception of what logic is about as well as a modest account of what is it to be found.

6 CONCLUSIONS

Philosophy of science in the twentieth century has been remarkably rich and varied in character, and has passed through a series of quite radical changes. This chapter began by analysing the Vienna Circle approach which dominated the discipline from the 1920s to the end of the 1950s. This approach was based on the idea that philosophy of science should consist of the logical analysis of scientific theories. The logic to be used in this task might include inductive logic with the associated concepts of probability and confirmation, but, as far as deductive logic was concerned, it was limited to classical logic with the occasional reference to intuitionistic logic. The distinction between the context of discovery and the context of justification was generally accepted, and this, together with the narrow notion of what constituted logic, led to the conclusion that philosophy of science should concentrate on the justification of scientific theories, leaving the question of discovery to other disciplines.

The dominance of the logical approach was challenged in the 1960s by the emergence (or perhaps better re-emergence) of the historical approach to the philosophy of science. This allowed questions concerned with scientific discovery to be tackled once again, though they were approached from a historical rather than logical point of view. Some of the proponents of the historical approach (notably Kuhn and Feyerabend) went as far as to reject the logical approach (particularly confirmation theory) even in the context of justification. Others, however, (particularly Popper and Lakatos) were more sympathetic to logic, and sought to combine logic with the new historical approach.

From the mid-1970s we drew attention to a new force at work on the development of philosophy of science, namely the increasing importance of the computer. Computer science led to cognitive psychology and both together gave rise to the concept of science as problem solving — a conception to be found in both Simon and in Popper's works of this period. Simon, however, had more involvement with the computer, and in particular raised the question of whether computers could be programmed to make scientific discoveries. This helped to bring scientific discovery to the fore as a central problem in the philosophy of science, but, in contrast to the earlier historical period, the analysis of discovery could now be carried out with logical and computational techniques. Further developments in computer science, cognitive science and logic itself, provided a new set of tools of a logical and computational nature. The rather limited logic used by the Vienna Circle was now augmented by the discovery of quite new systems of logic such as non-monotonic logics. These strengthened the trend towards including issues of discovery as part of the philosophy of science agenda in the eighties and nineties. This period has been characterised by results and concrete proposals, for the representation and modelling of common sense reasoning, scientific inference, knowledge growth and discovery. In part, the emergence of new research in logic in the field of artificial intelligence, has helped to analyse, in a fresh and more powerful manner, issues of scientific inference in a rigorous and a systematic way.

As for the connection between logic and history, with the emergence of computer science a new contact with history is presented, which gave rise to computational philosophy of science, a place in which history and computing meet on a par. This creates the possibility of a partial synthesis between the logical and the historical approaches with a new computational element added to the mix. This has been part of the research agenda of a number of contemporary philosophers of science since about the mid 80's, but is by no means a privileged topic.

To conclude, the present setting, in which we have all logical, historical and computational approaches to philosophy of science, fosters the view that what we need is a balanced philosophy of science, one in which we take advantage of a variety of methodologies, such as logical, computational and also historical, all together giving a broad view of science. This view was already presented by Suppes [1951–69] in the late sixties. We know at present that logical models (classical or otherwise) are insufficient to completely characterize notions like explanation, confirmation or falsification in philosophy of science, but this fact does not rule out that some problems in the history of science may be tackled from a formal point of view. For instance, "*some claims about scientific revolutions, seem to require statistical and quantitative data analysis, if there is some serious pretension to regard them with the same status as other claims about social or natural phenomena*" [Suppes, 1951–69, 97]. In fact, *Computational Philosophy of Science* may be regarded as a successful marriage between historical and formal approaches. It is argued that although several heuristic rules have been derived from historical reconstructions in Science, they are proposed to be used for future research" [Meheus and Nickles, 1999].

This is not to say however, that historical analysis of scientific practice could be done in a formal fashion, or that logical treatment should contain some kind of "historic parameter" in its methodology, but we claim instead that these two views should share their insights and findings in order to complement each other.

ACKNOWLEDGMENTS

We are grateful to Ladislav Kvasz who read an earlier draft of sections 1, 2 and 3 carefully and made many useful suggestions which have been incorporated into the final version.

BIBLIOGRAPHY

[Alchourrón *et al.*, 1985] C. Alchourrón, P. Gärdenfors, and D. Makinson. 'On the logic of theory change: Partial meet contraction and revision functions'. *Journal of Symbolic Logic*, 50: 510–530, 1985.

[Aliseda, 1997] A. Aliseda. *Seeking Explanations: Abduction in Logic, Philosophy of Science and Artificial Intelligence*. PhD Dissertation, Philosophy Department, Stanford University. Published by the Institute for Logic, Language and Computation (ILLC), University of Amsterdam (ILLC Dissertation Series 1997–4). 1997.

[Aliseda, 2000] A. Aliseda. Abduction as epistemic change: A Piercean model in Artificial Intelligence. In P. Flach and A. Kakas (eds.), *Abductive and Inductive Reasoning: Essays on their Relation and Integration*, pages 45–58. Kluwer Academic Press, 2000.

[Aliseda, 2004] A. Aliseda. Sobre la lógica del descubrimiento científico de Karl Popper. Suplemento 11 (Monográfico Karl Popper), *Signos Filosóficos*, pages 115–130. Universidad Autónoma Metropolitana. México, 2004. Translated as "On Karl Popper's Logic of Scientific Discovery", in L. Magnani (ed.), *Model Based Reasoning in Science and Engineering*, King's College Publications, 2006.

[Aliseda, 2005a] A. Aliseda. Lacunae, empirical progress and semantic tableaux. In R. Festa, A. Aliseda, and J. Peijnenburg (eds.), *Confirmation, Empirical Progress, and Truth Approximation: Essays in Debate with Theo Kuipers* (Volume 1), pages 169–189. Poznan Studies in the Philosophy of the Sciences and the Humanities, Vol. 83, 2005.

[Aliseda, 2005b] A. Aliseda. What is a logical system? A commentary. In Artemov *et al.* (eds.), *We Will Show Them! Essays in Honour of Dov Gabbay on his 60th Birthday*, Volume 2. College Publications. King's College, 2005.

[Aliseda, 2006] A. Aliseda. *Abductive Reasoning: Logical Investigations into Discovery and Explanation*. Synthese Library, Vol. 330. Springer-Kluwer Academic Publishers, 2006.

[Aravindan and Dung, 1994] C. Aravindan and P. M. Dung. Belief dynamics, abduction and databases. In C. MacNish, D. Pearce, and L. M. Pereira (eds.), *Logics in Artificial Intelligence. European Workshop JELIA'94*, pages 66–85. Lecture Notes in Artificial Intelligence 838. Springer-Verlag, 1994.

[Aristotle, 1941] Aristotle. *Nicomachean Ethics*. English Translation by W. D. Ross in R. McKeon (ed.), *The Basic Works of Aristotle*, pages 927–1112, Random House, 1941.

[Barwise and Etchemendy, 1995] J. Barwise and J. Etchemendy. *Hyperproof*. Center for the Study of Language and Information (CSLI). Lecture Notes Series 42. Stanford, CA, 1995.

[Batens, 2005] D. Batens. A logic of induction. In R. Festa, A. Aliseda, and J. Peijnenburg (eds.), *Cognitive Structures in Scientific Inquiry: Essays in Debate with Theo Kuipers*, Volume 1, pages 221–247. Poznan Studies in the Philosophy of the Sciences and the Humanities, vol. 83, 2005.

[van Benthem, 1984] J. van Benthem. Lessons from Bolzano. Center for the Study of Language and Information. Technical Report CSLI-84-6. Stanford University. 1984. Later published as 'The variety of consequence, according to Bolzano'. *Studia Logica*, 44: 389–403, 1985.

[van Benthem, 1994] J. van Benthem. General dynamic logic. In D. M. Gabbay (ed.), *What is a Logical System?*, pages 107–140, Clarendon Press. Oxford, 1994.

[van Benthem, 1996] J. van Benthem. *Exploring Logical Dynamics*. CSLI Publications, Stanford University, 1996.

[Bolzano, 1837] B. Bolzano. *Wissenschaftslehre*, Seidel Buchhandlung, Sulzbach, 1837. Translated as *Theory of Science* by B. Torrel, edited by J. Berg. D. Reidel Publishing Company. Dordrecht, The Netherlands. 1973.

[Brewka *et al.*, 1997] G. Brewka, J. Dix and K. Konolige. *Non-Monotonic Reasoning: An Overview*. Center for the study of Language and Information (CSLI), Lecture Notes 73, 1997.

[Burger and Heidema, 2002] I. C. Burger and J. Heidema. Degrees of abductive boldness. In L. Magnani, N. J. Nersessian, C. Pizzi (eds.), *Logical and Computational Aspects of Model-based Reasoning*, Kluwer Applied Logic Series, pages 181–198. Kluwer Academic Publishers, 2002.

[Carnap, 1950] R. Carnap. *Logical Foundations of Probability*. University of Chicago Press, 1950. 2nd Edition, 1963.

[Carnap, 1963] R. Carnap. Intellectual autobiography. In P. A.Schilpp (ed.), *The Philosophy of Rudolf Carnap*, Library of Living Philosophers, Open Court, pages 3–84, 1963.

[Chomsky, 1972] N. Chomsky. *Language and Mind* (Enlarged Edition). New York: Harcourt Brace & Jovanovich, 1972.

[Chomsky, 1976] N. Chomsky. *Reflections on Language*. Glasgow: Fontana/Collins, 1976.

[Chomsky, 1993] N. Chomsky. *Lectures on Government and binding: The Pisa Lectures.* Mouton de Gruyter, Berlin. 7^{nd} Edition, 1993.

[English, 1978] J. English. Partial interpretation and meaning change. *The Journal of Philosophy*, 75: 57–76, 1978.

[Feyerabend, 1970] P. Feyerabend. Consolations for the specialist. In Lakatos and Musgrave (eds.), pages 197–230, 1970.

[Feyerabend, 1975] P. Feyerabend. *Against Method. Outline of an Anarchist Theory of Knowledge*. 1975. Verso, 1984.

[Feyerabend, 1978] P. Feyerabend. *Science in a Free Society*. 1978. Verso, 1985.

[Feyerabend, 1987] P. Feyerabend. *Farewell to Reason*. Verso, 1987.

[Flach and Kakas, 2000] P. Flach and A. Kakas. *Abduction and Induction. Essays on their Relation and Integration*. Applied Logic Series. volume 18. Kluwer Academic Publishers. Dordrecht, The Netherlands, 2000.

[Fodor, 1975] J. Fodor. *The Language of Thought.* New York, Crowell, 1975.

[Frank, 1941] P. Frank. *Modern Science and its Philosophy.* 1941. Paperback edition. Collier Books, 1961.

[Fullbrook, 2004] E. Fullbrook (ed.) *A Guide to What's Wrong with Economics*. Anthem, 2004.

[Gabbay, 1985] D. M. Gabbay, Theoretical foundations for non-monotonic reasoning in expert systems. In K. Apt (ed.), *Logics and Models of Concurrent Systems*, pages 439–457. Springer-Verlag. Berlin, 1985.

[Gabbay, 1994] D. M. Gabbay (ed.). *What is a Logical System?* Clarendon Press. Oxford, 1994.

[Gabbay and Woods, 2005] D. M. Gabbay and J. Woods. *A Practical Logic of Cognitive Systems. The Reach of Abduction: Insight and Trial*. Amsterdam: North-Holland, volume 2, 2005.

[Gadol, 1982] E. Gadol (ed.). *Rationality and Science. A Memorial Volume for Moritz Schlick in Celebration of the Centennial of his Birth*. Springer Verlag, 1982.

[Galliers, 1992] J. L. Galliers. Autonomous belief revision and communication. In P. Gärdenfors (ed.), *Belief Revision*, pages 220–246. Cambridge Tracts in Theoretical Computer Science, Cambridge University Press, 1992.

[Gärdenfors, 1988] P. Gärdenfors. *Knowledge in Flux: Modeling the Dynamics of Epistemic States*. MIT Press, 1988.

[Gärdenfors and Rott , 1995] P. Gärdenfors and H. Rott. Belief revision. In D. M. Gabbay, C. J. Hogger and J. A. Robinson (eds.), *Handbook of Logic in Artificial Intelligence and Logic Programming*. Volume 4, Clarendon Press, Oxford Science Publications, 1995.

[Gentzen, 1934] G. Gentzen. *Recherches sur la deduction loguique*. Translaction of *Untersuchungen úber das logische schliessen 1934*, R. Feys and J. Ladriere P.U.F., Paris, 121955, 1934.

[Gillies, 1993] D. A. Gillies. *Philosophy of Science in the Twentieth Century. Four Central Themes*. Blackwell, 1993.

[Gillies, 1996] D. A. Gillies. *Artificial Intelligence and Scientific Method*. Oxford University Press, 1996.

[Gillies and Zheng, 2001] D. A. Gillies and Y. Zheng. Dynamic interactions with the philosophy of mathematics. *Theoria*, 16: 437–459, 2001.

[Ginsberg, 1988] A. Ginsberg. Theory revision via prior operationalization. In *Proceedings of the Seventh Conference of the AAAI*, 1988.

[Haack, 1993] S. Haack. *Evidence and Inquiry. Towards Reconstruction in Epistemology*. Blackwell, Oxford UK and Cambridge, Mass., 1993.

[Hanson, 1961] N. R. Hanson. *Patterns of Discovery*. Cambridge at The University Press, 1961.

[Harman, 1986] G. Harman. *Change in View: Principles of Reasoning*. Cambridge, Mass. MIT Press, 1986.

[Helman, 1988] D. H. Helman (ed.). *Analogical Reasoning : Perspectives of Artificial Intelligence, Cognitive Science and Philosophy*. Dordrecht, Netherlands: Reidel, 1988.

[Hempel, 1965] C. G. Hempel. *Aspects of Scientific Explanation and other Essays in the Philosophy of Science*. The Free Press, 1965.

[Hintikka and Remes, 1974] J. Hintikka and U. Remes. *The Method of Analysis: Its Geometrical Origin and Its General Significance*. D. Reidel Publishing Company. Dordrecht, Holland, 1974.

[Hintikka and Remes, 1976] J. Hintikka and U. Remes. Ancient geometrical analysis and modern logic. In R. S. Cohen (ed.), *Essays in Memory of Imre Lakatos*, pages 253–276. D. Reidel Publishing Company. Dordrecht Holland, 1976.

[Hobbs *et al.*, 1990] J. R. S. Hobbs, M. Stickel, D. Appelt, and P. Martin. Interpretation as abduction. *SRI International, Technical Note* 499. Artificial Intelligence Center, Computing and Engineering Sciences Division, Menlo Park, CA, 1990.

[Josephson and Josephson, 1994] J. R. Josephson and S. G. Josephson. *Abductive Inference. Computation, Philosophy, Technology.* Cambridge University Press, 1994.

[Kakas and Mancarella, 1990] A. Kakas and P. Mancarella. Knowledge assimilation and abduction. In *Proceedings of the European Conference on Artificial Intelligence, ECAI'90* International Workshop on Truth Maintenance. Springer-Verlag Lecture Notes in Computer Science. Stockholm, 1990.

[Kakas *et al.*, 1995] A. C. Kakas, R. A. Kowalski, F. Toni. Abductive logic programming. *Journal of Logic and Computation*, 2(6): 719–770, 1995.

[Kelly, 1997] K. Kelly. *The Logic of Reliable Inquiry (Logic and Computation in Philosophy).* Oxford University Press, 1997.

[Klein, 2000] N. Klein. *No Logo.* Flamingo, 2000.

[Konolige, 1990] K. Konolige. A general theory of abduction. In: *Automated Abduction, Working Notes*, pages 62–66. Spring Symposium Series of the AAA. Stanford University, 1990.

[Kowalski, 1979] R. Kowalski. *Logic for Problem Solving.* North-Holland, 1979.

[Kowalski, 1994] R. Kowalski. Logic without model theory. In D. M. Gabbay (ed.), *What is a Logical System?*, pages 35–72. Clarendon Press. Oxford, 1994.

[Kraus *et al.*, 1990] S. Kraus, D. Lehmann, M. Magidor. Non-monotonic reasoning, preferential models and cumulative logics. *Artificial Intelligence*, 44: 167–207, 1990.

[Kuhn, 1957] T. S. Kuhn. *The Copernican Revolution. Planetary Astronomy in the Development of Western Thought.* 1957. Vintage Books. 1959.

[Kuhn, 1959] T. S. Kuhn. The essential tension: Tradition and innovation in scientific research. 1959. In [Kuhn, 1977, 225–39].

[Kuhn, 1962] T. S. Kuhn. *The Structure of Scientific Revolutions.* The University of Chicago Press, 1962 , 7^{th} Impression, 1969.

[Kuhn, 1970] T. S. Kuhn. Reflections on my critics. In Lakatos and Musgrave (eds.), pages 231–78, 1970.

[Kuhn, 1974] T. S. Kuhn. Second thoughts on paradigms. 1974. In [Kuhn, 1977, 293–319].

[Kuhn, 1976] T. S. Kuhn. Theory change as structure change: Comments on the sneed formalism. 1976. In [Kuhn, 2000, 176–95].

[Kuhn, 1977] T. S. Kuhn. *The Essential Tension.* The University of Chicago Press, 1977.

[Kuhn, 1979] T. S. Kuhn. Metaphor in science. 1979. In [Kuhn, 2000, 196–207].

[Kuhn, 1983] T. S. Kuhn. Commensurability, comparability, communicability. 1983. In [Kuhn, 2000, 33–57].

[Kuhn, 1990] T. S. Kuhn. The road since structure. 1990. In [Kuhn, 2000, 90-104].

[Kuhn, 2000] T. S. Kuhn. *The Road since Structure.* The University of Chicago Press, 2000.

[Kuipers, 1999] T. Kuipers. Abduction aiming at empirical progress or even truth approximation leading to a challenge for computational modelling. *Foundations of Science*, 4: 307–323. 1999.

[Kuipers, 2001] T. Kuipers. *Structures in Science. Heuristic Patterns Based on Cognitive Structures.* Synthese Library 301, Kluwer AP, Dordrecht. The Netherlands, 2001.

[Kvasz, 1998] L. Kvasz. History of geometry and the development of the form of its language. *Synthese*, 116: 141–68, 1998.

[Kvasz, 1999] L. Kvasz. On classification of scientific revolutions. *Journal for the General Philosophy of Science*, 30: 201–32, 1999.

[Kvasz, 2000] L. Kvasz. Changes of language in the development of mathematics. *Philosophia Mathematica*, 8: 47–83, 2000.

[Kvasz, 2002] L. Kvasz. Lakatos' methodology between logic and dialectic. In G. Kampis, L. Kvasz and M. Stöltzner (eds.), *Appraising Lakatos. Mathematics, Methodology and the Man*, Vienna Circle Institute Library: Kluwer, pages 211–41, 2002.

[Lakatos, 1963-4] I. Lakatos. *Proofs and Refutations. The Logic of Mathematical Discovery.* 1963-4. Cambridge University Press, 1984.

[Lakatos, 1968] I. Lakatos. Changes in the problem of inductive logic. 1968. In [Lakatos, 1978b, 128-200].

[Lakatos, 1970] I. Lakatos. Falsification and the methodology of scientific research programmes. 1970. In [Lakatos, 1978a, 8–101].

[Lakatos, 1974] I. Lakatos. Popper on Demarcation and Induction. 1974. In [Lakatos, 1978a, 139–67].

[Lakatos, 1978] I. Lakatos. Newton's Effect on Scientific Standards. 1978. In [Lakatos, 1978a, 193–222].

[Lakatos, 1978a] I. Lakatos. *Philosophical Papers Volume I The Methodology of Scientific Research Programmes*. 1978. Edited by J. Worrall and G. Currie, Cambridge University Press, 1984.

[Lakatos, 1978b] I. Lakatos. *Philosophical Papers Volume II Mathematics, Science and Epistemology*. 1978. Edited by J. Worrall and G. Currie, Cambridge University Press, 1984.

[Lakatos and Musgrave, 1970] I. Lakatos and A. Musgrave (eds.), *Criticism and the Growth of Knowledge*. Cambridge University Press 1970.

[Langley, 2000] P. Langley. The computational support of scientific discovery. *International Journal of Human-Computer Studies*, 53: 393–410, 2000. Web: www.isle.org/∼langley/pubs.html.

[Langley *et al.*, 1987] P. Langley, H. Simon, G. Bradshaw, and J. Zytkow. *Scientific Discovery*. Cambridge, MA:MIT Press/Bradford Books, 1987.

[Laudan, 1977] L. Laudan. *Progress and its Problems*. Berkeley, University of California Press, 1977.

[Laudan, 1980] L. Laudan. Why was the logic of discovery abandoned?. In T. Nickles (ed.), *Scientific Discovery, Logic and Rationality*, pages 173–183. D. Reidel Publishing Company, 1980.

[Lipton, 2003] P. Lipton. Kant on wheels. *Social Epistemology*, 17: 215–19, 2003.

[McCarthy, 1980] J. McCarthy. Circumscription: A form of non-monotonic reasoning. *Artificial Intelligence*, 13: 27–39, 1980.

[McKie, 1935] D. McKie. *Antoine Lavoisier. The Father of Modern Chemistry*. Victor Gollancz, 1935.

[Massimi, 2005] M. Massimi. *Pauli's Exclusion Principle: The Origin and Validation of a Scientific Principle*. Cambridge University Press, 2005.

[Masterman, 1970] M. Masterman. The nature of a paradigm. In [Lakatos and Musgrave, 1970, 59–89].

[Mayer and Pirri, 1993] M. C. Mayer and F. Pirri. First order abduction via tableau and sequent calculi. In *Bulletin of the IGPL*, vol. 1: 99–117, 1993.

[Meheus, 1999] J. Meheus. Model-based reasoning in creative processes. In L. Magnani, N. J. Nersessian, and J. Thagrad (eds.), *Model-Based reasoning in Scientific Discovery*. Kluwer Academic/Plenum Publishers. 1999.

[Meheus, 2002] J. Meheus. Ampliative adaptative logics and the foundation of logic-based approaches to abduction. In L. Magnani, N. Nersessian, C. Pizzi (eds.), *Logical and Computational Aspects of Model-based Reasoning*, Kluwer Applied Logic Series. Kluwer Academic Publishers, 2002.

[Meheus and Nickles, 1999] J. Meheus and T. Nickles. The methodological study of discovery and creativity — Some background. *Foundations of Science*, 4(3): 231–235, 1999.

[Meheus *et al.*, 2002] J. Meheus, L. Verhoeven, M. van Dyck, and D. Provijn. Ampliative adaptative logics and the foundation of logic-based approaches to abduction. In L. Magnani, N. Nersessian, and C. Pizzi (eds.), *Logical and Computational Aspects of Model-based Reasoning*, Kluwer Applied Logic Series, pages 39-72. Kluwer Academic Publishers, 2002.

[Menger, 1980] K. Menger. Introduction to Hans Hahn. *Empiricism, Logic, and Mathematics: Philosophical Papers*, ed. Brian McGuinness, Reidel, *ix-xviii*, 1980.

[Michalski, 1994] R. Michalski. Inferential theory of learning: Developing foundations for multistrategy learning. In *Machine Learning: A Multistrategy Approach*, Morgan Kaufman Publishers, 1994.

[Miller, 1986] A. I. Miller. *Imagery in Scientific thought. Creating 20th-Century Physics*. MIT Press, 1986.

[Motterlini, 1999] M. Motterlini. *For and Against Method. Imre Lakatos and Paul Feyerabend*. The University of Chicago Press, 1999.

[Neurath *et al.*, 1929] O. Neurath *et al. The Scientific Conception of the World.The Vienna Circle*. 1929. English translation, Reidel, 1973.

[Ohlbach and Reyle, 1999] H. J. Ohlbach and U. Reyle (eds.). Research themes of Dov Gabbay. In *Logic, Language and Reasoning*, pages 13–30. Kluwer Academic Publishers. 1999.

[Peirce, 1931-35] C. S. Peirce. *Collected Papers of Charles Sanders Peirce*. Volumes 1–6 edited by C. Hartshorne, P. Weiss. Cambridge, Harvard University Press. 1931–1935; and volumes 7–8 edited by A. W. Burks. Cambridge, Harvard University Press. 1958.

[Polya, 1954] G. Polya. *Induction and Analogy in Mathematics*. Vol I. Princeton University Press, 1954.

[Polya, 1962] G. Polya. *Mathematical Discovery. On Understanding, learning, and teaching problem solving*. Vol I. John Wiley & Sons, Inc. New York and London, 1962.

[Poole *et al.*, 1987] D. Poole, R. G. Goebel, and Aleliunas. Theorist: a logical reasoning system for default and diagnosis. In Cercone and McCalla (eds.), *The Knowledge Fronteer: Essays in the Representation of Knowledge*. Springer Verlag Lecture Notes in Computer Science, pages 331–352, 1987.

[Pople, 1973] H. E. Pople. On the mechanization of abductive logic. In: *Proceedings of the Third International Joint Conference on Artificial Intelligence* (IJCAI-73). San Mateo: Morgan Kauffmann, Stanford, CA, pages 147–152, 1973.

[Popper, 1934] K. R. Popper. *The Logic of Scientific Discovery*. 1934, 6^{th} (revised) impression of the 1959 English translation, Hutchinson, 1972.

[Popper, 1960] K. R. Popper. The growth of scientific knowledge. 1960. In D. Miller (ed.), *A Pocket Popper*. Fontana Paperbacks. University Press, Oxford, 1983. This consists of Chapter 10 of K. Popper, (1963) *Conjectures and Refutations. The Growth of Scientific Knowledge*. 5th ed. London and New York, Routledge, 1963.

[Popper, 1963] K. R. Popper. *Conjectures and Refutations. The Growth of Scientific Knowledge*. Routledge & Kegan Paul, 1963.

[Popper, 1972] K. R. Popper. *Objective Knowledge. An Evolutionary Approach*. Oxford University Press, 1972.

[Popper, 1983] K. R. Popper. *Realism and the Aim of Science*. Hutchinson, 1983.

[Quine, 1953] W. V. O. Quine. *From a Logical Point of View*. 1953. 2^{nd} Revised Edition, Harper Torchbooks, 1963.

[Ramsey, 1929] F. P. Ramsey. Last papers F. philosophy. 1929. In R. B.Braithwaite (ed.), *The Foundations of Mathematics and other Logical Essays by F. P. Ramsey*, Routledge & Kegan Paul, 4^{th} Impression, 1965.

[Reggia *et al.*, 1985] J. A. Reggia, D. S. Nau and Y. Wang. A formal model of diagnostic inference I. Problem formulation and descomposition. *Inf. Sci.*, 37, 1985.

[Reiter, 1980] R. Reiter. A logic for default reasoning. *Artificial Intelligence*, 13, 1980.

[Reiter, 1987] R. Reiter. A theory of diagnosis from first principles. *Artificial Intelligence*, 32, 1987.

[Rescher, 1978] N. Rescher. *Peirce's Philosophy of Science. Critical Studies in His Theory of Induction and Scientific Method*. University of Notre Dame, 1978.

[Russell, 1914] B. Russell. *Our Knowledge of the External World as a Field for Scientific Method in Philosophy*. 1914. George Allen & Unwin, 1961.

[Sankey, 1994] H. Sankey. *The Incommensurability Thesis*. Avebury, 1994.

[Scott, 1971] D. Scott. On engendering an illusion of understanding, *Journal of Philosophy*, 68: 787–808, 1971.

[Shapere, 1964] D. Shapere. The structure of scientific revolutions. *Philosophical Review*, 73: 383–94, 1964.

[Sheehan, 1985] H. Sheehan. *Marxism and the Philosophy of Science: A Critical History*. Humanities Press, 1985.

[Shoham, 1988] Y. Shoham. *Reasoning about Change. Time and Causation from the standpoint of Artificial Intelligence*. The MIT Press, Cambridge, Mass, 1988.

[Simon, 1973] H. Simon. Does scientific discovery have a logic? In H. Simon, *Models of Discovery*, pages 326–337, 1973. A Pallas Paperback. Reidel, Holland. 1977. (Originally published in *Philosophy of Science*, 40: 471–480).

[Stadler, 2001] F. Stadler. *The Vienna Circle. Studies in the Origins, Development, and Influence of Logical Empiricism*. Springer, 2001.

[Suppes, 1951-69] P. C. Suppes. *Studies in the methodology and foundations of science. selected papers from 1951 to 1969*. D. Reidel. Dordrecht. The Netherlands, 1951-69.

[Tan, 1992] Y. H. Tan. *Non-Monotonic Reasoning: Logical Architecture and Philosophical Applications*. PhD Thesis, University of Amsterdam, 1992.

[Thagard, 1988] P. R. Thagard. *Computational Philosophy of Science*. Cambridge, MIT Press. Bradford Books, 1988.

[Thagard, 1992] P. R. Thagard. *Conceptual Revolutions*. Princeton University Press. 1992.

[Turing, 1950] A. M. Turing. computing machinery and intelligence. *Mind*, 59: 433–60, 1950.

[Williams, 1994] M. A. Williams. Explanation and theory base transmutations. In *Proceedings of the European Conference on Artificial Intelligence*, ECAI'94, pages 341–246, 1994.

[Wittgenstein, 1953] L. Wittgenstein. *Philosophical Investigations*. 1953. Basil Blackwell. 2^{nd} Edition. 1958.

[Wolters, 2003] G. Wolters. Carl Gustav Hempel. Pragmatic Empiricist. In P. Parrini, W. C. Salmon, M. H. Salmon (eds.), *Logical Empiricism: Historical and Contemporary Perspectives*, University of Pittsburgh Press, pages 109–22, 2003.

[Zahar, 1973] E. Zahar. Why did Einstein's programme supersede Lorentz's? *British Journal for the Philosophy of Science*, 24: 95–123 and 223–262, 1973.

DEMARCATING SCIENCE FROM NON-SCIENCE

Martin Mahner

1 INTRODUCTION

Every field of inquiry deals with some subject matter: it studies something rather than nothing or everything. Thus it should be able to tell, at least roughly, what sort of objects it is concerned with and how its objects of study differ from those studied by other disciplines. If a discipline were unable to offer a characterization of its subject matter, we would be entitled to suspect that its representatives do not really know what they are talking about. Evidently, what holds for all fields of inquiry also holds for a particular discipline such as the philosophy of science. Therefore, it belongs to the job description, so to speak, of the philosopher of science to tell us what that "thing" called science is.

Yet whereas everyone seems to know intuitively which fields of knowledge are scientific (such as physics and biology) and which are not (such as astrology and palmistry), it has proved difficult to come up with a satisfactory demarcation criterion. Indeed, many of the demarcation criteria proposed by philosophers of science have proved to be unsatisfactory, for being either too narrow or too wide. In addition, due to the historical and sociological studies of science, many contemporary authors believe that there simply is, or even can be, no single criterion or set of criteria allowing for a clear-cut characterization of scientific vis-à-vis non-scientific areas of human inquiry. In particular, most contemporary philosophers doubt that there is a set of necessary and sufficient conditions demarcating science from non-science. It comes as no surprise therefore that, in a survey conducted with 176 members of the Philosophy of Science Association in the US, about 89% of the respondents denied that any universal demarcation criteria have been found [Alters, 1997].

Does this vindicate relativist views like Feyerabend's [1975] well-known anything-goes epistemology? Must we give up the attempt to descriptively partition the landscape of human cognition into scientific and nonscientific areas, as well as to tell genuine science from bogus science (pseudo-science)? Is, then, the philosophy of science unable to address the normative problem of why some form of human inquiry arrives at (approximately) true knowledge, whereas some other, purporting to be equally scientific, must be judged to produce only illusory knowledge?

Handbook of the Philosophy of Science: General Philosophy of Science - Focal Issues, pp. 515–575.
Volume editor: Theo Kuipers. General editors: Dov M. Gabbay, Paul Thagard and John Woods.

This situation illustrates the problems that a reasonably comprehensive analysis of the demarcation problem should address. To this end, let us restate three questions formulated by Thagard [1988, p. 157], which will guide the following analysis:

1. Why is it important to demarcate science and from what should it be distinguished?

2. What is the logical form of a demarcation criterion?

3. What are the units that are marked as scientific or nonscientific, in particular as pseudoscientific?

2 WHY DEMARCATION?

Let us begin with the second conjunct of the first question. Evidently, demarcating science means to demarcate it from nonscience. Yet, in so doing, how widely do we have to conceive of nonscience? In the broad sense, simply anything that is not science is nonscience: driving, swimming, cooking, dancing, or having sex are nonscientific activities. Now, the philosopher of science is not particularly interested in demarcating science from nonscientific activities such as these, although they involve learning and hence some cognition, leading in particular to procedural knowledge. Naturally, he will be interested first of all in cognitive activities and practices leading to propositional knowledge, i.e., explicit and clear knowledge that can be either true or false to some extent. Thus, we are primarily interested in nonscience in the sense of *nonscientific cognitive fields* involving hypotheses and systems of such (i.e., theories) as well as the procedures by means of which these are proposed, tested, and evaluated. Consequently, distinguishing science from nonscience in this narrower sense is not restricted to the classical science/metaphysics demarcation attempted by the neopositivists and Popper, but extends to all nonscientific epistemic fields.

The first reason why we should strive for a demarcation of science is theoretical: it is the simple fact stated in the beginning that every field of knowledge should be able to tell roughly what it is about, what its objects of study are. Unless the philosophy of science simply is nothing but epistemology in general, it should be able to distinguish scientific from nonscientific forms of cognition. Note that such a basic distinction between science and nonscience is not pejorative: it does not imply that nonscientific forms of cognition and knowledge are necessarily bad or inferior. Nobody doubts the legitimacy and value of the arts and humanities, for example.

The second reason, or rather set of reasons, why we ought to demarcate science from nonscience concerns in particular the normative aspect of distinguishing science from pseudoscience. Moreover, it is practical rather than theoretical, comprising aspects of mental and physical health, as well as culture and politics. Should we entrust our own as well as other peoples' health and even lives to diagnostic or therapeutic methods which have no proven effect? Should public health insurance cover magical cures? Should we even consider the possibility that clairvoyants search for missing children? Should we have dowsers search for people buried by an avalanche? Should we make sure that tax payers' money be spent only on funding scientific rather than pseudoscientific research? Should we demand that people living in a modern democratic society base their political decisions on scientific knowledge rather than superstition? Examples such as these show that the distinction of science and pseudoscience is vital not just to our physical, but also to our cultural and political life.

This aspect leads us to a third reason: the need of science education to teach what science is and how it works. To this end, the science educator needs input from the philosopher of science as to the nature of science [Alters 1997; Eflin *et al.*, 1999]. The science educator cannot just tell her students that nobody knows what science is, but that they are nonetheless supposed to learn science rather than pseudoscience. For all these reasons, we ought not to give up too readily when facing difficulties with the demarcation of science from nonscience, and in particular from bogus science.

3 HOW DEMARCATION?

In the history of the philosophy of science various demarcation criteria have been proposed (see Laudan [1983]). Let us briefly recapitulate some of the classical attempts at demarcation, starting with logical positivism. In tune with their linguistic focus the neopositivists' foremost goal was to distinguish sense from nonsense. A sentence was deemed to be (semantically) meaningful if, and only if, it was verifiable; otherwise, it was nonsense. Whereas, according to neopositivism, the statements of science are verifiable and thus meaningful, those of metaphysics and all other kinds of bad philosophy were not; they were just nonsense (see, e.g., [Wittgenstein, 1921; Carnap, 1936/37; Ayer, 1946]). To verify a sentence means to find out whether it is true, which requires that it be tested empirically. The central tenet, then, was that testability is a necessary condition of (semantic) meaning: meaning $\rightarrow$ testability.

One problem with this view is that it has things the wrong way. Indeed, to test a statement empirically, must we not know what it means, i.e., what it says, in the first place? To devise a test for a statement such as "unemployment increases crime", we must already know what that sentence means. Only thus can we handle the variables involved. Hence, meaning is in fact a necessary condition of testability rather than the other way around: testability $\rightarrow$ meaning [Mahner and Bunge, 1997]. As a consequence, nonscientific discourse can be semantically meaningful, although it may not be testable empirically. If a Christian tells us "Jesus walked on water", we know quite well what that means, although we cannot test this statement, which we may moreover regard as purely mythical. We may reject it for many reasons, but not for being nonsensical.

A logical and methodological objection against the verifiability thesis is the fact that it is rarely possible to verify a statement in the strict sense, i.e., to show that it is true. For example, we may easily verify or falsify a spatiotemporally restricted existential statement, such as “There is a pink elephant in my office”. But if we are faced with general statements such as “For all X: if A then B”, observing B does confirm A, but only inductively, never conclusively. The most cherished scientific statements, then, namely law statements, are not strictly verifiable. Hence the strong concept of conclusive verification was soon replaced by the weaker notion of confirmation [Carnap, 1936/37]. Nonetheless, having always been a critic of induction, Popper [1934/1959] suggested giving up the verifiability condition in favor of a falsifiability principle. Indeed, according to the modus tollens rule, observing not-B entails not-A. Thus, logically, falsification is conclusive, whereas verification is not. This logical asymmetry is the basis for Popper’s famous demarcation criterion of falsifiability [Popper, 1963].

Critics have soon pointed out that not all scientific statements are universal: there are also unrestricted existential statements, such as “There are positrons” [Kneale, 1974; Bunge, 1983b]. These can be verified, e.g., in this case by coming up with at least one specimen of a positron; but they cannot be falsified, because we cannot search the entire universe to conclusively show that it does not contain even a single positron. Other critics have shown that scientists do not give up a theory as being unscientific just because there are some falsifying data, unless there is a better theory at hand, concluding that Popper’s criterion does not match scientific practice [Lakatos, 1973; 1974].

In the light of such critique, Popper has later clarified his position, emphasizing that it is not practical falsifiability which is his concern, but instead logical falsifiability [Popper, 1994]. That is, a statement is *logically falsifiable* if there is at least one *conceivable* observation statement contradicting it. In other words, a statement is scientific only if it is not consistent with every possible state of affairs. In proposing falsifiability as a demarcation criterion, Popper had in mind examples such as Freudian psychoanalysis. According to psychoanalysis, the Oedipus complex is either manifest or repressed, so no possible observable state of affairs can count against it: it is unfalsifiable as a matter of principle. Again, critics were quick to point out that this does not hold for all of psychoanalysis (e.g., [Grünbaum, 1984]): while some claims are indeed unfalsifiable, many others are falsifiable and others have actually been falsified. The same holds for many other pseudosciences, such as astrology and creationism. For example, the central tenet of creationism, that a supernatural being created the world, is indeed unfalsifiable: it is compatible with every possible observation statement, for any state of affairs can be seen as exactly what the creator chose to do. Other and more specific creationist claims, however, such as that the earth is only 6000 years old, are falsifiable and falsified. Thus, the falsifiability criterion may be useful to weed out some claims as pseudoscientific, but it accepts too many falsifiable and falsified statements as scientific, although there are good reasons to regard them as pseudoscientific. For all these reasons, falsifiability has been almost unanimously rejected as *the* demarcation criterion (e.g., [Kuhn, 1970; Kitcher, 1982; Bunge, 1983b; Laudan, 1983; Siitonen, 1984; Lugg, 1987; Thagard, 1978; 1988; Rothbart, 1990; Derksen, 1993; Resnik, 2000]).

Being first of all a logical condition, falsifiability is an ahistorical criterion. The historical turn in the philosophy of science has suggested taking into account both the development of theories and their relation to rival theories. In so doing, it shifted the focus of demarcation from individual statements or hypotheses to entire theories. The classic approach is certainly Lakatos's [1970] notion of a research program. A research program is a historical sequence of theories, where each subsequent theory results form a semantical reinterpretation of its predecessor, or from adding auxiliary assumptions or other modifications. A research program is called *theoretically progressive* if each new theory has a larger content, e.g., by having greater explanatory or predictive power, than its predecessor. It is also *empirically progressive* if it is confirmed, i.e., if it actually leads to the discovery of some new fact. A research program is then called *progressive* if it is both theoretically and empirically progressive; otherwise it is called *degenerative*. Finally, Lakatos [1970] takes a research program to be *scientific* if it is at least theoretically progressive; otherwise it is *pseudoscientific*.

Critics like Laudan [1983] have objected that progress might as well occur in nonscientific fields, such as philosophy, and that some branches of science did not progress much during some periods in their history. And what if some science actually had discovered and explained everything in its domain that is to be discovered and that needs explanation; in other words, what if there were such a thing as "the end of science" as envisioned as the ultimate goal of science by Einstein [Holton, 1993; Haack, 2003], and as was more recently speculated on by Horgan [1995]? Would such a theory or discipline be no longer scientific just because it does not or rather cannot progress any more? Similarly, there is the opposite problem of radically new theories: can they be scientific without being part of an existing research program? Consequently, however useful the criterion of growth and progressiveness will be in many cases, it too cannot provide *the* decisive demarcation criterion.

Kuhn [1970] has suggested that we focus not so much on the testability of theories as on their problem-solving capacity. He illustrates his point in the case of astrology. Many predictions of astrology are testable and have failed, but astrology is not therefore a science as Popper's falsifiability criterion would allow for. According to Kuhn, this is because astrology has no puzzles to solve: even its failed predictions did not entice the astrological community to engage in problem-solving activity. At most, astrology has rules to apply, for it is essentially a craft — or rather a pseudotechnology. But if applying rules is simply a characteristic of technology rather than science, what distinguishes a scientific technique from a nonscientific one? Finally, although earlier authors like David Hilbert have already dealt with problems and pointed out that a wealth of problems is an indicator of a good science, what about the end-of-science scenario mentioned above? Would even a true theory become nonscientific if all problems surrounding it were solved?

Scientists would hardly think so.

Nevertheless, other authors too suggested focusing on problems, in particular on the permissible rules of asking questions and stating problems [Siitonen, 1984]. Obviously, the solution of a problem should enrich our knowledge and contribute to stating and solving further problems. Moreover, we may learn something about the current state of a field of inquiry by asking questions such as "What are the problems?", "How are these problems formulated?", and "Which efforts have been and can be used to solve these problems?" [Siitonen l.c., p. 347]. But again, all these considerations are far from providing a new demarcation criterion.

In a book on creationism, Kitcher [1982] focuses on three characteristics of science. First, the auxiliary hypotheses involved in the testing of any scientific theory are *independently testable* themselves, i.e., independently of the theory it is supposed to protect or of the particular case for which they were introduced. Second, scientific practices are *unified* wholes, not patchworks of isolated and opportunistic methods: they apply a small number of problem-solving strategies (if preferred, exemplars) to a wide range of cases and problems. Third, good scientific theories are *fertile* in the sense that they open up new areas of research. Thereby, one of the sources of fecundity is the incompleteness of scientific theories, so that some problems remain unresolved. Incompleteness and some unresolved problems are therefore not shortcomings of scientific theories but instead sources of progress.

Thagard [1988] lists five features characterizing science. As a method of inference, scientists use "correlation thinking"; that is, by means of various statistical procedures they infer causation, if any, from correlation (rather than from mere resemblance). They seek empirical confirmation and disconfirmation, and evaluate theories in relation to alternative theories, whereby these theories are consilient and simple. Finally, science progresses over time, i.e., it develops new theories explaining new facts. Thagard does not regard these features as both necessary and sufficient, but only suggests that they belong to the conceptual profile of science.

Rothbart [1990] attempts to formulate a metacriterion (or adequacy condition) for any demarcation criterion. This condition is the testworthiness of a hypothesis or theory, i.e., its plausibility to be selected for experimentation in the first place. To this end a hypothesis must fulfill certain eligibility requirements prior to testing. If it does not fulfill even one of these requirements, a hypothesis is untestworthy and hence unscientific. Actual demarcation is then obtained by specifying such eligibility requirements. One such requirement is that the proposed theory must account for all the facts that its rival background theory explains; another is that it must yield test implications that are inconsistent with those of its rival theory.

Vollmer [1993] distinguishes necessary and merely desirable features of a good scientific theory. The necessary conditions are noncircularity, internal consistency (noncontradiction), external consistency (compatibility with the bulk of well-confirmed knowledge), explanatory power, testability, and test success (confirmation). Among the desirable features are predictability and reproducibility, as well as fecundity and simplicity (parsimony). Predictability and reproducibility are not among the necessary conditions, for otherwise historical sciences, such as evolutionary biology, geology, cosmology, and of course human history, would not count as scientific because both their predictability and reproducibility are limited. However, even in such overall historical fields, not all events are unique, but repeatable at least in the sense that events *of the same kind* may reoccur on a more or less regular basis. Consequently, if the very nature of some event is of the repeatable kind, irreproducibility may still indicate that something is wrong with the given field's claim of being a science.

Reisch [1998] attempts to resuscitate the unity of science ideal of logical positivism, though not in its reductionist form. He suggests identifying the various theoretical and methodological interconnections of the sciences, which should result in what he calls a *network unification* of science and hence a *network demarcation*. An epistemic field that cannot be incorporated into the existing network of the established sciences without destroying it should be rejected as pseudoscientific. Again, such network demarcation does not draw a fixed boundary around the sciences, but allows for changes in what belongs to that network and what not. Finally, the neopositivist aspect of Reisch's approach consists in the claim that the specification of the interconnections among scientific fields is essentially a *scientific* form of demarcation rather than a philosophical one.

The result of the preceding overview is clear: neither is there a single criterion such as falsifiability to demarcate science from nonscience, nor is there a generally accepted set of necessary and sufficient criteria to do this job. However, *pace* Laudan [1983], this does not imply that no demarcation is possible. To see why, it will be useful to make a brief foray into the philosophy of biology, which faces a similar problem.

In the philosophy of biological systematics there has been a long debate concerning the ontological status and definition of biological species (see, e.g., [Mahner and Bunge, 1997]). The classical, essentialist view regards species as natural kinds defined by a set of necessary and sufficient properties. Against this view the anti-essentialists have argued that, due to the high genetic and morphological variety of organisms, there simply is no set of necessary and sufficient characters possessed by all and only the organisms of a given species, let alone higher taxonomic units (see, e.g., [Dupré, 1993]). Nevertheless, the organisms of a given species usually are both similar among each other and distinct from organisms belonging to different species.

The radical answer to this problem says that species should therefore not be conceived of as kinds at all, but rather as concrete supraorganismic individuals. Now, science too can be viewed as a concrete system, namely as a research community. In this case it is relatively easy to determine who is part of this community and who is not. But of course, science is more than that: in contrast to the sociologist of science, the philosopher of science is more interested in science as a collection of reliable knowledge items produced by following certain methodological standards. To this end, science is better regarded as a special *kind* of knowledge production, which can be demarcated from other kinds of knowledge acquisition. Now, if traditional essentialism with respect to kinds cannot be upheld at least in biology, we might try what in the philosophy of systematics could be called *moderate* species essentialism. This is the idea that biological species *can* be viewed as natural kinds, if only in a weaker sense defined by a variable *cluster* of features instead of a strict set of necessary and sufficient properties (see, e.g., [Boyd, 1999; Wilson, 1999]). Thus, whereas no single property need be present in all the members of the given species, there are always "enough" properties making these organisms belong to the given kind. (Forerunners of moderate essentialism are Wittgenstein's family resemblance concept, which was suggested for demarcation purposes by Dupré [1993], and Beckner's [1959] polythetic species definitions.)

Despite the unsolved problems concerning the formalization of such disjunctive characterizations [Mahner and Bunge, 1997], applying this approach to the demarcation of science might allow us to define science through a variable cluster of properties too, rather than through a set of necessary and sufficient conditions. For example, if we came up with ten conditions of scientificity (all of equal weight), we might require that an epistemic field fulfill at a minimum seven out of these ten conditions in order to be regarded as scientific, but it would not matter which of these ten conditions are actually met. According to the formula $N!/n!(N-n)!$, where $N = 10$ and $n = 7$, and adding the permutations for $n = 8, n = 9$, and $n = 10$, there would in this case be a total of 176 possible ways of fulfilling the conditions of scientificity.

In a similar vein, many authors have argued that, for demarcation purposes, we must do with a reasonable *profile* of any given field rather than with a clear-cut distinction (e.g., [Thagard, 1988; Derksen, 1993; Eflin *et al.*, 1999]). In other words, it will be worthwhile to attempt to come up with a whole battery of science indicators. Such a cluster of criteria should be as comprehensible as possible, and enable us to examine every possible field of knowledge by a list of marks noting the presence or absence of the relevant features, or the compliance or noncompliance with some, e.g. methodological, rule. On this basis we should be able to come to a well-reasoned (and hence rational) conclusion concerning the scientific or nonscientific status of a cognitive field.

4 CHARACTERIZING FIELDS OF KNOWLEDGE

As is obvious from the preceding section, scientificity has been ascribed to many items: individual statements, problems, methods, systems of statements (theories in the strict sense), entire practices (theories in the broad sense), historical sequences of theories and/or practices (research programs), and fields of knowledge. Given the notorious problems with the traditional demarcation criteria, it seems promising to try the most comprehensive approach, for it allows us to consider the many facets of the scientific enterprise, namely the fact that science is at the same time a body of knowledge and a system of people including their activities or practices, and hence something that did not come into existence ex nihilo, but has developed over several centuries from a mixed bag of ordinary knowledge, metaphysics and non- or at most pre-scientific inquiry. This most comprehensive approach is the one focusing on fields of knowledge (see, e.g., [Thagard, 1988]). As we shall see at the end, this approach has the advantage that, by demarcating entire fields of knowledge as scientific or nonscientific, it allows us to also evaluate individual components of such a field, like characteristic principles and methods, as being scientific or not.

Before we begin to determine whether or not a field of knowledge is scientific, we must first define what a field of knowledge is. In a chapter on pseudoscience, Thagard [1988] just refers to fields of knowledge, without, however, offering much of a characterization. In their work on "interfield theories", Darden and Maull [1977] point out that fields are characterized, for example, by a certain domain of facts as well as a number of problems, methods and theories concerning that domain. However, they do not use their characterization to demarcate between scientific and nonscientific, let alone pseudoscientific, fields. The most comprehensive characterization of epistemic fields has been proposed by Bunge [1983a; b], who has moreover explicitly used it for demarcation purposes [Bunge, 1982; 1983b; 1984]. For this reason, I shall rely heavily on his analysis, but will readily modify it whenever necessary to make it better suited to the task at hand.

4.1 Epistemic Fields

Roughly speaking, an epistemic field is a group of people and their practices, aiming at gaining knowledge of some sort. Thus, physics and theology, astronomy and astrology, psychology and parapsychology, evolutionary biology and creationism, art history and mathematics, medicine and economics, philosophy in general and epistemology in particular, as well as biology in general and genetics in particular are examples of epistemic fields. These examples show that epistemic fields, or, if preferred, cognitive disciplines can be more or less inclusive; in other words, they may be structured hierarchically. (Note that in the following we shall not distinguish between "field" and "discipline", although one might argue that the term "discipline" be reserved for denoting generally acknowledged or institutionalized fields.) They also indicate that the knowledge acquired in an epistemic field need neither be factual nor true: we may acquire knowledge about purely fictional rather than factual entities, and our knowledge may be false or illusory. (Thus, we do not adopt the classic definition of "knowledge" as "justified true belief", but rather the Popperian view that all knowledge is hypothetical, so that it can turn out to be either true or false.) Finally, it is immaterial whether the aim of our cognitive activities is either epistemic or practical, or both.

These examples of fields of knowledge just serve as a starting point for a more detailed characterization. In his characterization of epistemic fields, Bunge [1983a] considers ten aspects:

1. the group or *community* C of knowers or knowledge seekers;

2. the *society* S hosting the activities of C;

3. the *domain* or universe of discourse D of the members of C, i.e., the collection of factual or fictional objects the members of C refer to in their discourse;
4. the *philosophical background* or general outlook G, which consists of
   (a) an *ontology* or general view on the nature of things,
   (b) an *epistemology* or general view on the nature of knowledge, and
   (c) a *methodology*, *axiology* and *morality* concerning the proper ways of acquiring and handling knowledge;
5. the *formal background* F, which is a collection of logical or mathematical assumptions or theories taken for granted in the process of inquiry;
6. the *specific background* B, which is a collection of knowledge items (statements, procedures, methods, etc.) borrowed from other epistemic fields;
7. the *problematics* P, which is the collection of problems concerning the nature, value or use of the members of D, as well as problems concerning other components listed here, such as G or F;
8. the *fund of knowledge* K, which is the collection of knowledge items (propositions, theories, procedures, etc.) obtained by the previous and current members of C in the course of their cognitive activities;
9. the *aims* A, which are of course the cognitive, practical or moral goals of the members of C in the pursuit of their specific activities;
10. the *methodics* M, which is the collection of general and specific methods (or techniques) used by the members of C in their inquiry of the members of D.

Note that these aspects come in a certain logical order. For example, the method used to find out something in a given field depends on the problem to be solved, on what we already know and on our aims. Thus, Bunge analyzes an epistemic field E, for any given time, as an ordered set or, more precisely, a ten-tuple

$$\mathcal{E} = \langle C, S, D, G, F, B, P, K, A, M \rangle.$$

Since our emphasis here is on the usefulness of these coordinates for demarcation purposes, we can disregard the question of whether their order is optimal or whether an alternative order would be more adequate (e.g., exchanging P and K). Bunge calls the first three components of this ten-tuple the *material framework* of the given epistemic field, although he admits that this is a misnomer in the case of fields like mathematics and the humanities whose domains consists mostly or even exclusively of nonmaterial objects. In any case, C and S do consist of concrete objects, namely persons and systems of persons. Consisting mostly of abstract objects, the last seven components make up the *conceptual framework* of the field, which may as well be equated with Kuhn's notion of a paradigm or disciplinary matrix. This name too is a misnomer in some cases, because the methodics M need not only consist of rules and procedures as conceptual entities, but may also comprise material objects (artifacts) such as measuring instruments.

Most of the members of E will be obvious, such as D, G and M, but some remarks may nonetheless be helpful. For example, the two coordinates C and S indicate that cognition and knowledge are not self-existing, but activities of real people in a particular social environment. Only by taking these aspects into account can we do justice to the history, psychology, and sociology of knowledge. But why distinguish C from S, since C is actually a subsystem of S? Because the community C may have interesting sociological features worth examining and because it may emerge or go extinct, without necessarily having a serious effect on the entire society in which it exists or had existed. Think of L.R. Hubbard's scientology movement.

The problematics P and the aims A of an epistemic field are important characteristics, because the same domain may be studied by asking different questions, and with different aims. For example, biochemistry and molecular biology study virtually the same objects, namely certain classes of molecules, but they concern different problems: whereas biochemistry studies these molecules under purely chemical auspices, molecular biology is interested in the biological function of these molecules in living organisms. Similarly, the same object may be studied to simply learn more about it, or to control it by technical means. For example, the phylogeneticist may just be interested in the evolution of mosquitoes, whereas the applied entomologist and especially the ecotechnologist may be interested in how to control their population and restrict their geographical distribution.

4.2 Scientific epistemic fields

When speaking of science we are first of all interested in the *factual* (often called empirical) sciences, such as physics and chemistry, biology and psychology, as well as the social sciences. (Note that we prefer the expression "factual science" over "empirical science" because the advanced sciences are not just empirical, but have well-developed theoretical branches.) An epistemic field $\mathcal{S}$ is a (factually) *scientific* field if the elements of any ten-tuple $\langle C, S, D, G, F, B, P, K, A, M\rangle$ approximately satisfy the following conditions [Bunge, 1983b].

1. The community C of the field is a *research* community: it is a system of persons who share a specialized training, hold strong information links amongst each other, and initiate or continue a certain tradition of inquiry. Thus, every researcher belongs to either a local, regional, national, or international community of colleagues.

2. The society S hosting C supports or at least tolerates the activities of the persons in C. In particular, it allows for research free from authority, in that it does not proscribe which of its results have to be accepted as true, or else be rejected as false.

3. The domain D of a factual science deals exclusively with *concrete* entities (past, present and future), their properties and changes. These entities may be elementary particles, living beings, human societies, or the universe as a whole. Some of the entities hypothesized in a factual discipline may turn out not to exist really, but if they were real, they would be concrete (as opposed to abstract) entities.

4. That science rests on certain philosophical assumptions is rather uncontroversial. There is less agreement, however, as to which particular assumptions are characteristic of science. Let us therefore discuss some of the philosophical principles that are good candidates for membership in the general philosophical outlook G of any scientific field. To this end, consider a simple physiological experiment, which can be done in biology class (Fig. 1).

   Where is the hidden philosophy in this experiment? Unlike the solipsist or the follower of George Berkeley, the normal scientist does not assume that, when she is actually carrying out this experiment, it is occurring only in her mind. Nor does she suppose that a supernatural entity is producing the entire situation in her mind. We cannot prove that this is not actually the case, but it simply does not belong to the scientists' presuppositions. By contrast, the scientist takes it for granted that this experiment is occurring in an outer world existing independently of her mind, but including her as a part.

   Imagine we repeat this experiment several times under the same conditions. The first time, the gas produced would be helium, the second time oxygen, the third time no gas at all would appear. The fourth time, the entire setup would explode before even adding hydrogen peroxide, and the fifth time four of the test tubes would turn into chewing gum, whereas the fifth would fly off to the ceiling. For some reason such weird things do not happen. Instead, things remain the same under the same conditions. Moreover, the outcome of the experiment is *ceteris paribus* the same: the gas always consists of oxygen. Furthermore, its amount depends on the pH in the test tube, whereby the highest amount is produced at a pH of 8. Obviously, the properties of the things involved are constantly (i.e., lawfully) related. Imagine further that for some reason we do not get any gas out of the test tubes at all. In this case the scientist would not believe that the gas has disappeared into nothingness, but that there must be something wrong with the setup.

   Excluding effects coming out of nothingness or from some supernatural realm, the scientist further assumes that it is her adding hydrogen peroxide which causes the production of oxygen. In other words, by manipulating some part of the setup a certain effect can be produced, whereby the steps in this process are ordered: the steps in the causal chain follow each other rather than occurring capriciously. Furthermore, the scientist takes it for granted not only that no supernatural entities, like friendly fairies or evil

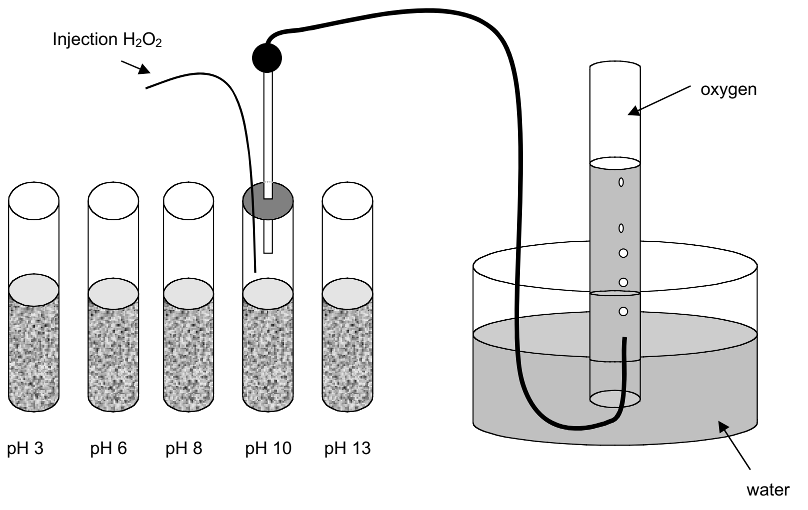


Figure 1. Take five test tubes filled with water and add a certain amount of yeast. Furthermore, by adding different amounts of hydrochloric acid (HCl) or caustic soda (NaOH) respectively, we arrange for a different acidity or alkalinity respectively in each tube, say, pH 3, pH 6, pH 8, pH 10, and pH 13. The yeast cells contain the enzyme catalase, which enables them to break down hydrogen peroxide into water and oxygen (i.e., $2H_2O_2 \rightarrow 2H_2O + O_2$). Upon adding a certain amount of hydrogen peroxide into one test tube after the other (by means of a syringe, for example), we each time close the tube and measure the amount of gas produced after 2 minutes by collecting it in a measuring tube, which is connected to the given test tube by a thin rubber hose. We do not need to specify the precise amounts and conditions here, because the basic setup of this experiment will be clear anyway (redrawn and modified from Knodel 1985, p. 39).

demons, meddle with the experiment either in the positive or in the negative, but also that they do not influence her own thinking, e.g., by making her hallucinate. And finally, she assumes that neither she herself nor anybody else can affect the setup by pure thinking or wishing alone, but only by acting; in other words, she takes it for granted that it is neither her own mind nor the mind of her colleague nor that of some little green alien on another planet which causes the outcome of the experiment.

In all this experimenting our scientist believes of course that she can get to know something about what is going on. Moreover, she also believes that the setup can be improved if necessary, and that thereby the precision of the measurement can be increased. Indeed, by varying and improving the experiment, she will find out that her earlier datum "The most oxygen is produced at a pH of 8.0" was not quite true, but that the maximum production occurs at a pH of 8.5. In other words, the initial finding was only an approximation to the real fact.

Let this sample of tacit assumptions suffice. It is time to extract some of the ontological, epistemological, and semantic *isms* or principles involved here.

a) Ontological assumptions

Despite the efforts of the positivists to denounce metaphysics as nonsense, it has long been acknowledged that science and metaphysics, though different, are related — and often even fruitfully so (see, e.g., [Agassi, 1964]). After all, one might argue that science is the emancipated daughter of metaphysics. As is indicated by the experiment described in the preceding, some minimal set of ontological tenets is presupposed even by modern science.

The first candidate is of course ontological realism, i.e., the thesis that there is a mind-independent world, whose inhabitants may become the subject matter of scientific investigation. Ontological realism is among the least controversial philosophical presuppositions of science, as is also indicated by Alters's [1997] survey mentioned earlier, in which about 90% of the interviewed philosophers agreed with the thesis that science presupposes realism. Note that ontological realism says nothing about whether this real world can be known and, if so, how and to which degree. This is a matter of epistemological realism. (It is, by the way, mostly the latter which is the target of antirealist criticism.)

The next assumption is ontological naturalism. This is the thesis that the inhabitants of the real world are exclusively natural as opposed to supernatural. Whether or not there is a transcendent world beyond our universe (if this very idea makes sense in the first place), our universe is causally closed, that is, there is no interaction with any possible other-worldly entities. Many philosophers of science would go even further and posit that there can be no interaction of concrete and spiritual as well as abstract entities either, even if the latter were natural ones — which reduces naturalism to materialism (e.g., [Armstrong, 1995; Mahner and Bunge, 1997]). Note that naturalism involves the parsimony principle (see Sect. 4.2, 4c). The least parsimonious view would be some sort of non-interventionism. This is the thesis that the universe is full of supernatural entities, but these have somehow agreed never to interfere with scientific measurements or experiments. Evidently, this view is quite arbitrary and nonparsimonious.

The third ontological ingredient of science is the principle of lawfulness. This is the hypothesis that the real world is not capricious, but behaves in a regular fashion. Indeed, if things behaved lawlessly, the world would resemble a cartoon movie in which everything can change into anything, forward and backward in time, in a completely arbitrary fashion. Presumably, there would be no living beings, no knowledge and no technology if the world were lawless. Note that the principle of lawfulness does not presuppose Laplacean determinism, because there are also stochastic processes — which follow probabilistic laws. Note further that, if laws as ontic regularities are distinguished from law statements purported to represent such laws, the various criticisms of the concept of natural law in science (e.g., by [Cartwright, 1983] and [Giere, 1999]) mostly concern the latter, i.e., the epistemological notion of a law. A too rigid traditional conception of natural law statements held by many philosophers of science, and our difficulties with idealization and approximation in representing real laws must not lead us to conclude that, as a consequence, there are no laws in the ontic sense, i.e., that the world behaves irregularly or even miraculously.

The fourth ontological presupposition is the principle of antecedence, which is often conflated with the causality principle. The antecedence principle maintains that causes precede their effects or, alternatively, that the presence is (causally or stochastically) determined by the past. By contrast, the principle of causality in the strict sense states that every event has an (external) cause producing the given event; more precisely, for every event e in some thing x, there is another event e' in some (other) thing $x' \neq x$, such that e' causes e. But since there are spontaneous (uncaused) events, such as exemplified by certain quantum events like radioactive decay, it is false as a universal principle. Nonetheless, in the case of our above experiment, we also need some version of the causality principle to account for the fact that our actions have some effect on the world.

The fifth ontological presupposition of science may be called the genetic or ex-nihilo-nihil-fit principle. Going back at least to Epicurus and Lucretius, this principle says that nothing comes out of nothing and nothing disappears into nothingness. Note that "nothing" here really means "nothing": even the curious vacuum field filling up empty space is *something* rather than nothing, for it can affect other things. (Note, incidentally, that this ontological assumption also affects physical cosmology: although one might be prepared to make exceptions for the universe as a whole, the genetic principle should encourage us to explore and prefer cosmological models, even big bang models, that do not assume a *creatio ex nihilo*, but presuppose some pre-existing state of the universe.)

Finally, there is the "no psi" principle [Broad, 1949; Bunge, 1983b], which is the postulate that minds or brain processes do not act directly on the things out there, but only through some motoric action of our body. Nobody could trust the readings of any measurement instrument or the results of any experiment if immediate mental forces and causes permeated the world.

These ontological principles must not be seen in isolation: they are a package deal. The idea that there are real and natural things, behaving lawfully and not popping out of, or into, nothing, is certainly the major metaphysical guide line of factual scientists. Note that these ontological and epistemological principles could all be false, which is why they are hypotheses or postulates, not ideological dogmas, as some critics of science tend to claim. However, both their eminent fertility and the extraordinary success of science justify that we accept them as true — for the time being. We might therefore call them the ontological default assumptions or, in some cases, metaphysical null hypotheses of factual science.

b) Epistemological assumptions

In order to do factual science, ontological realism must be combined with epistemological realism, i.e., the thesis that the real world can be known, if only approximately and imperfectly. Otherwise, scientists would just study the figments of their imagination, and technologists were unable to successfully alter real things, because this presupposes that at least some relevant properties of those things are known correctly.

Now, epistemological realism comes in different versions and strengths (see, e.g., the overview by Kuipers 2001, Ch. 2). We need not commit ourselves here to any position, although the most widely accepted version is likely to be what is often called scientific realism, which stipulates that we can know not just observables, but also unobservables. Elementary physics and evolutionary biology, for example, would make little sense without this assumption.

But what about instrumentalism, conventionalism, and other antirealist epistemological positions held on occasion by both scientists and philosophers of science? Are they not more parsimonious than realism? It is not just the claim that the majority of working scientists adopts realism in their daily work, but also the fact that, both in science and metascience, we should accept that position that has the greatest explanatory power and fecundity. In this regard realism beats instrumentalism, because the latter can explain neither the success nor, more importantly, the failure of scientific theories.

Moreover, whereas the instrumentalist cannot explain what the realist does and thinks, the realist is able to explain what the instrumentalist does. Thus realism subsumes instrumentalism [Vollmer, 1990; Kuipers, 2000]; see also [Kitcher, 1993] for an analysis of various antirealist arguments).

c) Methodological principles

A very general methodological maxim of any scientific approach is the principle of parsimony, also known as Ockham's Razor. It enjoins us not to multiply explanatory assumptions (entities, processes, causes, etc.) beyond necessity, in particular with respect to theoretical entities. It does not tell us, however, when such necessity obtains. Note that this principle is methodological, not ontological: it does not presuppose that nature is always and perhaps necessarily parsimonious, but that as inquirers we should begin with parsimonious assumptions. Note further that parsimony should not be readily equated with simplicity, such as the injunction to always prefer the simpler of two theories. After all, a theory can be simpler than another in many respects: it may be referentially simpler (having less qualitatively different referents), mathematically simpler, methodologically simpler (easier to test), or pragmatically simpler (easier to apply in a technological context). Simplicity in one such respect does not guarantee simplicity in another.

A second methodological principle is fallibilism or methodological skepticism. It is the acknowledgment of the fact that error is possible in all cognitive matters, so that our knowledge may be subject to criticism and, if possible, improvement and, if necessary, revision. We may highlight the latter by explicitly adding a "meliorist principle" [Bunge, 1983b] or a "principle of improvement of theories" [Kuipers, 2001].

d) Semantic assumptions

Most factual scientists maintain that their hypotheses, models and theories are true if they adequately represent the facts they refer to. That is, they subscribe to a correspondence theory of truth. Needless to say, the notion of truth is as tricky as many other concepts, so that there is no agreement among philosophers as to the appropriate truth concept in science [Weingartner, 2000]. Nevertheless, scientific realism is quite naturally associated with a correspondence concept of truth [Bunge, 1983b; Thagard, 1988; Devitt, 1996; Wilson, 2000]. Such a notion becomes easier to defend when we realize that the concept of correspondence truth provides just a semantic *definition* of "truth": it says nothing about how, and in particular how well, the truth of a hypothesis can be *known*. In other words, it does not provide a truth criterion. Truth criteria, such as evidential support, are not the business of semantics, but of methodology.

The concept of correspondence truth fits scientific practice even better when we realize that factual truth is in many cases not a dichotomy between true and false, but a matter of degrees. Models and theories often represent facts only in certain respects and moreover imperfectly so. Thus they correspond to facts only partially. Similarly, quantitative properties (represented by magnitudes) may be known only approximately, which is why scientists attempt to improve their measurement techniques. A realistic philosophy of science will therefore try to do justice to the idea of partial or approximate truth [Bunge, 1983b; Weston, 1992] and hence methods of truth approximation [Niiniluoto, 1987; Kuipers, 2000].

e) Axiological and moral assumptions

Most norms of science are built into its methodology. However, there are not only methodological values and norms, but also attitudinal and moral ones. Merton's [1973] expression "the ethos of science" captures this fact aptly, although his work is mostly concerned with attitudinal and moral norms that are not immediately relevant to the production of true knowledge (see below). To stress the fact that science has an internal system of values and corresponding norms, it may be useful to treat them all together. Thus, the researchers in a scientific field of knowledge are expected to accept the following values:

- *Logical values* such as the principle of noncontradiction and noncircularity. Together with the entire canon of valid reasoning, these are of course basic principles of rationality.
- *Semantical values* such as meaning definiteness, clarity, and maximal truth. Of course, a young or emerging scientific field may teem with vague and fuzzy concepts. But as it progresses and matures, in particular when it develops a theoretical branch, clarity and exactness are supposed to replace fuzziness. However heuristically fruitful vagueness may be in the beginning or in certain contexts, it may as well indicate that a field is degenerative rather than progressive.
- *Methodological values* such as testability (including the testability of the methods used in testing hypotheses, as well as the independent testability of auxiliary assumptions), explanatory power, predictability, reproducibility, and fecundity. Since these and other methodological categories are the main business of the philosophy of science, we shall not elaborate on them here.
- *Attitudinal*- and *moral values* such as critical thinking (or rationality in general), open-mindedness (but not blank-mindedness), universalism or objectivity (i.e., the requirement that ideas be evaluated independently of the personal, social or national characteristics of their proponents),

truthfulness, and acknowledgment of the work of others (e.g., by adequate citation).

As stated above, Merton's [1973] classic ethos of science concerns mostly attitudinal and moral values or norms, respectively. These are often abbreviated by the acronym CUDOS, which stands for four main norms: *c*ommunism (research results should be public property and accessible to everybody), *u*niversalism (see above), *d*isinterestedness (research should be uninfluenced by extra-scientific interests, and scientists should be emotionally detached from their subject matter), and finally *o*rganized *s*kepticism (scientists should be critical in particular towards their own work, and point out on their own weak spots or problematic parts). However, Merton's norms have been criticized for being too idealized and geared to an academic ivory tower situation (see [Ernø-Kjølhede, 2000] for an overview). Indeed, the history, psychology and sociology of science provide many examples that scientists have failed to follow one or more of these values. Like everyone else scientists are only human after all. Thus, individual scientists may be biased and jealous; they may intrigue against colleagues, or engage in nepotism; they often are emotionally attached to their subject matter in being passionate researchers, and they sometimes do not see the weak spots, if not flaws, in their own work; in particular, *pace* Popper, they are usually interested in having their hypotheses and theories confirmed, not refuted — after all, Nobel prizes are not awarded for the falsification of a theory. Moreover, the social and economic organization of scientific research has changed drastically during the past 50 years in that research institutions including universities are now run more like businesses, so that there is severe competition for funds and a strong pressure to focus on applied science and technology at the expense of basic science (see [Ziman, 1994]). For all these reasons Merton's classic ethos no longer describes realistically the behavior of scientists, however desirable his norms may still be from an ethical point of view (see also [Kuipers. 2001]). Finally, most of Merton's norms concern the professional social behavior of scientists in general, whereas the primary interest of the philosopher of science concerns those values and attitudes that are epistemologically relevant by contributing to gaining true knowledge, such as rationality, objectivity, and truthfulness.

In sum, the system of logical, semantical, methodological, and attitudinal ideals constitutes the *institutional rationality* of science [Settle, 1971], even though individual scientists may more or less often fail to behave rationally. (More on the problems of the rationality of science in [Kitcher, 1993].) And, however biased the individual scientist may be, the above values are also the basis for the *institutional objectivity* of science. As a consequence, basic science is value-free only in the sense that it does not make value judgments about its objects of study. In other words, basic science has no external value system.

This completes our extensive analysis of the philosophical outlook of a scientific field (condition 4), so that we proceed at last with our list of conditions characterizing an epistemic field as scientific.

5. The *formal background* F of a scientific field is a collection of up-to-date logical and mathematical theories used by the members of C in studying the items of D. This does not imply that scientificity is to be equated with formalization. All this criterion demands is that formal tools have to be handled correctly, and they must be adequate to tackle any given theoretical problem.

6. The *specific background knowledge* B is a collection of up-to-date and reasonably well-confirmed data, hypotheses, theories, or methods borrowed from adjacent fields. Every scientific field uses some knowledge from other scientific fields. For example, biology borrows knowledge from physics and chemistry. A science that borrows little from other fields is either very fundamental or very backward.

7. The *problematics* P is of course the collection of problems to be solved in the given field. It consists exclusively of epistemic questions on the nature and in particular on the lawful behavior of the objects in its domain D. It may also comprise problems concerning other components of its conceptual framework (e.g., the adequacy of methods, formalisms, and other background assumptions). If a discipline deals with practical problems, it is a technology, not a basic science.

8. The *fund of knowledge* K is a *growing* collection of up-to-date, testable and well-confirmed knowledge items (data, hypotheses, theories), gained by C and compatible with those in B. Even a young scientific field will possess some fund of knowledge, either taken over from ordinary knowledge or inherited from a parent science.

9. The *aims* A of the members of C of a field in basic science (as opposed to technology) are purely cognitive. They include, for example, the discovery and use of the laws of the members of D; the systematization of the knowledge in K (e.g., by constructing general theories); and the refinement of the methods in M.

10. The *methodics* M is a collection of empirical methods or techniques which may be used by the researchers in C in their study of the members of D, whereby "method" means a rule-directed procedure for collecting data or testing a theory. (Note that methods of reasoning, such as rules of inference or rules for evaluating theories, have been treated as belonging in G. Whence the distinction between methodics and methodology.) A scientific technique may be either concrete (i.e., involving instruments), such as electron microscopy, or conceptual (formal), such as the various statistical methods.

And they may be quite specific, such as Hennig's method of reconstructing phylogenies, or else more or less general, i.e., applicable in several fields, or for different purposes.

Among the methodological requirements for a technique to be scientific are the following. The functioning of these methods should be scrutable (e.g., by alternative procedures) and explainable by well-confirmed theories. (This may not be the case in a young field, but it should be achieved as the field matures. For example, when Galileo used his telescopes, optics was still too immature to fully explain their functioning.) And the techniques must be objective in the sense that every competent user is able to obtain roughly the same results.

It has been quite controversial whether there is such a thing as a scientific method in general (see, e.g., [Laudan, 1983; Haack, 2003]). If such a general method is expected to be a fool-proof procedure for delivering true and certain knowledge, then there is of course no such method. However, if we view the scientific method as an extremely general research *strategy*, then there may very well be a scientific method. For example, the sequence "problem–hypothesis–test–evaluation" reflects the general structure of any empirical scientific paper (introduction, methods, results, discussion), and may thus be seen as representing *the* scientific method [Bunge, 1983b]. However, if this definition is accepted, the scientific method is at best a necessary, but not a sufficient condition of scientificity: its application does not automatically turn one's inquiry into a scientific inquiry. Moreover, being extremely general, it is not an empirical method proper, so that it may as well be seen as belonging to the methodological rules in G.

In addition to the ten conditions of the ten-tuple $\langle C, S, D, G, F, B, P, K, A, M\rangle$ used in the preceding to characterize a scientific field, Bunge (1983b) requests that a scientific field satisfy two further conditions. These conditions take into account two aspects of science that have been emphasized by many philosophers of science: unity (consilience) and progressiveness.

11. The *systemicity condition*. There is at least one other field of research $\mathcal{S}'$ such that $\mathcal{S}$ and $\mathcal{S}'$ share some items in G, F, B, K, A and M; and either the domain D of one of the two fields $\mathcal{S}$ and $\mathcal{S}'$ is included in that of the other, or each member of the domain of one of the fields is a component of a system in the domain of the other [Bunge, 1983b, p. 198]. In simpler words, every scientific field has connections with other fields — a fact which allows for multi- and interdisciplinary research. This is due to the fact that nature is organized into several levels of complexity — levels that scientific disciplines may approach from various perspectives and with different aims and methods. Thus, despite all the differences in our cognitive interests, scientific disciplines form a network of approaches, striving for a unified — a consilient or convergent — view of nature, which need not be a reductionist

one [Kitcher, 1982; Bunge, 1983b; Bechtel, 1986; Thagard, 1988; Vollmer, 1993; Reisch, 1998]; for a dissenting view see [Dupré, 1993]. For this reason, new theories are evaluated not only on the basis of empirical tests, but also with regard to their overall compatibility with the well-confirmed background theories (external consistency). Although a new theory cannot by definition be compatible with every other theory, in particular its rivals, because it would otherwise not be a new theory, it must somehow allow to be accommodated within the totality of our knowledge. In Kuhnian terms: even if revolutionary, a new theory will cause only local or regional revolutions, never a total revolution turning upside down all existing fields at once.

12. The *changeability* or *progressiveness condition*. The membership of the conditions 5–10 changes, however slowly and meanderingly at times, *as a result of research* in the same field or as a result of research in neighboring disciplines. In Lakatosian words, the history of a scientific discipline must be progressive, at least on the whole. Even if science were to come to an end in the distant future, the history of a scientific discipline would have to show a certain amount of progress. (How the view that science is progressive can be defended against various antirealist objections has been shown by [Kitcher, 1993].)

This concludes the characterization of scientific epistemic fields. Note, firstly, that this characterization applies first of all to contemporary science, because many of its features have developed into their current state over the past 400 years. Consequently, it may not be fully applicable to 17th century science, for example. As for its future development, I doubt that the basic features and principles discussed above will evolve in a way that leads to their replacement by completely different principles, in particular their contraries. However, future development might consist in their improvement as well as in the discovery of some as yet unknown features and principles.

Note, secondly, that this characterization comprises both descriptive and normative aspects. Whereas the descriptive conditions provide diagnostic indicators, the normative ones will be the foundation for any judgment on the scientificity, or nonscientificity respectively, of an epistemic field.

What about science as a whole? Science as a whole is of course the totality of all individual scientific disciplines. If, as in the preceding, we represent each scientific field as a ten-tuple $\mathcal{S}_1 = \langle C_1, S_1, D_1, G_1, F_1, B_1, P_1, K_1, A_1, M_1\rangle$, $\mathcal{S}_2 = \langle C_2, S_2, D_2, G_2, F_2, B_2, P_2, K_2, A_1, M_2\rangle, \ldots, \mathcal{S}_n = \langle C_n, S_n, D_n, G_n, F_n, B_n, P_n, K_n, A_n, M_n\rangle$, science as a whole can be conceived of as the sum of these ordered sets: $\Sigma = \mathcal{S}_1 + \mathcal{S}_2 + \ldots + \mathcal{S}_n$. Similarly, we could characterize a *multidiscipline*, consisting of two or more scientific fields, as the sum of two or more ten-tuples representing them [Bunge, 1983b, p. 219]. In the case of a two-field multidiscipline this would be represented by: $\mathcal{S}_1 + \mathcal{S}_2 = \langle C_1 \cup C_2, S_1 \cup S_2, D_1 \cup D_2, G_1 \cup G_2, F_1 \cup F_2, B_1 \cup B_2, P_1 \cup P_2, K_1 \cup K_2, A_1 \cup A_2, M_1 \cup M_2\rangle$. (Note that we represent the concrete systems C and S by their composition, i.e., the set of their components.

Otherwise we would need an operation of physical or mereological addition rather than simply the one of set-theoretical union.)

By contrast, an *interdiscipline* does not just consist of at least two fields retaining their identity, but it is a merger of fields attempting to approach a common domain from a unified point of view rather than from different angles. Therefore, an interdiscipline may be conceived of as the *intersection* of two or more fields.

The analysis of a scientific field as a ten-tuple also allows us to elucidate the notion of a scientific research project. In section 4.1 we have defined the conceptual framework of an epistemic field as a septuple $\mathcal{S}_c = \langle G, F, B, P, K, A, M\rangle$. A *research project* π within a scientific field $\mathcal{S}$ characterized by a conceptual framework $\mathcal{S}_c = \langle G, F, B, P, K, A, M\rangle$ is then the septuple $\pi = \langle g, f, b, p, k, a, m\rangle$, where every component is a subset of the corresponding component of S_c [Bunge, 1983b, p. 176].

How does Lakatos's notion of a research program fit into this conceptualization? According to Lakatos [1970], a research program is a historical sequence of theories. Now theories surely belong to the fund of knowledge K of a scientific discipline. But we must also include the reference class of the theory belonging in D, as well as the formalism used to built the theory, which belongs in F. Further, Lakatos also counts auxiliary and other relevant assumptions as belonging to a theory. These may belong either in B or in K. Thus, a theory ϑ at any given time t might be construed at least as a quadruple $\vartheta(t) = \langle d(t), f(t), b(t), k(t)\rangle$, and a research program ρ over a period τ, where $\tau = [t_1, t_n]$, as an ordered set of such quadruples, $\rho(\tau) = \langle\langle d(t_1), f(t_1), b(t_1), k(t_1)\rangle, \langle d(t_2), f(t_2), b(t_2), k(t_2)\rangle, \ldots, \langle d(t_n), f(t_n), b(t_n), k(t_n)\rangle\rangle$. Depending on what we take to belong to a theory, we might as well regard a research program as a sequence of research projects as defined in the previous paragraph. Or, disregarding the historical focus of Lakatos's concept, we might simply redefine "research program" in the broad sense of "research project" or even "conceptual framework" or "disciplinary matrix" as explicated above (see, e.g., [Kuipers, 2001] for an even broader conception of "research program"). I take this broader approach to be more useful for demarcation purposes than Lakatos's idea of a series of theories in themselves.

So much for a possible characterization of the notion of a scientific epistemic field, which views science in the sense of basic factual science. It is now time to take a look at other research fields which, though not factual sciences, are related to them: mathematics, technology, and the humanities.

4.3 Other Research Fields

4.3.1 Mathematics

In contrast to the factual sciences, mathematics as well as formal logics and semantics are often called *formal sciences*. Although they have much in common with the factual sciences, the question is whether these commonalities justify to regard them as sciences. In other words, the question is whether we should use the label “science” in the strict sense of factual science or in a broader sense including formal science and perhaps technology.

Let us quickly analyze the status of mathematics with regard to the twelve conditions listed in Section 4.2. In so doing, we shall mention only those conditions that show significant differences.

Clearly, the domain D of mathematics shows an important difference with factual science: all the referents of mathematics are abstract objects. Although we can apply mathematical concepts and theories to concrete things, their properties and processes, we do so only by interpreting them in factual terms. In this way we represent factual properties in formal terms. Pure mathematics does not deal with concrete objects.

The philosophical background G of mathematics is also quite different. To begin with, mathematics can do without ontological realism: it would work just as well if there were no mind-independent reality. Of course, most mathematicians are *de facto* also ontological realists, but this is not a necessary assumption for doing mathematics: mathematics can be done on the basis of a Platonist, nominalist, or constructivist ontology (see, e.g., [Agazzi and Darvas, 1997]). Being just as ontologically neutral as logics [Nagel, 1956], mathematics has no use for the other ontological assumptions of factual science either, except for the principle of lawfulness. Indeed, mathematicians also assume that the referents of their discourse “behave” lawfully, whether they be found in a Platonic realm of ideas or whether they be constructed by our minds. Depending on the philosophy of mathematics adopted, the mathematical Platonist will need a form of epistemological realism, whereas the constructivist can do without it.

A major difference lies in the semantic concept of truth in mathematics: dealing with abstract objects and thus purely formal properties, mathematics is in no need of a correspondence theory of truth and hence can do with a coherence theory of truth (recall Leibniz’s *verités de raison*; see also [Bunge, 1983b]). Only the mathematical Platonists and empiricists may have use for a correspondence theory of mathematical truth. Still, mathematical truth is *de facto* established by formal coherence.

The methodological, attitudinal and moral values are by and large the same as in factual science. The major difference here lies in the notion of testability, which can only mean conceptual testability, not empirical testability. Moreover, testability in mathematics is stronger than empirical testability, because it allows for conclusive proof and disproof, whereas empirical testability only provides confirming or disconfirming instances.

As a consequence of the differences mentioned so far, there is another difference in the methodics M: mathematics uses no empirical, but only conceptual methods. (Even though some proofs obtained with the help of computers, such as that of the four color problem, may imitate empirical means in certain respects, they are still virtual and hence conceptual. Likewise, thought experiments, whether in mathematics or in the factual sciences, are conceptual means.) However, being extremely general, the scientific method, as defined in Sect. 4.2, seems to be used in mathematics as well.

As is obvious from the preceding, the main differences between mathematics and the factual sciences lie in the fact that it deals exclusively with abstract objects. On the other hand, mathematics too is a rigorous and progressive research field, consisting of a set of fruitfully interacting subfields.

4.3.2 Technology

In popular thinking, science and technology are often conflated. Worse, industrial production and marketing of technical goods is often equated with technology, which is in turn equated with science. So science gets often blamed for everything negative associated rightly or wrongly with the Western capitalist way of living. However closely these areas may be related *de facto*, the philosopher of science or of technology is of course interested in the question of whether science and technology can be distinguished *de jure*.

Borrowing again from Bunge [1983b], I shall propose the following distinctions. To begin with, the investigation of cognitive problems with *possible* practical relevance will be termed *applied science*. Thus, an applied science differs from its basic science partner mostly in its problematics (P) and aims (A). Further, its domain D will be narrower. For example, in contrast to human biology, medical research studies only those properties of humans that concern, directly or indirectly, matters of health. The same holds for clinical psychology as opposed to psychology in general.

If we now add the requirement that, on top of having discovered or studied some X which may be useful to produce (or else prevent) some Y, we actually *design* an artifact or a procedure to produce or else prevent Y, we arrive at technology. More precisely, *technology* may be defined as "the design of things or processes of possible practical value to some individuals or groups with the help of knowledge gained in basic or applied science" [Bunge, 1983b, p. 214].

Note first that, by making technology dependent on science, this definition distinguishes technology from the traditional crafts or *technics*, which are based solely on ordinary knowledge. Note further that this definition is so wide that it includes not only the classic fields of physical and chemical engineering, but also biological, psychological and social technologies. Thus, medicine, psychiatry, pedagogy, law, city planning, and management "science" are all technological fields.

Let us briefly review the coordinates of the ten-tuple $\langle C, S, D, G, F, B, P, K, A, M\rangle$ as to the differences between science and technology. As in the preceding section, only those showing significant differences will be mentioned. To begin with, although C is a research community, it is not as international and universalist as in the case of basic science, because patents and industrial secrets limit the circulation of technological knowledge. The domain D is both narrower and wider than in the case of applied science: it is narrower because it is concerned only with natural things which are useful for us, and it is wider because it includes not only natural things and processes but also artificial ones. The general outlook G

shares a realist and naturalist ontology and epistemology with basic science, as well as most of the other philosophical assumptions and values. The main difference lies in the fact that technology does not test so much for truth as for efficiency. Truth is relevant only as a means for design and planning. Finally, the ethos of technology differs from that of basic science: usually, it consists not in the free and disinterested search for knowledge, but in task-oriented work, often depending on the economic interests of some employer (see also [Ziman, 1994]). Obviously, the problematics P and the aims A are among the main differences: the problems and aims are practical rather than cognitive. Moreover, the aim of technology is not to discover new laws: it suffices to make use of known ones. Finally, technology is characterized by a coordinate of its own: in contrast to basic science, technology has not only an internal value system, but also an external one (V). That is, it attributes positive or negative values to natural or artificial things or processes, be it raw material or finished product. Thus, a technology is actually characterized by an eleven-tuple $\langle C, S, D, G, F, B, P, K, A, M, V \rangle$.

4.3.3 Humanities

In contrast to the social sciences, which study social systems (composed of human individuals) and their activities by empirical means, the humanities mostly abstract from these concrete individuals and groups as well as their activities and study their intellectual (including artistic) products, i.e., ideas or concrete artifacts. Inasmuch as the humanities study the activities of groups or individuals, these are usually of an artistic nature, such as a theatrical or musical performance. Accordingly, literature and literary criticism, languages (philology) and part of linguistics, art history and criticism, musicology, the history of ideas, religious studies, and philosophy belong to the humanities. On the other hand, some fields like history and archeology, as well as the history and sociology of religion belong — or should belong — to the social sciences. Similarly, part of linguistics is a social science too. And according to our classification, the law (jurisprudence) and pedagogy are not humanities but sociotechnologies (Sect. 5.2). These examples show that quite often there is an overlap between some social sciences and the humanities. In particular, some fields starting out as humanities may develop into sciences.

Again, a quick review of the ten coordinates of an epistemic field will be in order. To begin with, the humanities are clearly research fields with a specialized research community C. As just mentioned before, their domain D consists of ideas and artifacts rather than natural things and processes. Consequently, the humanities are consistent with either a naturalist-materialist or a Platonist outlook. As for epistemology, the natural approach is most likely a constructivist one, which can be either realist or antirealist. Furthermore, the humanities are open to the influence of subjectivist philosophies like phenomenology and hermeneutics. (And of course, in the field of philosophy, which has to provide its own metaphilosophy, just anything goes.) In sum, the philosophical outlook of the humanities is much more variegated than that of the sciences, and necessary connections, if any, with particular philosophical presuppositions are much less obvious. Presumably, the more aspects of the much straighter scientific outlook are adopted, the better the chances of bridging a humanistic field with a scientific one. Think of linguistics and comparative religion (*Religionswissenschaft*), which make contact with sociology, history, evolutionary biology, psychology and, more recently, even the neurosciences.

As for methodology and semantics, since the humanities deal with ideas and artifacts, which are not to be explained by natural laws and mechanisms but instead interpreted and comprehended, it is unclear which role the parsimony principle plays in the humanities. More complex views and interpretations may be preferred to simpler ones, just as conversely. Similarly, fallibilism may not be that important because there may be different reasonable perspectives and interpretations, without implying that therefore one of them is erroneous. Consequently, the notion of truth in the humanities is often contextual or relative rather than factual. The fact that Othello killed Desdemona is (fictionally) true only in the context of Shakespeare's story. Another author could easily write an alternative play in which Desdemona kills Othello, so that in this context the opposite would be true. On the other hand, inasmuch as the humanities are descriptive of certain (e.g., historical) facts, these descriptions can be correct or not in the correspondence sense.

What about the internal value system of the humanities? Rationalist humanities will certainly respect the standard logical values. But there are also irrationalist branches, in particular in philosophy and certain postmodernist cultural studies (see Sect. 5.2). Very often the semantical values of clarity and exactness cannot be heeded. This is due to the very nature of human thought and communication, which is far from unambiguous, whence the need for interpretation arises. However, if these semantical values are not accepted even as remote ideals, and fuzziness is instead turned into method, the line to obscurantism may easily be crossed.

Evidently, the methodological values of testability and explanatory power in the scientific sense are not part of the humanities. A certain view, reading, or interpretation may be open to criticism, but since it is neither true nor false, it cannot be tested for truth. At most, it is reasonable, plausible, sensible, or apposite. Explanatory power may be replaced by "comprehensive power" if we admit the hermeneutic goal of understanding in the humanities. On the other hand, fecundity is certainly also a value in the humanities, because humanistic understanding can be increased if some approach opens up new perspectives.

Whereas some attitudinal values are of course the same as in the sciences, others are different. For example, just as there may exist competing theories in the sciences, there may be competing interpretations in the humanities. Honesty requires at least mentioning the existence of such competing approaches, even though the researcher wants to focus on her own. The same holds for the adequate citation of sources, although the standards appear to be lower than in the natural sciences. For example, it seems to be much easier to survive peer-review when disregarding the work of disliked colleagues in a philosophical article than in a science paper. Furthermore, the value of universalism plays only a minor role, if any, in the humanities. For, naturally, the humanities are more inclined towards relativism, because many cultural items cannot be evaluated independently of the personal and cultural characteristics of their creators: they must be seen and understood in context. Finally, like technology, many humanities have an external value system: they attribute, for example, aesthetic values, meanings, and purposes to the objects in their domain D, because the latter are studied in their relation to humans.

The formal background of the humanities, if any, is of course small. Exceptions occurring for example in philosophy, such as mathematical logics and formal semantics, may be classified as formal sciences. On the other hand, other branches of analytical philosophy too are formal (like ontology), which indicates that they are science-oriented, though not full-fledged sciences.

The aims of the humanities can be either cognitive or practical, or both. In contrast to the sciences, however, they usually do not seek to find laws. Indeed, the "sciences of the mind" (*Geisteswissenschaften*) have been regarded as descriptive (idiographic) rather than law-finding (nomothetic). On the other hand, we have seen before that some humanities make contact with the sciences, so that such multi- and interdisciplinary ventures may be able to find some cultural or even aesthetic laws.

Obviously, a major difference with the sciences is found in the methodics M of the humanities. Naturally, except for some observation, their methods are mostly conceptual. Among these are some general methods unique to the humanities, such as the hermeneutic and dialectic "method" [Poser, 2001], although these are not methods in the strict sense of rule-guided procedures to attain a certain goal. (Here, "hermeneutics" does not mean philosophical hermeneutics, but only the traditional concept of text interpretation, or understanding of works of art, respectively. And the dialectic method concerns first of all the discoursive triad thesis-antithesis-synthesis, without presupposing the whole of dialectic philosophy.) If not objective in the sense that every competent user will get roughly the same results, these "methods" are at least intersubjective in that their results can be communicated to, and understood by, other people. The humanistic scholar may also borrow or apply certain techniques from the factual sciences, but this does not yet turn her field into a science. For example, the art historian may have some paint chemically analyzed, or some cloth radiocarbon-dated, without thereby changing the nature of her discipline.

In sum, compared to formal science and technology, the humanities show the greatest distance from factual science. But again, we emphasize that this is not a value judgment. When saying, for example, that the arts and humanities are not scientific, nobody claims that they are therefore objectionable or bad.

4.4 Conclusion

The factual and formal sciences, the technologies, and the humanities are all research fields producing genuine knowledge, which on the whole is either (approximately) true or else useful, and contributes to the understanding of the world and its inhabitants. For this reason, one might argue that they should all be included in a broad conception of science. This is for example done in the German intellectual tradition, where the name of almost any field of knowledge is dignified by the ending "-wissenschaft" (-science), including the humanities, which are called *Geisteswissenschaften* (sciences of the mind). So there is bioscience alongside "music science", just as there is computer science alongside "literature science". Consequently, if a practitioner of a *Geisteswissenschaft* is told that what he does is not science, he will most likely be offended. It comes as no surprise that such a broad, if not inflationary, construal of "science" aggravates the problem of demarcation (see, e.g., [Poser, 2001]).

By contrast, most other traditions and languages separate the arts and humanities from the sciences already terminologically, so that no offense is given by calling the humanities nonscientific. Yet even so, the question remains of what to do with mathematics and technology. While some authors include both of them in the sciences (e.g., [Kuipers, 2001] classifies them as explicative research programs and design programs, respectively, within a broad conception of a *scientific* research program), others assert that neither mathematics [Lugg, 1987] nor technology [Bunge, 1983b] are sciences. In any case, taking into account the preceding overview, the common post-positivist picture, which admits more categories than just sense (i.e., science) and nonsense (i.e., all the rest), may look like the one given in Fig. 2. One the one hand, there is science including mathematics and technology; on the other there is nonscience including the arts and humanities as good nonscience, so to speak, for it too is viewed as producing true, reliable, or at least valuable knowledge, respectively, and finally pseudoscience as bad nonscience, for its knowledge claims are unjustified.

We may refine this picture by adding protoscience and prototechnology, as well as ordinary knowledge. These straddle the lines between pseudoscience and science. A protoscience is expected to develop into a science proper by leaving behind its nonscientific (or even pseudoscientific) roots (see Sect. 6). And ordinary knowledge is mostly nonscientific and reliable, but contains illusory items on the one hand, and some knowledge adopted from the sciences on the other (Fig. 3). It is the task of the science educator to increase the share of the latter and to decrease that of pseudoscience and superstition.

We shall further refine this picture later on to reflect the distinctions made above between factual science, mathematics, technology, and the humanities. Before, however, we need to take a closer look at that kind of knowledge which is not just nonscientific but in fact unscientific or pseudoscientific.

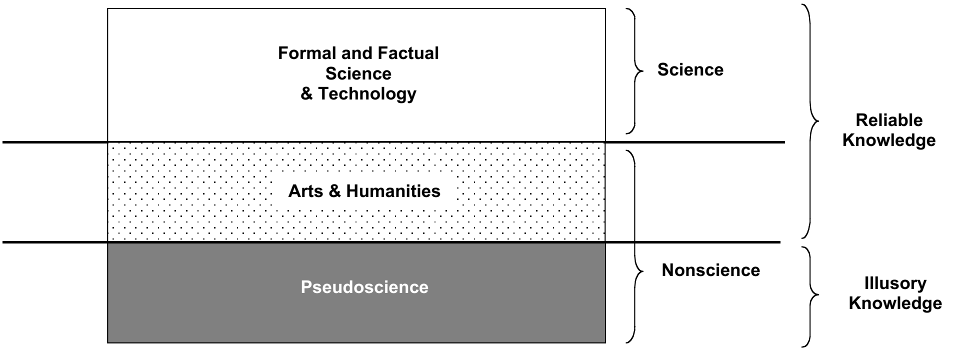


Figure 2. A common post-positivist picture of science and nonscience. As scientific research fields, mathematics, factual science (including psychology and social science), and technology are subsumed under the general label of "science". Nonscience divides into the arts and humanities (including philosophy) on the one hand, producing reliable or at least valuable knowledge, and pseudoscience on the other, offering nonreliable or illusory knowledge.

5 UNSCIENTIFIC FIELDS

As emphasized previously, calling an epistemic field *nonscientific* is not pejorative but descriptive. Calling it *unscientific*, however, is judgmental: it indicates that the given field cannot live up to its cognitive claims. Since there is no noun "unscience", an unscientific field is called a "pseudoscience". As usually defined, a pseudoscience is a particular form of nonscience, namely a nonscientific field whose practitioners, explicitly or implicitly, *pretend* to do science. Thus, to say that a field is pseudoscientific amounts to saying that it is a fake. In other words: While there is reliable or, if preferred, approximately true theoretical and practical nonscientific knowledge, the knowledge produced by pseudoscience is illusory. And since spreading bogus knowledge amounts to deception, pseudoscience has a moral dimension that other nonscientific fields lack. Therefore, a demarcation of science versus nonscience in general does not yet tell us how legitimate nonscientific fields are to be demarcated from pseudoscientific ones.

5.1 Characterizing Pseudoscience

For this reason, several authors have attempted to give not only a characterization of science as opposed to nonscience, but also of pseudoscience in particular [Thagard, 1978; 1988; Radner and Radner, 1982; Bunge, 1982; 1983b; 1984; Grove, 1985; Lugg, 1987; Derksen, 1993; 2001; Hansson, 1996; Wilson, 2000; Kuipers, 2001]. It will come as no surprise that the criticisms of such attempts parallel those leveled against any quick and clear-cut demarcation of science: though dealing with important aspects of pseudoscience, the proposed demarcation crite-

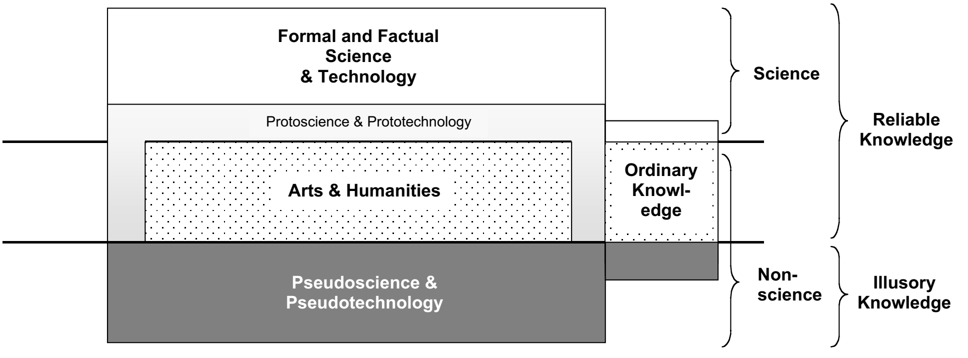


Figure 3. A refined post-positivist picture of science and nonscience, making room for ordinary knowledge as well as protoscience and prototechnology, which range from the pseudoscientific to the scientific.

ria do not combine to form a set of necessary and sufficient conditions, because they always leave some pseudosciences unscathed. Let us briefly review some such demarcation attempts.

Improving on his earlier demarcation proposal [Thagard, 1978], Thagard [1988, p. 170] contrasts his five characteristics of science mentioned in Section 3 with five features typical of pseudoscience. In pseudoscience, scientific correlation thinking is replaced by primitive resemblance thinking; empirical matters of confirmation and disconfirmation are neglected; practitioners of the field are oblivious to alternative theories; the theories are nonsimple and contain many ad hoc hypotheses; and there is no progress in doctrine and application. Thagard points out that these are indicators of pseudoscientificity, not necessary and sufficient criteria.

Grove [1985] gives four characteristics of pseudoscience. The first is the lack of an "independently testable framework of theory capable of supporting, connecting, and hence explaining their claims" (p. 237). The second is the lack of progress. Third, a pseudoscience is usually constructed in such a way that it is able to resist any possible counter-evidence; in other words, it is practically irrefutable (though it may be logically falsifiable). And fourth, according to Grove, not just irrefutability is a mark of pseudoscience but, more generally, their "total resistance to criticism".

Lugg [1987, p. 228] suggests regarding pseudosciences as "radically flawed practices, i.e., as radically flawed complexes of theories, methods and techniques". He maintains that, in the case of the pseudosciences, empirical matters are relatively unimportant, because their being conceptually flawed makes them unworthy of serious attention, whether or not their claims could actually be confirmed or disconfirmed. This is similar to Rothbart's [1990] claim that pseudoscientific theories are not testworthy. If we can already show by means of formal or informal logic that an argument or an approach is fallacious, there is no need to empirically test the hypotheses involved. Finally, according to Lugg, if pseudosciences are practices, they are social institutions, and realizing that they are such helps to explain their longevity and resilience.

Rationalistic approaches, such as Lugg's and Rothbart's, are likely to be rejected for smacking of dogmatism by those inclined towards empiricism. Can we really declare some theory untestworthy in an apriori manner? Is not empirical confirmation or disconfirmation the final arbiter of a theory? For example, Thagard [1988, p. 170] generously admits that, despite all the previous failures of astrology, future studies might find empirical support for astrology, although he takes that to be rather unlikely. By contrast, Kanitscheider [1991] maintains that there can be no such evidential support, because astrology is so defective theoretically that, even if there were strong empirical correlations between the star positions and human character and fortune, it could never explain these data by way of mechanisms that do not involve sheer magic. In other words, the empirical situation is irrelevant if the theory in question cannot even begin to explain the data at hand.

Derksen [1993] rejects the idea that it is theories, practices, or entire fields that are pseudoscientific. Instead he recommends examining the attitude or the pretensions of the individual pseudoscientist. After all, it is not a field that can have scientific pretensions, but only its practitioners, and only the latter can be blamed for not making good on these pretensions. Similarly, Kitcher [1993, p. 196] holds that "[t]he category of pseudoscientists is a psychological category. The derivative category of pseudosciences is derivatively psychological, not logical as philosophers have traditionally supposed. Pseudoscientists are those whose psychological lives are configured in a particular way. Pseudoscience is just what these people do." Whereas Kitcher has in mind the inflexible epistemic performance of American creationists, Derksen's analysis concerns the work of Freud. In his analysis Derksen [1993] lists seven attitudinal sins of the pseudoscientist. The first is the "dearth of decent evidence". Having scientific pretensions, the pseudoscientist will have to show respect for empirical evidence. But what he claims to be good evidence for his theory is in fact defective. For example, it is unclear how reliable Freud's clinical data are, because he did not ensure that they were not the result of his own suggestive questioning. (See also [Grünbaum, 1984].) The second sin consists in "unfounded immunizations", which result from selecting and tailoring the data until they fit the given theory; in other words, only particular interpretations of the data are accepted. This also happens in science but, there, immunization is based on well-confirmed theories rather than on unfounded ad hoc hypotheses.

Derksen calls the third sin the "ur-temptation of spectacular coincidence", which consists in ascribing a deeper significance to *prima facie* spectacular coincidences. The fourth sin is the application of a "magic method". That is to say, the pseudoscientist always has some magic method at hand by means of which he can generate all the data he needs. With regard to Freud, Derksen mentions the method of free association, the analysis of symbols and the interpretation of dreams, by which Freud was able to get any data he needed to support his ideas.

The fifth sin is the "insight of the initiate". This is not the claim that only the person with a specialized training can do proper research, since this holds also for science. Rather, it is the claim that the researcher has to overcome certain impediments and prejudices in order to be able to gain the knowledge and insight to be had in the given field. Thus, only the Freudian who underwent psychoanalysis himself is said to be able to practice psychoanalysis.

The sixth sin refers to the presence of an "all-explaining theory", i.e., "a theory that has ready answers to whatever happens". The seventh sin, finally, consists in "uncritical and excessive pretension". Here, "excessive" refers to the fact that, first, the pseudoscientist claims a much greater reliability of his knowledge than allowed for by the evidence (or rather the lack of evidence), and, second, that his pretensions concerning the importance of his theory are far too great. In a later paper, Derksen [2001] elaborates on these sins, offering seven further strategies typical of the "sophisticated pseudoscientist". In any case, although Derksen is right that, strictly speaking, only a person can have scientific pretensions, it seems rather unproblematic to abstract from these individual "sinful attitudes" and treat them as methodological rules, as is commonly done. The same holds in my view for Kitcher's [1993] psychologistic approach.

In a complex study of scientific research, which can be summarized only in a rather simplified way, Kuipers [2001, p. 247] defines pseudoscience as the combination of scientific pretensions and the neglect of the "principle of improvement of theories". The latter enjoins us to aim at more successful theories by eliminating the less successful ones. This improvement is supposed to occur within a research program (in the broad sense), i.e., we aim at better theories while keeping the hard core of the program intact. Only if this strategy fails should we try to adapt the hard core; and only if this strategy fails too, should we look out for a new research program. According to Kuipers, these rules may be seen as constituting *scientific* (or *methodological*) *dogmatism*. By contrast, *unscientific dogmatism* is characterized by the strict adherence to one or more central dogmas which are deemed to be in no need of improvement.

Although these authors do not quite agree on the characterization of pseudoscience, they provide important indicators of pseudoscientificity, useful for any analysis of any theory, practice, or field suspected of being pseudoscientific.

5.2 Pseudoscience or Parascience?

There is a fundamental problem, however, with the very definition of the term "pseudoscience". If it is an essential connotation of "pseudoscience" that it be a nonscientific field *with scientific pretensions*, what do we do with nonscientific fields that appear to be as defective as the classic pseudosciences, but do not claim to be scientific in the first place? As Hansson [1996] has rightly pointed out, many fields that are often subsumed under the label "pseudoscience" are not really such. Indeed, many areas in the vast realm of esoterics, occultism and New Age thinking do not pretend to be scientific at all. Some are even outright antiscientific: they reject the scientific approach to knowledge in favor of various "alternative ways of knowing". If not as completely wrong, the scientific world view is regarded at best as short-sighted and hence in dire need of "complementary" forms of cognition, such as "holistic", "spiritual" or "mystical" ones. Examples of such fields are various forms of "alternative healing" such as shamanism, or esoteric world views like anthroposophy (for further examples see [Carroll, 2003; Hines, 2003]; as well as the various articles in [Stalker and Glymour, 1989]). Obviously, the standard definition of a pseudoscience as a nonscientific field with scientific pretensions does not apply to such areas. Yet these esoteric fields do compete with science in claiming to produce, or have at their disposal, important factual knowledge that the "narrow-minded" scientific approach necessarily must overlook. Moreover, the alleged knowledge produced in these areas often collides head-on with well-confirmed scientific knowledge. For this reason, we must suspect that the "alternative knowledge" produced in such fields is just as illusory as that of the standard pseudosciences.

For these reasons it will be useful to have a different term which subsumes both the pseudosciences proper and all the other fields producing bogus knowledge. I suggest using the term *parascience* for this purpose. Note, though, that the term "parascience" is often used in a different sense, namely descriptively for a field of knowledge whose status as either a pseudoscience or a protoscience is still under debate. I shall disregard this descriptive usage here in favor of the normative one. Alternatively, we could as well give up the standard meaning of "pseudoscience" as a nonscientific field with scientific pretensions and conceive it in a broader sense to also cover all those areas dealing with bogus knowledge.

However, I shall stick here to the name "parascience", because it allows us to explore further distinctions, which are usually neglected in the demarcation literature. Thus, as a matter of principle, we can not only distinguish science from pseudoscience, but also pseudotechnology from paratechnics, and pseudohumanities from parahumanities. Recalling our earlier distinction between technology and technics, a pseudotechnology then would be a technological field based on some pseudoscience, whereas a paratechnic would just be a crackpot technic without any elaborate pseudoscientific background, or at most with a traditional magical background theory. A pseudohumanistic field would be one pretending to produce humanistic knowledge, although its business actually consists in sheer intellectual imposture or obscurantism. And a parahumanistic field, finally, would be the same, except for the fact that it does not pretend to be a field which should belong in the circle of the humanities. Finally, there is a category which contains all those fields that are neither pseudoscientific nor pseudo- or paratechnological nor pseudo- or parahumanistic. We have no choice but to call them parasciences in the narrow sense, in contradistinction to parascience in the broad sense as defined above (see Fig. 4). Having two notions of parascience is one of the disadvantages of the present analysis.

To see whether this extended typology is of any use, let us take a look at some examples. Considering these examples here does not imply that all of them

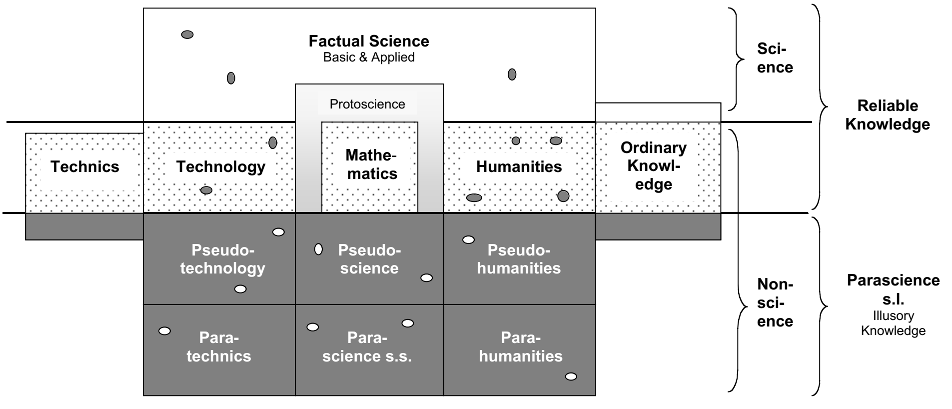


Figure 4. An extended typology of epistemic fields. In this typology only the basic and applied factual sciences are considered as strictly scientific, whereas technology, mathematics and the humanities are classed as nonscientific fields, though still close to the factual sciences. In any case they belong to the class of epistemic fields providing reliable knowledge. By contrast, the knowledge claims of the parasciences (*sensu lato*) are illusory: they do not enrich human knowledge, but pollute it. Protosciences are epistemic fields shading from the dubious into the scientific. The light gray shading indicates that by and large they are on the right track, although they are still burdened with nonscientific ideas or procedures. Ordinary (or everyday) knowledge and technics also lie in between the reliable and the mistaken. Note the gray spots on science's bright vest and the white spots on the dark attire of the parasciences. This indicates that the science/parascience distinction is not really a clear-cut black and white demarcation line, as suggested by this idealized diagram. There are pseudoscientific pockets within otherwise good sciences. These are sometimes labeled *pathological science*. And of course, some knowledge produced in science, technology, and the humanities has turned out to be false (without therefore being pseudoscientific), and not all knowledge in the parasciences need be false. Further explications in the text.

are correctly placed in the proposed category. Some of them certainly are, but the status of others is still under debate, so we may prefer to call them parascience *candidates*. Standard examples of pseudoscientific theories or fields are parapsychology, scientific creationism and intelligent design, psychoanalysis (as basic psychological theory), astrology (as a theory of human character), cryptozoology, Lyssenkoism, New Age physics, ufology, Däniken's archeology, Afrocentric history, and Sheldrake's morphogenetic fields theory (see [Shermer, 1997; 2002; Carroll, 2003; Hines, 2003]). A more recent suspect is the constructivist-relativist sociology of science [Gross and Levitt, 1994; Sokal and Bricmont, 1998; Bunge, 1999; Wilson, 2000]. All these fields pretend to be scientific, e.g., in using scientific methods.

By contrast, a parascience (in the narrow sense) does not claim to be scientific: it is just a field involving some (often traditional) theory about certain matters of facts. For example, traditional Chinese medicine involves a "biological" theory of the life energy *qi* flowing in meridians through the human body. The Indian theory of chakras asserts that the human body contains thousands of energy centers (chakras), which may be influenced by meditation (e.g., tantra). Similarly, the Western esoteric theory of reincarnation states that a personal soul really survives the body's death and can be reborn in some other body. (Note that the traditional Buddhist concept of reincarnation does not involve the survival of some spiritual substance.)

As for pseudotechnology, recall from section 4.3.2 that technology does not just consist of the classic physico-mechanical or engineering disciplines, but also of biological, psychological, and social technologies. All the fields attempting to come up with perpetua mobilia and other so-called free energy machines, with antigravitation devices and earth ray protection gizmos, count as pseudo-physicotechnologies. Likewise, sophisticated dowsing, which is based on pseudogeological assumptions, and water energizing on the basis of "quantum transformation" or other bogus concepts belong in pseudo-physicotechnology.

Examples of bio-medical pseudotechnologies are homeopathy, chiropractic, iridology, and biorhythmology. Candidates for psychological pseudotechnologies are psychoanalytical therapy, phrenological and graphological diagnosis, astrotherapy and horoscopes, neurolinguistic programming, and applied kinesiology. Finally, as pseudo-sociotechnologies have been regarded: Marxism as scientific socialism [Popper, 1959] as well as feminist technology and the so-called New Evidence Scholarship relying on subjective probabilities in jurisprudence [Bunge, 1999]. By contrast, mere paratechnics, i.e., procedures not based on some pseudoscience but at best on some parascience (in the narrow sense), are naive dowsing, faith healing, magic, voodoo, and prophetic techniques such as palmistry, Tarot, and I Ging.

What about pseudo- and parahumanities? Are there any examples at all? In section 4.3.3 we listed only some of the major differences between the humanities and the factual sciences. Since this does not constitute a positive and comprehensive characterization of the humanities, it does not enable us to demarcate genuine humanities from pseudo- and parahumanities. Thus, the following examples merely give some possible suspects, not the results of a detailed analysis. As pseudohumanities have been regarded: anthroposophy, theology, irrationalist philosophy (pseudophilosophy), and postmodernist cultural studies. Scientology may be another candidate. Parahumanities on the other hand might be hermetics, gnosticism, mysticism, and maybe traditional religions inasmuch as they make cognitive claims. These examples show the highly controversial nature of demarcating pseudo- and parahumanities. Even if this demarcation proves to be untenable or useless, it should at least provoke a detailed examination of the suspects involved before admitting them into the humanities or else refusing them entry.

Indeed, only few authors (e.g., [Kuipers, 2001]) have dared ask the question of whether, for example, theology is a pseudoscience, and whether there is such a thing as pseudophilosophy. Whereas Kuipers does not give an answer with respect to theology in his 2001 (see, however, [Kuipers, 2004]), he suggests that pseudophilosophy is the combination of philosophic pretensions with unscientific dogmatism. Philosophy reducing to nothing but exegesis, or the attempt to preserve the teachings of some master instead of developing and improving on them, would be examples of pseudophilosophy. Another example, not mentioned by Kuipers, could be irrationalist philosophy. For example, it is well known that Schopenhauer and many others accused Hegel of being a pseudophilosopher for writing utter nonsense, and the positivists, the critical rationalists and others have criticized some of the German philosophical tradition (e.g., Heidegger) for being obscurantist (see, e.g., [Albert, 1985; Edwards, 2004]). And recently, the French deconstructionists and others have been accused of being intellectual impostors [Sokal and Bricmont, 1998]. Be this as it may, if there is pseudophilosophy, it will be a pseudohumanistic field rather than a pseudoscientific one.

Theology is somewhat different, because the work of theologians ranges from the social sciences to the humanities. While working, for instance, in the field of comparative religion, text analysis, or sociology of religion, theologians do proper scientific and humanistic work — *de facto* and as individual researchers. Hence their individual work need not differ from religious studies or comparative religion (*Religionswissenschaft*), which can just as well be done by nontheologians. Presumably, the main problem with theology is institutional, because theology is by its very essence denominational: the theologian is the representative of some particular religion and is therefore expected to accept its creed as a given. The core of this belief system is not open to revision as a matter of principle, wherefore it must be regarded as a form of unscientific dogmatism. Thus, it is impossible that, as a result of internal progress in research, Christian theology will come to the conclusion that Christianity is actually false and Hinduism is true after all. For example, in the past 200 years the research of many theologians has contributed to demolishing the authority of the scriptures by putting them in a proper historical perspective, but this has not led them to abandon Christianity. Rather, it has spawned a hermeneutic industry of apologetics, attempting to save the Christian faith by reinterpreting and re-reinterpreting its tenets, often in unintelligible terms [Albert, 1985, Ch. 5]. Of course, the individual theologian may eventually change his mind and give up his belief, adopting another one or even becoming an atheist. But, unless he gets fired upon so doing, he has to leave his field if he wants to be consistent. Thus, it seems that, due to its fundamentally denominational and dogmatic nature, theology as an epistemic field is pseudoscientific or pseudohumanistic, respectively.

What about pathological science? In which category does it belong, or is it a category of its own? As mentioned in the legend of Fig. 4, pathological science concerns pockets or niches of pseudoscience still located within the sphere of science. In Fig. 4, this is indicated by the dark spots marring the field of science. Classic examples are the N-rays and polywater affairs. More recently cold fusion has been added to this list. But other theories and approaches within the sciences too have been regarded as pseudoscientific, such as steady state cosmology, the anthropic principle, the subjectivist interpretation of quantum theory, the quantum theory of measurement, evolutionary psychology, information processing psychology, and the research on race and IQ (see, e.g., [Bunge, 1982; 1983b; 1984; 1999; Shermer, 2002]). Some fields, like holocaust denial, have even somewhat branched off from academic historiography to form a specialized field of their own, which enforces the impression that they have turned into full pseudosciences [Shermer, 1997].

As for the corresponding white spots in the parascientific fields, they indicate that not every piece of knowledge in the parasciences need be false: we may find some true or useful items on occasion. An example is acupuncture. Although there is no hope for the magical theory of traditional Chinese medicine underlying the practice of acupuncture, there is some evidence that putting needles here and there has some effect on relieving certain forms of pain [Ernst *et al.*, 2001]. If this turns out to be true, acupuncture will become an area of biomedical research and explanation, which most likely will not have much in common with its parascientific origins. Finally, some parasciences, such as parapsychology, do use scientific methods for example, so that not everything occurring in an overall parascientific field need be unscientific.

So much for some possible examples illustrating the distinctions suggested in Fig. 4, and some qualifications concerning the idealizations involved. The purpose of this extensive typology is to show that in its standard definition the label "pseudoscience" fails to do justice to the wide variety of the parasciences. On the other hand, if we are only interested in distinguishing the genuine article from bunk, a simpler analysis will of course do, such as the one depicted in Fig. 3, in which, however, one might want to replace the terms "pseudoscience" and "pseudotechnology" by "parascience" and "paratechnology", respectively.

Having dealt with various parascience suspects, let us proceed at last with the characterization of parascience.

5.3 Characterizing Parascientific Fields

In the following analysis we shall try to develop a profile of parascience (in the broad sense) by applying the twelve criteria of scientificity listed in Sect. 4.2.

1. *Community C.* Faced with a parascience candidate, we need to examine whether there is in fact a real research community continuing a research tradition, or just a loose collection of individuals. If there really is a genuine system of persons, we need to check further whether this community engages in research, or whether it is just a group of believers.

   One of the few parasciences that does have a research community is parapsychology. Many others, by contrast, are belief communities: there is a single guru or a small number of authorities, surrounded by a more or less numerous crowd of followers, who do not engage in research, but at most in exegesis or application. Think of Immanuel Velikovsky's pseudocosmology, Erich von Däniken's pseudoarcheology and pseudohistory, Charles Berlitz's Bermuda triangle mystery, or Ron Hubbard's scientology.

2. *Society S.* The society hosting a community of researchers or else believers must at least tolerate its activities. However, political power can turn an epistemic field into a pseudoscience if it starts to proscribe what is to be accepted as true knowledge and what not, and if the people working in that field follow suit. Examples are *Deutsche Physik* (German physics) or, more generally, Aryan science in the Third Reich, and Lyssenkoism during the time of Stalinism and after. A contemporary example is creationism, which is adopted at the national level in official theocracies, or at least pushed at the regional or local level where conservative churches or fundamentalist religious groups of any color wield enough power (e.g., in Turkey, Iran, the US, and Russia). In the same vein, it is legitimate to ask whether the calls for a feminist science, based on the relativist-sociological "finding" that science is just an enterprise of white Western males, belong in the same category [Gross and Levitt, 1994; Bunge, 1999]. It may well be that women have somewhat different research interests, so that they focus on different problems. But as soon as we get to questions of method, testing, validity, and justification, there seems to be no leeway for "alternative" forms of science.

3. *Domain D.* The domains of parascientific fields often comprise dubious and ill-defined items, such as mysterious energies or vibrations, which have so far escaped detection. In other words, many parasciences still have to prove that the objects and processes they refer to in their discourse do exist really. Therefore, much of their domain is factually empty and consists mostly of speculative entities. An example is parapsychology, which has not been able to come up with a single unambiguous finding concerning the real existence of "psi" [Alcock, 2003; Hines, 2003].

   At first sight, hypothesizing unobserved or unobservable entities appears to be analogous to the theoretical entities posited in many scientific fields.

However, the difference is ontological, semantical and methodological: if not supernatural, the entities posited in many parascientific fields are by definition paranormal or, if preferred, paranatural, and they are often idle, arbitrary, or nonparsimonious, for not being embedded in some explanatory theory proper. Hence they are often ill-defined, i.e., they are so vague that it is unclear what is being tested — if there are serious tests at all. An example is the mysterious "psi" occurring in parapsychology, which is defined but negatively [Alcock, 2003]. For example, precognition is defined as seeing future events in a way that *cannot* be explained by contemporary science. Likewise, psychokinesis and telepathy involve interactions that *cannot* be accounted for by any mechanisms known to normal science. Moreover, parascientific entities are not hypothesized in a search for the best explanation (i.e., abductively, as it is often called), but they are often objects of prior beliefs, for which a justification is sought only if the belief is questioned by some skeptic. So whatever *prima facie* explanatory function they may have, the very same function could often be exerted by any other paranatural entity. In other words, paranatural entities are usually not specific enough for a satisfactory explanation (see, e.g., [Flew, 1990; Kanitscheider, 1991; Humphrey, 1999]).

4. *Philosophical background G.*

   (a) *Ontology.* The ontological aspects of parascience are often neglected in favor of its methodological problems. An early exception was the philosopher Charlie D. Broad, who was a firm believer in parapsychology. He pointed out that both science and our everyday practice presuppose various philosophical assumptions, which he called "basic limiting principles" [Broad, 1949]. He gave four main examples, three of which are ontological, one epistemological. His ontological principles were (i) the antecedence principle (effects cannot precede their causes); (ii) mind cannot directly act on matter without involving a brain event; and, conversely, (iii) the mind depends on the brain, i.e., a necessary condition of any mental event is an event in the brain of a living body. An epistemological consequence is (iv) that our ways of acquiring factual knowledge are limited to sensory experiences, i.e., a physical event does not directly act on our mind, but only through some intermediate events in our sensory organs and finally in our brain. (Note that (ii) and (iii) sound dualist — Broad was sympathetic to epiphenomenalism — but may be reformulated so as to be compatible with monistic mind-body theories.) Since he maintained that the existence of the various parapsychological phenomena like telepathy and precognition was established beyond doubt, Broad concluded that these basic limiting principles of science are refuted.

   The fact that some of the research Broad referred to was later shown to be fraudulent [Ludwig, 1978; Kurtz, 1985; Hines, 2003], and that sophisticated parapsychologists try to conceive of telepathy and precognition in a somewhat different manner, so as to retain at least *prima facie* a naturalist interpretation [Duran, 1990], does not invalidate this as a useful example of the ontological problems faced by most parasciences. Indeed, many of their claims can only be upheld by giving up basic ontological convictions, which have so far proven to be extraordinarily fruitful for scientific research.

The most radical departure from the ontological paradigm of factual science is the open supernaturalism espoused by creationism. Inasmuch as creationism stipulates a *creatio ex nihilo*, it also violates Lucretius's principle. It is unclear whether or not many other parascientific claims can be accommodated within ontological naturalism. In any case, they still violate much of what we know about the lawful behavior of things. Homeopaths, for example, claim that high dilutions that no longer contain even a single molecule of the given substance still have a potent pharmacological effect. If what we know about chemistry is roughly true, there can be no such effect. Homeopaths have learned to concede this objection, but now forward the protective hypothesis that, in the mandatory process of shaking the dilutions (called "dynamization"), somehow the relevant "information" of the given substance gets transferred to the solvent. So what produces the therapeutic effect is this "information". It goes without saying that this supposed information is ill-defined and perhaps even immaterial, because water chemistry tells us that any molecular structure formed by H_2O-clusters is too short-lived to do any informational work. Moreover, if water (or alcohol or whatever fluid) had a memory, why would it specifically remember only the information of the homeopathic substance rather than that of all the other chemicals it had contained previously?

Another example is Therapeutic Touch. By moving her hands about 10cm over the patient's body, the healer attempts to adjust the patient's "vital energy", whose "imbalance" is always among the causes of whatever disease is to be healed. Needless to say, biology has abandoned any idea of vital energies long ago.

These examples show that many of the ideas occurring in the parasciences and paratechnologies are not necessarily supernatural in the traditional sense of involving powerful personal entities like gods or demons, but nevertheless *paranatural* [Kurtz, 2000], in the sense that they are not compatible with the naturalist-materialist outlook of the factual sciences. If we enrich this standard naturalism with more and more paranatural elements, it remains unclear, when this results in destroying it altogether.

The only ontological principle that is rarely violated by the parasciences is ontological realism. Even the weirdest entities occurring in the domain of the parasciences are deemed to exist really after all. The same holds for epistemological realism, which is why we proceed with a look at the methodological principles in the following subsection.

These examples illustrate that the parasciences not only suffer from the methodological problem of lacking evidential support, but also from their incompatibility with the major metaphysical background assumptions, which belong to the general hard core — the hard hard core, so to speak — common to any scientific approach. (For an analysis of the ontological presuppositions of esoterics see [Runggaldier, 1996].)

(b) *Methodology.* It is rather obvious that both Ockham's razor and fallibilism are widely neglected in the parasciences. Indeed, many parasciences populate the universe with (often occult) entities that are not needed for a scientific explanation of the world around us. Examples are the many life or other energies and forces postulated by quack medicine and pseudophysics. Dowsers believe that there are not only earth rays, but that these also occur in certain grids, which can be measured and mapped. And occultism teems with ghosts and spirits. There is no indication that the nature or the number of such entities is restricted by considerations of parsimony in hypothetico-deductive reasoning: their only restriction seems to be due to the limits of their authors' imaginative powers. This is not to say that they serve no explanatory function: they certainly do. The point is, as mentioned earlier, that almost any other arbitrary alternative or additional entity would do just as well.

As for fallibilism, it too is evident that most parascientists are not willing to seriously consider the possibility that they may be in error. If we extend Settle's [1971, p. 185] diagnosis of magic to the parasciences in general, we might say that many parasciences are explanatorily complete and thus come with the air of certainty, whereas factual science is explanatorily incomplete and thus accompanied by corrigibility. This difference helps to explain why the former are so much more appealing to many than the latter. Obviously, an explanatorily complete field has no need for research and hence for improvement, let alone revision (see Kuiper's [2001] definition of pseudoscience mentioned above). As we shall see later on again, some parascientific fields do allow for some limited improvement, such as parapsychology and astrology. However, these changes are not due to an internal tradition of fallibilism, but they are the result of massive external criticism by mainstream scientists.

(c) *Semantics.* As a truth definition the correspondence notion of truth, being simply a companion of ontological realism, is adopted in most parascientific fields. The major difference between science and parascience lies in the question of what is acceptable as truth indicators. Now this belongs in methodology, not semantics, so it may suffice here to add that, beside the main question of what can be regarded as legitimate objective evidence, the parasciences often accept as indicators of truth also subjective “evidence”, such as sheer belief or feeling, mystical vision, or other paranatural forms of experience.

(d) *Axiological and moral assumptions.* Different values manifest themselves in different behaviors of the individuals adopting these values. Thus, as mentioned in Section 5.1, Derksen [1993] has suggested analyzing the behavior and attitudes of the individual pseudoscientist, and Kitcher [1993] has recommended focusing on the psychology of the pseudoscientist. However enlightening this may be in some cases, in particular when taking a closer look at the founding father (or mother as the case may be) of some field, as Derksen did with Freud, it does not suffice to characterize the entire epistemic field. For example, it is possible for an individual to behave rationally within a magical belief system [Settle, 1971], whereas an individual scientist working in a rational tradition may on occasion behave irrationally. For this reason we better focus on the *institutional rationality*, or irrationality respectively, exhibited by the community C of some epistemic field, which is done best by examining the latter's general ethos or value system.

- *Logical values.* The canon of valid reasoning and thus the basic principles of rationality may be accepted officially, but they can be suspended whenever needed to save some claim. Lots of logical blunders occurring in the parasciences have been collected by various authors (see, e.g., [Schick and Vaughn, 1999; Wilson, 2000]). Since many of these occur in the context of justification, we shall give a sample in the subsection on methodological values.
- *Semantical values.* Meaning definiteness and clarity are rarely among the semantical values of the parasciences. Instead, vagueness and fuzziness are rampant, if not even seen as virtues by those cherishing the mysterious. We must also be prepared to encounter the meaningless, i.e., nonsense. (Note that scientists often are too quick in calling something nonsense, just because it is false. However, something that is false cannot be nonsense, because nonsense can be neither true nor false, for it has no semantic meaning in the first place.) Regrettably, since for most laymen many scientific theories are more or less incomprehensible, unintelligibility on the part of a parascientific theory may easily be mistaken for a sign of an authentic science.
- *Methodological values.* Many parasciences are characterized by methodological values and hence procedures of their own. These consist, for example, in certain rules of inference or rules of evaluating evidence which are quite often regarded as fallacious by philosophers of science. For this reason they either have been eliminated from science, or, if they occasionally reappear in some reasoning, are quickly detected and denounced as mistakes by the scientific community. Indeed, fallacious methods were described already by

19th century philosophers of science like Mill and Peirce, and many modern authors who attempted to demarcate pseudoscience by its peculiar inferential methods, have collected various fallacies as indicators of pseudoscientificity (e.g., [Radner and Radner, 1982; Giere, 1984; Thagard, 1988; Schick and Vaughn, 1999; Wilson, 2000]). Since these fallacies do constitute important parascience indicators, a quick sample will be in order.

The a priori method: Accept only those beliefs that are such that it is impossible to imagine that the contrary is true [Wilson, 2000]. In other words, a hypothesis is accepted and considered worthy of use for explanation not on the basis of empirical evidence, but because its proponents regard alternatives as inconceivable. Examples: von Däniken keeps repeating that he simply cannot imagine how some artifact could have been produced by ancient man without extraterrestrial help. The creationists (including the more recent branch of Intelligent Design) keep repeating that it is inconceivable how the natural process of evolution could have produced certain complex organs without divine design or even intervention.

The fallacy of competition: This is the claim that some parascientific theory should be admitted because it might become an alternative theory in the future. Yet, as Radner and Radner [1982] point out, competition is only among current alternatives: by referring to some unknown future science, one actually refuses to compete. Their very apt analogy is the attempt to participate in a marathon on roller skates, arguing that the marathon might be changed to a skating race in the future.

Simplistic elimination [Giere, 1984; Wilson, 2000]: Assuming there are two rival theories A and B, and they are the only possible alternatives, we may infer that A is true if B is false. Yet in reality there usually are many possible alternative theories that might explain the same fact. So if we are faced with two or more alternative theories, we must first make sure that they really are the only alternatives, and that they are not false all together. Thus, many supposed eliminations are fallacious, because they do not consider all possible alternatives. The creationists argue, for instance, that there are only two alternatives: evolutionary theory and the theory of divine creation. But if evolutionary theory, including all we know about the history of the universe, is false, then divine creation is not the only remaining alternative: it may well be then that life is coeval with an uncreated eternal universe. Ufologists argue that, since some strange sightings cannot be explained by the usual candidates such as satellites, balloons, aircraft, or bright planets, they must be due to extraterrestrial visitors. Yet there may also be unknown natural atmospheric processes causing a given UFO-sighting.

Anything-goes method [Wilson, 2000]: This is the argument that, since even a well-confirmed theory might possibly be false, we should not dismiss alternatives to it. So everything goes. If this were correct, the corollary would be that in fact nothing goes, because these supposed alternatives might likewise be false.

Method of authority [Wilson, 2000]: As pointed out earlier, many parasciences are belief systems rather than research fields. It comes as no surprise therefore that a rule "to accept as true what the relevant authority tells you" is wide-spread. Naturally, this holds in particular for religious or quasi-religious fields such as creationism, scientology, anthroposophy, or transcendental meditation.

Resemblance thinking [Thagard, 1988; Wilson, 2000]: This is the habit, already pointed out by John S. Mill, of inferring from the observation that *A resembles B*, that therefore *A causes B*. Prime examples of fields relying heavily on resemblance thinking are astrology and homeopathy. The latter's "law of similars", stating that like heals like, is even enshrined in the very name "homeo-pathy" (from the Greek *homoios*, similar).

The grab-bag approach to evidence [Radner and Radner, 1982]; see also the *blunderbuss argument* in [Wilson, 2000]): In evaluating the evidential support for some theory, we should not just look at the quantity of confirming instances, but first of all at their quality. Thus, we do not have to keep shooting canon balls in order to confirm the laws of motion. Of particular value, on the other hand, are data that were gathered after a theory had been proposed, and that were possibly even predicted by the theory; likewise with evidence that was produced under a variety of different conditions. Classical examples with regard to Newton's theory are the discovery of Uranus and Neptune, and the prediction of the return of Halley's comet. By contrast, it is typical of many parasciences that the sheer quantity of "evidence" makes up for the lack of quality of the individual data. For example, von Däniken pulls out artifact after artifact in favor of his "alien hypothesis"; the creationists keep listing complex biotic structures which impossibly could have come into existence naturally, i.e., by evolution; and the ufologists will report strange sighting after sighting. Moreover, as soon as one piece of such evidence has been rejected, either for being fallacious or forged, or for having been explained within a standard scientific context, the parascientist will simply continue to pull out data of the same kind and quality from his evidential grab bag, thereby keeping the skeptic busy for all times. Worse, the fact that scientists cannot always readily refute each and every item pulled out of the grab bag, is taken as a further reason for belief in the parascientific tenet in question.

- *Attitudinal values.* The attitudinal value system of the parasciences is as varied as the parasciences themselves. Thus, again, there are no universal features characterizing all the parasciences. Nonetheless, an attitudinal profile of parascience may include the following aspects. Parascientists pretend to be critical thinkers, but their canon of critical thinking is not the same as that of science and philosophy. In fact, many are just believers, not investigators. They also claim to be open-minded, but their open-mindedness does not extend to the possibility that the standard scientific view of nature is the correct one. Instead, it includes sympathy for the most outlandish claims, because to the parascientist open-mindedness often means "anything nonscientific goes", so that it amounts to blank-mindedness. Universalism and objectivity are not values in those fields dominated by authorities, or in which only the initiate has special access to the truth. Think of the various branches of occultism.

5. *Formal background F.* Concerning the formal background of any suspected parascience, we may ask questions such as the following: Are there any mathematical models? Is the mathematics in these models handled correctly? This is often not the case. In particular, in some pseudophysics such as the attempts at refuting the theory of relativity, the mathematics is defective, if not phoney. The same occurs in some social sciences, in particular sociology and economics, where pseudoquantitation may go unnoticed [Sorokin, 1956; Blatt, 1983; Bunge, 1999, Ch. 4]. The latter example illustrates once more that some research fields which on the whole are regarded as scientific may nonetheless exhibit some occasional pseudoscientific feature (Fig. 4).

6. *Specific background knowledge B.* In contrast to scientific fields, which borrow amply from adjacent disciplines, the parasciences are typically isolated enterprises. They presuppose some ordinary knowledge, and of course they borrow some science when needed. But note that the function of the scientific knowledge borrowed consists mostly in justifying the scientific pretensions of the given pseudoscience: it is easier to imitate science when you also use some well-accepted scientific knowledge. The scientific input is often not needed to advance the own field. Note also that the converse input does not obtain: scientific fields have hardly any use for knowledge produced in a parascientific field.

   Astrology, for example, accepts of course some basic astronomic facts, but disregards many others, in particular those that refute its own claims. Creationists rely heavily on biological knowledge, but only to prove the falsity of evolutionary theory. However, no scientific knowledge whatsoever can shed any light on the totally occult mechanism of divine creation. In other words, no scientific knowledge can advance creationist "theory".

The theory probably most often borrowed from the sciences is quantum theory, which has become an explanatory panacea for many parasciences, from New Age physics through parapsychology to holistic medicine [Grove, 1985; Stalker and Glymour, 1989]. For example, sophisticated parapsychologists have long abandoned stories of moving tables and telepathically communicating people. The naturalistically oriented part of current parapsychology claims that paranormal effects are microeffects rather than macroeffects, and that they can be accounted for by quantum theory. Telepathy, for instance, is no longer seen as a form of human communication, but at most as an instance of nonlocal correlations between some quantum events in two peoples' brains, or between a person's brain and some other object like a random number generator. It will come as no surprise then that the use of quantum theory in the parasciences often involves a serious distortion, in particular a return to long abandoned subjectivist interpretations. Moreover, one often uses the vocabulary of quantum theory but rarely its concepts [Stenger, 1995; Spector, 1999]. In sum, the motto is: if you don't know what it is and how it works, call in quantum theory to describe and explain it.

Note, incidentally, that in sophisticated parapsychology this move is due to the attempt to stay within the bounds of a naturalist ontology. At the same time, it presupposes a radically reductionist view, because it disregards the level structure of the world, i.e., the fact that macroobjects such as neural assemblages have systemic properties, so that their behavior is usually not influenced by microevents occurring at the quantum level. For example, neuroscientists know that mental processes, such as perception and thinking in general, involve millions, if not billions, of complexly interacting neurons and their coordinated activities at different organizational levels. The idea that quantum events occurring at the level of elementary particles or at most atoms should be able to influence these highly complex neuronal systems in a coordinated manner is extremely implausible [Beyerstein, 1987; Humphrey, 1999; Kirkland, 2000].

Parasciences sometimes also borrow ideas from other parasciences. A prime example is Carl G. Jung's concept of synchronicity, which is made use of both in sophisticated astrology and parapsychology. This is the idea that two events which have no causal connection are nonetheless "meaningfully" related (McGowan 1994; Carroll 2003; Hines 2003). Thus, if the quantum physical notion of nonlocal correlation cannot be called in as an ad hoc device to establish a connection between two (simultaneous) events, because what we have is just a coincidence, synchronicity will do the trick. For example, sophisticated astrologers have learned from the many scientific objections hurled at them during the past centuries: they nowadays admit frankly that the relation between humans and the various constellations of stars and planets is not a causal one. What saves the business though is the claim that the relation between the stars and humans is nonetheless a meaningful one, namely an instance of synchronicity. This neo-astrology then finds and interprets these meanings and explains them to its customers, turning the field into a form of astro-counseling. Note that this strategy is clearly ad hoc: it is not due to internal progress in astrology but a move to avert external criticism, making astrology immune against the standard astronomic objections without having to give up the "astro" in "astrology".

7. *Problematics P.* In the parasciences the collection of problems is usually small and mostly practical, for many parasciences are actually paratechnologies or paratechnics. Important questions about any parascience candidate are: Does it solve or help to solve problems other than its own? Do its problems arise from natural contexts, or are they artificial (fabricated)? Three examples might illustrate this problem concerning parascientific problems.

   Astrology mostly solves problems that would not exist without astrology in the first place. The only general and natural question that astrology tries to answer, namely the question why different people have different characters, is better answered by genetics, developmental psychology, and sociology. Moreover, the astrological answer is incompatible with the scientific one and thus does not enrich scientific knowledge. For the most part, however, astrology is a pseudotechnology, which has rules to apply, but no puzzles to solve [Kuhn, 1970]. In particular, the many failures of astrological predictions do not entice any problem-solving activity in the astrological community.

   The problems of von Däniken's pseudoarcheology too are fabricated rather than natural, because he preys on the natural problems of normal archeology and turns them into mysteries, which he claims can only be solved by his hypothesis about extraterrestrial visitors. Thus, von Däniken's hypothesis does not yield any new problems on its own: it is entirely parasitic on the pre-existing problems in other fields.

   Parapsychology started out with the natural problem of unusual human experiences, in particular at a time when spiritualism was *en vogue*. Some people sometimes do have anomalous (though nonpathological) experiences. The basic question therefore is whether all such anomalous experiences can be explained naturally (i.e., within the normal paradigm of scientific biopsychosociology), or whether we do need to enrich this paradigm with paranormal entities and processes to account for these unusual experiences. Yet, the more successful the normal sciences, including in more recent times the neurosciences, became in explaining anomalous experiences, the less needed were explanations referring to paranormal entities or processes. In this way, parapsychology practically lost its source of spontaneous or natural problems, although people keep experiencing unusual things. Not willing to give up the psi hypothesis in favor of the null hypothesis, parapsychologists started to fabricate new problems: they began studying arbitrary correlations between human subjects and virtually every possible other object, desperately looking for statistically significant deviations from chance expectation (i.e., anomalies), which can then be interpreted as evidence for psi. Since all the results from such — often quite sophisticated — studies are, if not negative, at best inconclusive, the consequence is the perpetual call for further research. Thus, parapsychology generates arbitrary problems of the sort "Could there be an anomalous correlation between x and y_1 or y_2 or $\ldots y_n$?" in order to keep itself alive. As Alcock [2003, p. 34] observes, the anomalies parapsychologists search for have never popped up in normal research. Thus, again, the contemporary problems of (sophisticated) parapsychology would not exist if it were not for the existence of parapsychology itself.

This may be the place to take a brief look at the role of anomalies in science and parascience. Normal scientists do not look for anomalies, they "hit them in the face" [Radner and Radner, 1982, p. 33; Alcock, 2003]. Indeed, every scientist who performs some measurement or experiment has certain expectations as to its outcome, in particular if the outcome is predicted by some theory. If the resulting data seriously deviate from these expectations, they constitute an anomaly. Although it takes more than just a few anomalies to initiate a scientific revolution, the importance of anomalies for theory change and hence scientific progress has been well known and discussed ever since the work of Kuhn [1962]. However, scientists are conservative in the sense that they will not give up an otherwise well-confirmed theory, let alone an entire research program, in favor of some alternative theory whose only merit is its ability to explain a certain anomaly. On top of explaining the given anomaly, the new rival theory must at least explain as much as does the standard theory.

By contrast, parascientists rejoice when they find anomalies. Their expectations are not those of an orderly and lawful world, but of a world teeming with mysteries. Therefore, they actively search for anomalies, which they can then turn into problems to be solved by their respective "alternative" theories. And these alternative theories are expected to revolutionize science. In so hoping, parascientists forget that no scientific revolution has ever been triggered from without. Nonetheless, there is even a field or rather a multi-field called *anomalistics*, which is exclusively devoted to the study of anomalies supposedly neglected by mainstream science. The main player in this field is the *Society of Scientific Exploration.*

8. *Fund of knowledge K.* The fund of knowledge of a parascience is not a growing collection of up-to-date and well-confirmed data and theories: it is usually small, it stagnates, it contains statements that are incompatible with well-confirmed scientific knowledge, and its hypotheses lack evidential support. For this reason, the knowledge in these fields is purely speculative and cannot be said to even approximate the truth, i.e., to roughly represent any real facts.

   A frequent feature of parascientific knowledge is its anachronistic character [Radner and Radner, 1982]. What many parascientists propagate as revolutionary new insights or at least as rival "scientific" theories is in fact very old news, so old indeed that they have long been discarded by science. For example, alternative medicine teems with mysterious vital energies that supposedly are out of balance when we are sick. Thus, the basic ideas of homeopathy only make sense when we go back 200 years when vitalism was still going strong in biology and medicine. Traditional Chinese medicine presupposes the existence of some vital energy (*qi* or *ch'i*), flowing in channels (meridians) unknown to biology. And the practitioners of therapeutic touch and reiki (*ki* is the Japanese equivalent of *qi*) claim that they treat the imbalances in the "human energy field", whereas the so-called prana healers refer to the Hinduist equivalent *prana*. The creationists still defend views that may have been legitimate 200 years ago. Then there are the pseudophysicists who still try to build perpetua mobilia or other so-called free energy machines as though thermodynamics were nonexistent, or who desperately strive to refute Einstein's two relativities in order to re-establish good old Newtonianism. Finally, astrology is another prime example of a world view that has been superseded for several hundred years.

9. *Aims A*. The aims of the parasciences are sometimes cognitive, but for the most part practical. That is, many parasciences are paratechnics or paratechnologies, such as astrology and alternative medicine. Yet even when the aims appear to be cognitive, the ultimate goal of many parasciences is often anthropocentric and quasi-religious (Alcock 1985), if not explicitly religious as in the case of creationism. *Prima facie* the goals of the creationists, such as the establishment of an alternative cosmology and history, appear to be cognitive rather than practical. But we may suspect that the ultimate goal is in fact personal salvation, which, in the fundamentalist world view, can only be achieved by a consistent way of life according to biblical literalism. Similarly, the spiritualist approach of esoterics wants to establish the multifarious spiritual connections of humans with the rest of the world. Often the ultimate goal is quite explicitly stated: the materialist world view of science is to be replaced with a spiritualist one. For example, one of the main figures in 20th century parapsychology, Joseph Banks Rhine, asserted that "little of the entire value system under which human society has developed would survive the establishment of a thoroughgoing philosophy of physicalism" (Rhine [1954/1978, p. 126]). This exemplifies how the aims of both science and parascience often depend on — conflicting — metaphysical outlooks.

10. *Methodics M*. The empirical methods used in the parasciences often are just as occult as the theoretical background assumptions. For example, an instrumental technique such as a pendulum used to diagnose some disease, presupposes some occult mechanism mediating between the healer and, say, the patient's "life energy". How can this method be checked? Interestingly, it can partly be checked scientifically, but it cannot be checked within the own theoretical system of the given field. In other words, in can partly be

tested externally, but not internally. For example, in a double-blind setup, someone claiming to be able to diagnose some specific disease by simply holding a pendulum over a photo of a patient, is given 25 photos of healthy persons and 25 photos of persons suffering from the given disease (i.e., neither the healer nor the experimenter knows which of the photos belongs in which group, and it is impossible to diagnose the given disease from merely looking at the peoples' faces on the photos). As yet, all experiments of such a kind have had negative results, i.e., the candidate's success rate has never been significantly above chance expectation.

Now this is of course a basic and objective scientific test which only checks whether or not the given technique works (not how it works if it did work in the first place). And it was imposed from the outside, because it does not belong to the methodics of the given parascience. So how can the functioning of the method be checked internally? Unsurprisingly, the healer herself might claim that she is able to check her diagnostic technique with alternative means. She may, for instance, use a dowsing rod, or perhaps just put her hand on the picture. In her normal environment all this will most likely combine with confirmation bias and subjective validation into the belief that her method is successful and reliable. However, as a matter of fact even within the own outlook of such a parascientific approach, the given method cannot be checked by other persons in the field, because her colleagues will not be able to reproduce her diagnosis. Indeed, every other person claiming the same ability will very likely come up with a different diagnosis, provided of course she does not know the earlier diagnosis of her colleague. There may be some overlap in the results due to chance, but by and large the success rate will not differ from mere guessing. In short, many techniques used in the parasciences are not objective in the sense that everyone applying the method will get the same results. This holds *a fortiori* for openly subjective methods like spiritual means of communication or mystical vision. The latter are not even methods in the sense of rule-guided procedures.

By contrast, in their attempt to imitate science, the pseudosciences often do use scientific methods. For example, the statistical methods used in sophisticated parapsychology are sometimes impeccable. Moreover, often even the general scientific method is followed as is obvious from the parapsychological journals. In so doing, many pseudosciences, in particular parapsychology and astrology, often exhibit a naive empiricist view of science: they believe that the application of scientific methods and techniques, including the scientific method as defined above, is sufficient to warrant the scientific status of their field. Indeed, in particular parapsychologists have learned a lot from their critics and have thus improved both their statistical sophistication and the precautions against fraud and self-deception. (Note again that these improvements are largely due to external pressure, not internal progress.) So they believe that what they do is proper science, and they reject the various methodological and other philosophical objections as sheer ideological dogmatism, failing to realize that conceptual criticism is part and parcel of science too.

11. *Systemicity.* The systemicity condition is one of the stronger indicators of parascientificity (recall Reisch's criterion of network demarcation mentioned in Sect. 3). Indeed, parasciences are isolated fields. They do not form a consilient system of knowledge; in particular, they make no contact with normal science. It is precisely because parascientific knowledge must be rejected as unfounded that it cannot enrich scientific knowledge. Moreover, parascientific knowledge often collides head-on with scientific knowledge: if parascientific theories were true, their scientific alternatives including those theories to which they are connected would be false. Thus, many parascientific theories would cause total or global revolutions: the entire edifice of scientific knowledge including the scientific paradigm as a whole would collapse. By contrast, contemporary scientific revolutions, if any, will only be local or regional revolutions, because too many things we have come to know during the past 400 years are reliable and must therefore be at least approximately true. Examples of fields calling for global revolutions are creationism and parapsychology. As for the latter, recall C. D. Broad's basic limiting principles, which underlie all modern science.

12. *Progressiveness.* According to the criterion of progressiveness, the membership of the conditions 5–10 changes, however slowly and meanderingly at times, *as a result of research* in the same field or as a result of research in neighboring disciplines. Obviously, many parascientific fields are plainly stagnant, which can be detected rather easily. This is due to the fact that many of them are not really research fields but instead belief systems.

    But of course, there are also some parasciences in which there is at least some minor change, and there are others which are actually research-oriented, such as parapsychology. Indeed, as mentioned before, research keeps parapsychology busy. However, despite its age of more than 120 years, it has not come up with a single conclusive finding [Kurtz, 1985; Hyman, 1989; Alcock, 2003]. Thus after 120 years it is still a field in search for its domain, and it desperately tries to gather hard data. Nonetheless, it has even produced some theories to explain certain supposedly paranormal events or experiences, respectively. It has also introduced plenty of ad hoc hypotheses to protect itself from criticism. An example is the idea of psi missing. If some experiment yields a score slightly above chance expectation, this is of course regarded as evidence for psi. Likewise, if some trial yields a below chance result, this too is seen as evidence for psi: in this case the subject's psi abilities somehow operate to avoid the target (psi missing). In this way any fluctuation around the exact chance expectation becomes evidence for psi. Given this situation, it seems that parapsychology is able to generate the appearance of progress, although a closer look reveals that this progress is just as illusory as the very domain of parapsychology. After all, can there be genuine progress when the given field does not even have a real domain?

5.4 Conclusion

We have now listed and examined a number of features characterizing parascientific fields. The features used in this characterization are of course of unequal weight: some are more decisive than others, so that their presence is a stronger indicator of a field's status. For example, a violation of some of the basic limiting principles in G carries more weight than some methodical flaw in M, which may be repairable more easily, provided the practitioners of the field care to. Since the above features are not jointly necessary and sufficient conditions, another open question is how many of these characteristics must at a minimum be present for a field to be parascientific. Insofar as such a condition is a necessary one, such as the logical requirement of noncontradiction, we may reject the given field as irrational on this one count. In most cases, however, a simple characterization of a parascience such as "it's all a matter of X", where X may stand for falsifiability, method, or attitude will not do. Indeed, we ought to be more careful and always attempt to prepare a comprehensive profile of the suspected field. Such a profile should allow us to come to a well-reasoned conclusion as to the scientific or parascientific status of the given field, although every such conclusion will differ in the reasons used as its premises.

The preceding analysis focused on epistemic fields as the central units of demarcation. However, a comprehensive profile of some parascientific epistemic field should also allow us to diagnose smaller units as parascientific, if they are the bearers of one or more characteristic features occurring in the profile. Such smaller units may be theories as systems of statements, which may be inconsistent or circular, or incompatible with the accepted background knowledge; individual hypotheses, which may be logically unfalsifiable; individual methods, which may have long been weeded out from the sciences for being defective; or some behavior or attitude of the representatives of the field, and so on. In this way we are justified in calling a theory, a hypothesis, a method, or a behavior unscientific. This is of particular importance when we are dealing with an epistemic field which we normally regard as scientific. For in such a case the philosopher of science may still detect some unscientific feature and denounce it as being pseudoscientific, calling for its repair or, if impossible, its elimination.

6 PROTOSCIENCE AND HETERODOXY

Calling some theory, approach or entire epistemic field parascientific is a strong and damning verdict. For this reason we must be quite careful in our judgment, which ought to be based on a diligent examination of the suspected theory or field. Now, whereas the philosopher of science may be more careful in such pursuit, scientists are sometimes less careful. Thus, many authors have warned us that the history of science should teach us sobering and humbling lessons concerning the science/pseudoscience demarcation (e.g., Toulmin 1984). First, it has always been too easy a temptation to reject a theory or approach as pseudoscientific just because it is heterodox, or maybe just because we do not like or understand it. Second, some theories that are declared pseudoscientific may actually turn out to be protoscientific, so that their possibly bright future could be endangered by an unfair judgment. Third, there is the historical problem of judging a certain field in retrospect: some field that may be clearly pseudoscientific today, may have been protoscientific at an earlier time and hence in a different scientific landscape.

A prime example is Alfred Wegener's hypothesis of contintental drift, which was initially rejected when proposed in 1915 and sometimes even derided, but eventually became the basis for the plate tectonics revolution in the 1960ies. Wegener's ideas were indeed protoscientific rather than pseudoscientific because he did not refer to untestable myths and mysteries like Velikovsky or von Däniken, but instead to geological and climatological data. And he did not behave like a pseudoscientist, for he admitted that his ideas were conjectural and that the main problem of his hypothesis was the unknown mechanism of continental drift. However, his geological colleagues also acted rationally in rejecting his hypothesis for being too implausible at that time (see [Kitcher, 1982; Radner and Radner, 1982]). Apart from the historical vindication of Wegener's protoscientific ideas, an assessment of Wegener's hypothesis in a pseudoscience profile would most likely have shown that even at their time his views were not pseudoscientific, but merely unorthodox [Edelman, 1988]. This indicates that it is not always true that we can determine the scientific status of a certain theory or field only retrospectively, e.g., by observing its historical progress or else degeneration.

A less favorable example is phrenology, which has been regarded as a protoscience leading to neuropsychology (Young 1970). Phrenology advanced the correct and fruitful idea that mental functions are localized in the brain, but was badly mistaken in the claim that these functions manifest themselves craniologically, i.e., as bulges on the skull. The latter made phrenological diagnosis a pseudotechnology, which, however, had some beneficial side-effects on the treatment of prisoners and the mentally ill [Hines, 2003]. In this case a retrospective analysis shows that a small part of phrenology led to progress, if only in a field that quickly emancipated itself from phrenology, whereas the larger part degenerated into a pseudoscience.

In the case of astrology opinions are divided. Apart from its defenders of course, even some philosophers of science are willing to grant astrology the status of a former protoscience (e.g., [Thagard, 1978]). Others maintain that astrology never was a protoscience, because even in antiquity educated people, like Strabo, Cicero and Ptolemy, clearly distinguished between astronomy and astrology, whether or not they believed in the latter [Culver and Ianna, 1988]. Moreover, it was obvious to many even back then that astrological predictions are unreliable for failing too frequently. And although some early scientists like Kepler practiced some astrology, they too kept it apart from science. Thus, it seems that despite various connections and flirtations between early astronomers and astrologers, astrology has long, if not always, been para- or even pseudoscientific, contributing nothing to astronomy or any other science.

These historical examples illustrate the need for a comprehensive analysis of any field or theory suspected of being a parascience. Even if we were wrong with our judgment at a given time, a genuine protoscience will sooner or later prove its fruitfulness and potential by developing into a full-fledged scientific field, propelled by successful research, or at least by giving rise to some scientific field.

But what exactly does "sooner or later" mean? We must ask this question because one of the most intriguing and sophisticated pseudosciences, namely parapsychology, has always claimed that it is actually a protoscience (or a pre-paradigmatic science, as some parapsychologists prefer to call it in Kuhnian terms), so that its classification as a pseudoscience would be unjustified. Now the birth of parapsychology as a field of research is usually taken to coincide with the establishment of the *Society for Psychical Research* in 1882, although earlier research in the area of spiritualism dates back to the 1850s [Kurtz, 1985]. Should a field still be regarded as a protoscience after more than 120–150 years? As mentioned several times in this chapter, parapsychology is a field still in search for a proper domain, because it has not succeeded in producing any findings that would convince its critics from mainstream psychology of the existence of some paranormal entities or processes [Hyman, 1989; Hines, 2003]. Worse, as Alcock [2003, p. 32] summarizes the situation: "...to the extent that parapsychology constitutes a 'field' of research, it is a field without a core knowledge base, a core set of constructs, a core set of methodologies, and a core set of accepted and demonstrable phenomena...". Does this not rather indicate that there is no such thing as psi (in other words, that the null hypothesis is true) and that the field is degenerative rather than protoscientific?

The same holds for astrology and creationism, which have also learned to exploit the "humbling lessons of history", claiming to be actually protosciences, which deserve to be granted their due chance of proving themselves full-fledged sciences. Yet if we are suspicious of a 120-150 years old protoscience, we are entitled to be even more skeptical of alleged protosciences that are thousands of years old.

A comprehensive profile of the epistemic field under consideration should also help to solve the problem of how to distinguish fruitful scientific heterodoxy from pseudoscientific deviation. In his foreword to the book "Scientists Confront Velikovsky" [Goldsmith, 1977], the famous science fiction author Isaac Asimov has coined the terms *endoheresy* and *exoheresy.* These terms capture nicely the gist of Section 5.3, namely the condition that a heresy must stay within the bounds of the scientific superparadigm, so to speak, if it is to be considered legitimate, even though the majority of the scientific community may reject it as mistaken or misguided. For example, in developmental biology there is a school called "developmental structuralism" [Webster and Goodwin, 1996], which takes genes to be relatively irrelevant for development, and hence seeks to explore the role of "universal laws of form" or "transformation laws" in development. Thus, it is attempted to describe the developing organism by field equations, reviving the earlier notion of a morphogenetic field. This structuralist approach is rejected or ignored by most developmental biologists, but it stays within the bounds of science, although some of the philosophical considerations of these authors seem to be in need of repair [Mahner and Bunge, 1997]. By contrast, the morphogenetic field hypothesis of the former biochemist Rupert Sheldrake is clearly an exoheresy, for it shows too many marks of pseudoscience and is irreparably esoteric [Carroll, 2003].

The preceding considerations result in the recommendation that both the sciences and the humanities ought to welcome endoheresies, because they form a valuable stock of alternative views, however implausible they may be at a given time. After all, it is too easy to be blinkered by orthodoxy which is reinforced by the routine of normal research. On the other hand, scientists must judge for themselves whether they wish to spend any time on investigating exoheresies. However, if not for scientific reasons, they should on occasion study exoheresies for educational purposes, explaining to the public why certain claims are parascientific and hence unworthy of serious attention. Although scientists may have very good reasons for rejecting exoheresies, they must keep explaining these reasons to the public in order to avoid the impression that their refusal to pay attention to parasciences is due to sheer dogmatism and arrogance. Thus, the advancement of the public understanding of science requires that we deal not only with science, but also with parascience.

7 CONCLUSION

Looking at the figures 2, 3 and 4, we notice that there are two main demarcation lines: the one between science and nonscience, and the other between reliable (approximately true) and illusory knowledge. Now some authors maintain that it is the latter which is the more important one (e.g., [Laudan, 1983; Haack, 2003]). After all, proper inquiry and proper standards of reasoning and evidence exist also outside science. For example, not only the philosopher arguing his case, but also the policeman investigating a crime knows (or at least should know) how to reason properly and how to distinguish good from bad evidence. As a consequence science would not differ in kind from other epistemic areas where common standards of rational and objective inquiry are practiced, but at most in the degree and thoroughness of their application [Haack, 2003]. Since determining when knowledge is gained in a proper way is the task of epistemology in general, it seems that the basic epistemological demarcation between knowledge and illusion is more important than that between science and nonscience.

This view usually rests on the idea that science is but an extended form of common sense, as both scientists like Thomas Huxley and Albert Einstein, and philosophers like John Dewey and Gustav Bergmann believed [Haack, 2003]. But unless the common sense of philosophers is totally different from everybody else's, this view is doubtful: there are good arguments for the contrary thesis that, in important respects, science transcends common sense and ordinary language, and therefore is quite “unnatural” [Wolpert, 1992]. The fact that so many people have serious difficulty in understanding scientific concepts, theories, and methods renders noncommonsensism more plausible than commonsensism. Yet even if scientific thinking were just extended common sense, it would still be the task of the philosopher of science to tell us how scientific cognition and knowledge differ from nonscientific cognition and knowledge.

In any case, wherever we eventually draw our lines, the important thing is to draw some line at all, so as not to surrender to relativism, arbitrariness, and irrationalism.

ACKNOWLEDGEMENTS

I am indebted to Mario Bunge, Michael Kary, Amardeo Sarma, Thomas Waschke, and Theo Kuipers for valuable suggestions and comments.

BIBLIOGRAPHY

[Agassi, 1964/1999] J. Agassi. The Nature of Scientific Problems and Their Roots in Metaphysics. In M. Bunge (ed.) *Critical Approaches to Science and Philosophy*, pp. 189–211. Transaction Publishers: New Brunswick, NJ 1999.

[Agazzi and Darvas, 1997] E. Agazzi and G. Darvas, eds. *Philosophy of Mathematics Today*. Kluwer: Dordrecht, 1997.

[Albert, 1985] H. Albert. *Treatise on Critical Reason*. Princeton University Press. Princeton, NJ, 1985.

[Alcock, 1985] J. Alcock. Parapsychology as a “Spiritual” Science. In: P. Kurtz (ed.), pp. 537–565, 1985.

[Alcock, 2003] J. Alcock. Give the Null Hypothesis a Chance. Reasons to Remain Doubtful about the Existence of Psi. In: J. Alcock, J. Burns, A. Freeman (eds.) *The Psi Wars: Getting to Grips with the Paranormal*, pp. 29–50. Imprint Academic: Exeter, 2003.

[Alters, 1997] B. J. Alters. Whose Nature of Science? *Journal of Research in Science Teaching* 34: 39–55, 1997.

[Armstrong, 1995] D. M. Armstrong. Naturalism, Materialism, and First Philosophy. In: P.K. Moser & J.D. Trout (eds.) *Contemporary Materialism*, pp. 35–50. Routledge: London, 1995.

[Ayer, 1946] A. J. Ayer. *Language, Truth and Logic*. Dover: New York, 1946.

[Bechtel, 1986] W. Bechtel, ed. *Integrating Scientific Disciplines*. Martinus Nijhoff: Dordrecht, 1986.

[Beckner, 1959] M. Beckner. *The Biological Way of Thought*. Columbia University Press: New York, 1959.

[Beyerstein, 1987] B. L. Beyerstein. Neuroscience and psi-ence. *Behavioral and Brain Sciences* 10: 571–572, 1987.

[Blatt, 1983] J. M. Blatt. How Economists Misuse Mathematics. In: A.S. Eichner (ed.) *Why Economics is not yet a Science*, pp. 166–186. M.E. Sharpe: Armonk, NY, 1983.

[Boyd, 1999] R. Boyd. Homeostasis, Species, and Higher Taxa. In: R.A. Wilson (ed.) *Species: New Interdisciplinary Essays*, pp. 141–185. MIT-Press: Cambridge, MA, 1999.

[Broad, 1949] C. D. Broad. The Relevance of Psychical Research to Philosophy. *Philosophy* 24: 291–309, 1949. [Also in J. Ludwig, ed. 1978]

[Bunge, 1982] M. Bunge. Demarcating Science from Pseudoscience. *Fundamenta scientiae* 3: 369–388, 1982.

[Bunge, 1983a] M. Bunge. *Treatise on Basic Philosophy, vol. 5: Epistemology and Methodology I: Exploring the World*. D. Reidel: Dordrecht, 1983.

[Bunge, 1983b] M. Bunge. *Treatise on Basic Philosophy, vol. 6: Epistemology and Methodology II: Understanding the World*. D. Reidel: Dordrecht, 1983.

[Bunge, 1984] M. Bunge. What Is Pseudoscience? *Skeptical Inquirer* 9(1): 36–46, 1984.
[Bunge, 1999] M. Bunge. *The Sociology-Philosophy Connection.* Transaction Publishers: New Brunswick, NJ, 1999.
[Carnap, 1936–37] R. Carnap. Testability and Meaning. *Philosophy of Science* 3: 419–471, 4: 1–40, 1936–37.
[Carroll, 2003] R. T. Carroll. *The Skeptic's Dictionary.* Wiley: New York, 2003. [Also online: www.skepdic.com]
[Cartwright, 1983] N. Cartwright. *How the Laws of Physics Lie.* Clarendon Press: Oxford, 1983.
[Culver and Ianna, 1988] R. B. Culver and P. A. Ianna. *Astrology: True or False? A Scientific Evaluation.* Prometheus Books: Buffalo, NY, 1988.
[Curd and Cover, 1998] M. Curd and J. A. Cover, eds. *Philosophy of Science. The Central Issues.* W.W. Norton: New York, 1998.
[Darden and Maull, 1977] L. Darden and N. Maull. Interfield Theories. *Philosophy of Science* 44: 43–64, 1977.
[Derksen, 1993] A. A. Derksen. The Seven Sins of Pseudo-Science. *Journal for General Philosophy of Science* 24: 17–42, 1993.
[Derksen, 2001] A. A. Derksen. The Seven Strategies of the Sophisticated Pseudo-Scientist: A Look into Freud's Rhetorical Toolbox. *Journal for General Philosophy of Science* 32: 329–350, 2001.
[Devitt, 1996] M. Devitt. *Realism and Truth.* Princeton University Press: Princeton, NJ.
[Dupré, 1993] J. Dupré. *The Disorder of Things. Metaphysical Foundations of the Disunity of Science.* Harvard University Press: Cambridge, MA, 1993.
[Duran, 1990] J. Duran. Philosophical Difficulties with Paranormal Knowledge Claims. In P. Grim (ed.), pp. 232–242, 1990.
[Edelman, 1988] N. Edelman. Wegener and Pseudoscience: Some Misconceptions. *Skeptical Inquirer* 12(4): 398–402, 1988.
[Edwards, 2004] P. Edwards. *Heidegger's Confusions.* Prometheus Books: Amherst, NY, 2004.
[Eflin *et al.*, 1999] J. T. Eflin, S. Glennan, and G. Reisch. The Nature of Science: A Perspective from the Philosophy of Science. *Journal of Research in Science Teaching* 36: 107–116, 1999.
[Ernø–Kjølhede, 2000] E. Ernø–Kjølhede. *Scientific Norms as (Dis)Integrators of Scientists?* MPP Working Paper 14, pp. 1–18. Copenhagen Business School: Copenhagen, 2000.
[Ernst *et al.*, 2001] E. Ernst, *et al.*, eds. *The Desktop Guide to Complementary and Alternative Medicine.* Mosby: Edinburgh, 2001.
[Feyerabend, 1975] P. Feyerabend. *Against Method.* New Left Books: London, 1975.
[Flew, 1990] A. Flew. Parapsychology: Science or Pseudoscience? In P. Grim (ed.), pp. 214–231, 1990.
[Giere, 1984] R. N. Giere. *Understanding Scientific Reasoning.* Holt, Rinehart & Winston. New York, 1984.
[Giere, 1999] R. N. Giere. *Science Without Laws.* University of Chicago Press: Chicago, IL, 1999.
[Goldsmith, 1977] D. Goldsmith, ed. *Scientists Confront Velikovsky.* Cornell University Press: Ithaca, NY, 1977.
[Grim, 1990] P. Grim, ed. *Philosophy of Science and the Occult.* State University of New York Press: Albany, NY, 1990.
[Gross and Levitt, 1994] P. Gross and N. Levitt. *Higher Superstition.* Johns Hopkins University Press: Baltimore, MD, 1994.
[Grove, 1985] J. W. Grove. Rationality at Risk: Science against Pseudoscience. *Minerva* 23: 216–240, 1985.
[Grünbaum, 1984] A. Grünbaum. *The Foundations of Psychoanalysis. A Philosophical Critique.* University of California Press: Berkeley, CA, 1984.
[Haack, 2003] S. Haack. *Defending Science — Within Reason.* Prometheus Books: Amherst, NY, 2003.
[Hansson, 1996] S. O. Hansson. Defining Pseudo-Science. *Philosophia naturalis* 33: 169–176, 1996.
[Hines, 2003] T. Hines. *Pseudoscience and the Paranormal.* Prometheus Books: Amherst, NY, 2003.
[Holton, 1993] G. Holton. *Science and Anti-Science.* Harvard University Press: Cambridge, MA, 1993.
[Horgan, 1995] J. Horgan. *The End of Science.* Addison-Wesley: Reading, MA, 1995.

[Humphrey, 1999] N. Humphrey. *Leaps of Faith. Science, Miracles and the Search for Supernatural Consolation*. Copernicus-Springer: New York, 1999.

[Kanitscheider, 1991] B. Kanitscheider. A Philosopher Looks at Astrology. *Interdisciplinary Science Reviews* 16: 258–266, 1991.

[Kirkland, 2000] K. Kirkland. Paraneuroscience? *Skeptical Inquirer* 24(3): 40–43, 2000.

[Kitcher, 1982] P. Kitcher. *Abusing Science. The Case Against Creationism*. MIT-Press: Cambridge, MA, 1982.

[Kitcher, 1993] P. Kitcher. *The Advancement of Science. Science without Legend, Objectivity without Illusion*. Oxford University Press: New York, 1993.

[Kneale, 1974] W. C. Kneale. The Demarcation of Science. In: P.A. Schilpp (ed.), pp. 205–217, 1974.

[Knodel, 1985] H. Knodel, ed. *Neues Biologiepraktikum Linder Biologie* (Lehrerband). [New Practical Instruction in Biology, Teachers' Edition] J.B. Metzler: Stuttgart, 1985.

[Kuhn, 1962] T. S. Kuhn. *The Structure of Scientific Revolutions*. University of Chicago Press: Chicago, IL, 1962.

[Kuhn, 1970] T. S. Kuhn. Logic of Discovery or Psychology of Research? In: I. Lakatos & A. Musgrave (eds.), pp. 1–24, 1970.

[Kuipers, 2000] T. A. F. Kuipers. *From Instrumentalism to Constructive Realism*. Kluwer: Dordrecht, 2000.

[Kuipers, 2001] T. A. F. Kuipers. *Structures in Science. Heuristic Patterns Based on Cognitive Structures*. Kluwer: Dordrecht, 2001.

[Kuipers, 2004] T. A. F. Kuipers. De logica van de G-hypothese. Hoe theologisch onderzoek wetenschappelijk kan zijn. In: K. Hilberdink (ed.) *Van God los? Theologie tussen godsdienst en wetenschap*, pp. 59–74. KNAW: Amsterdam, 2004.

[Kurtz, 1985] P. Kurtz, ed. *A Skeptic's Handbook of Parapsychology*. Prometheus Books: Buffalo, NY, 1985.

[Kurtz, 2000] P. Kurtz. The New Paranatural Paradigm. *Skeptical Inquirer* 24(6): 27–31, 2000.

[Lakatos, 1970] I. Lakatos. Falsification and the Methodology of Research Programmes. In: I. Lakatos & A. Musgrave (eds.), pp. 91–197, 1970.

[Lakatos, 1973] I. Lakatos. Science and Pseudoscience. In: M. Curd & J.A. Cover (eds. 1998), pp. 20–26, 1973. `http://www.lse.ac.uk/collections/lakatos/scienceAndPseudoscienceTranscript.htm;21.2.05`

[Lakatos, 1974] I. Lakatos. Popper on Demarcation and Induction. In: P.A. Schilpp (ed.), pp. 241–273, 1974.

[Lakatos and Musgrave, 1970] I. Lakatos and A. Musgrave, eds. *Criticism and the Growth of Knowledge*. Cambridge University Press: New York, 1970.

[Laudan, 1983] L. Laudan. The Demise of the Demarcation Problem. In: M. Ruse, ed. *But Is It Science? The Philosophical Question in the Creation/Evolution Controversy*, pp. 337–350. Prometheus Books: Buffalo, NY, 1988.

[Ludwig, 1978] J. Ludwig, ed. *Philosophy and Parapsychology*. Prometheus Books: Buffalo, NY, 1978.

[Lugg, 1987] A. Lugg. Bunkum, Flim-Flam and Quackery: Pseudoscience as a Philosophical Problem. *Dialectica* 41: 221–230, 1987.

[Mahner and Bunge, 1997] M. Mahner and M. Bunge. *Foundations of Biophilosophy*. Springer-Verlag: Berlin, Heidelberg, New York, 1997.

[McGowan, 1994] D. McGowan. *What is Wrong with Jung?* Prometheus Books, Buffaly, NY, 1994.

[Merton, 1973] R. Merton. *The Sociology of Knowledge*. University of Chicago Press: Chicago, 1973.

[Nagel, 1956] E. Nagel. *Logic Without Metaphysics*. Free Press: Glencoe, IL, 1956.

[Niiniluoto, 1987] I. Niiniluoto. *Truthlikeness*. D. Reidel: Dordrecht, 1987.

[Popper, 1959] K. R. Popper. *The Logic of Scientific Discovery*. Hutchinson: London, 1959.

[Popper, 1963] K. R. Popper. *Conjectures and Refutations*. Basic Books: New York, 1963.

[Popper, 1994] K. R. Popper. Zwei Bedeutungen von Falsifizierbarkeit [Two Senses of Falsifiability]. In: H. Seiffert & G. Radnitzky (eds.) *Handlexikon der Wissenschaftstheorie*, pp. 82–85. Deutscher Taschenbuch Verlag: München, 1994.

[Poser, 2001] H. Poser. *Wissenschaftstheorie. Eine philosophische Einführung* [Theory of Science: A Philosophical Introduction]. Reclam: Stuttgart, 2001.

[Radner and Radner, 1982] D. Radner and M. Radner. *Science and Unreason*. Wadsworth Publishing Company: Belmont, CA, 1982.

[Reisch, 1998] G. A. Reisch. Pluralism, Logical Empiricism, and the Problem of Pseudoscience. *Philosophy of Science* 65: 333–348, 1998.

[Resnik, 2000] D. B. Resnik. A Pragmatic Approach to the Demarcation Problem. *Studies in History and Philosophy of Science*, 31: 249–267, 2000.

[Rhine, 1954] J. B. Rhine. The Science of Nonphysical Nature. In: J. Ludwig (ed. 1978), pp. 117–127, 1954.

[Rothbart, 1990] D. Rothbart. Demarcating Genuine Science from Pseudoscience. In: P. Grim (ed.), pp. 111–122, 1990.

[Runggaldier, 1996] E. Runggaldier. *Philosophie der Esoterik* [Philosophy of Esoterics]. Kohlhammer: Stuttgart, 1996.

[Schick and Vaugn, 1999] T. Schick and L. Vaughn. *How to Think About Weird Things*. Mayfield Publishing Company: Mountain View, CA, 1999.

[Schlipp, 1974] P. A. Schilpp, ed. *The Philosophy of Karl Popper*, vol. 1. Open Court: La Salle, IL, 1974.

[Settle, 1971] T. Settle. The Rationality of Science versus the Rationality of Magic. *Philosophy of the Social Sciences* 1: 173–194, 1971.

[Shermer, 1997] M. Shermer. *Why People Believe Weird Things*. W. Freeman: New York, 1997.

[Shermer, 2002] M. Shermer, ed. *The Skeptic Encyclopedia of Pseudoscience*, 2 vols. ABC-Clio: Santa Barbara, CA, 2002.

[Siitonen, 1984] A. Siitonen. Demarcation of Science from the Point of View of Problems and Problem-Stating. *Philosophia naturalis* 21: 339–353, 1984.

[Sokal and Bricmont, 1998] A. Sokal and J. Bricmont. *Intellectual Impostures*. Profile Books: London, 1998.

[Sorokin, 1956] P. A. Sorokin. *Fads and Foibles in Modern Sociology and Related Sciences*. H. Regnery: Chicago, IL, 1956.

[Spector, 1990] M. Spector. Mind, Matter, and Quantum Mechanics. In: P. Grim (ed.), pp. 326–349, 1990.

[Stalker and Glymour, 1989] D. Stalker and C. Glymour, eds. *Examining Holistic Medicine*. Prometheus Books: Buffalo, NY, 1989.

[Stalker and Glymour, 1989b] D. Stalker and C. Glymour. Quantum Medicine. In: D. Stalker & C. Glymour (eds.), pp. 107–125, 1989.

[Stenger, 1995] V. Stenger. *The Unconscious Quantum*. Prometheus Books: Buffalo, NY, 1995.

[Thagard, 1978] P. Thagard. Why Astrology is a Pseudoscience. In: P. Asquith & I. Hacking (eds.) *PSA* 1978, vol. 1, pp. 223–234. East Lansing, MI: Philosophy of Science Association, 1978. [Also in: M. Curd & J.A. Cover (eds. 1998), pp. 27–37.]

[Thagard, 1988] P. Thagard. *Computational Philosophy of Science*. MIT-Press: Cambridge, MA, 1988.

[Toulmin, 1984] S. Toulmin. The New Philosophy of Science and the "Paranormal". *Skeptical Inquirer* 9(1): 48–55, 1984.

[Vollmer, 1990] G. Vollmer. Against Instrumentalism. In: P. Weingartner & G.J.W. Dorn (eds.) *Studies on Mario Bunge's Treatise*, pp. 245–259. Rodopi: Amsterdam, 1990.

[Vollmer, 1993] G. Vollmer. Wozu Pseudowissenschaften gut sind [What Pseudosciences Are Good For]. In: G. Vollmer, *Wissenschaftstheorie im Einsatz* [Philosophy of Science in Action]. Hirzel-Verlag: Stuttgart, 1993.

[Webster and Goodwin, 1996] G. Webster and B. C. Goodwin. *Form and Transformation*. Cambridge University Press: Cambridge, UK, 1996.

[Weingartner, 2000] P. Weingartner. *Basic Questions on Truth*. Kluwer: Dordrecht, 2000.

[Weston, 1992] T. Weston. Approximate Truth and Scientific Realism. *Philosophy of Science* 59: 53–74, 1992.

[Wilson, 2000] F. Wilson. *The Logic and Methodology of Science and Pseudoscience*. Canadian Scholars' Press: Toronto, 2000.

[Wilson, 1999] R. A. Wilson. Realism, Essence, and Kind: Resuscitating Species Essentialism? In: R.A. Wilson (ed.) *Species: New Interdisciplinary Essays*, pp. 187–207. MIT-Press: Cambridge, MA, 1999.

[Wittgenstein, 1921] L. Wittgenstein. *Tractatus Logico-Philosophicus*. Suhrkamp: Frankfurt 1921/1960.

[Wolpert, 1992] L. Wolpert. *The Unnatural Nature of Science*. Faber & faber: London, 1992.

[Young, 1970] R. M. Young. *Mind, Brain and Adaptation in the Nineteenth Century: Cerebral Localization and Its Biological Context from Gall to Ferrier*. Oxford University Press: Oxford, 1970.

[Ziman, 1994] J. M. Ziman. *Prometheus Bound*. Cambridge University Press. Cambridge, UK, 1994.

HISTORY OF THE PHILOSOPHY OF SCIENCE. FROM *WISSENSCHAFTSLOGIK* (LOGIC OF SCIENCE) TO PHILOSOPHY OF SCIENCE: EUROPE AND AMERICA, 1930–1960

Friedrich Stadler

"Philosophy of science without history of science is empty,
history of science without philosophy of science is blind"
Imre Lakatos, 1974

PRELIMINARY REMARKS

There seems a general consensus in the scientific community that modern philosophy of science — as a subdiscipline of (scientific) philosophy — has emerged as a genuine academic research and teaching field as well as an institution only since the middle of the 20^{th} Century.

Accordingly, we can reconstruct a process of differentiation and professionalization of philosophy of science from the ancient Greek philosophy (Pre-Socratics, Plato and Aristotle) via the rationalist and empiricist philosophers of the "Scientific Revolution" to the Enlightenment up to the (Neo-)Kantian versions of science-oriented philosophy. These developments lead to re-evaluations of a "Theory of Science" (Wissenschaftslehre) in the 19th century in close interaction with the rise of the empirical sciences between physics, physiology and psychology — as is typically illustrated with the philosopher-scientists Ernst Mach and Ludwig Boltzmann. In parallel, this dynamics of departure from, and interaction with traditional philosophy as a universal normative discipline was accompanied by a specific focus on the methods of the natural sciences in general, but also in the cultural and social sciences: historism as well as the "probabilist revolution" in the cultural sciences [Ringer, 1997]. There is a re-conceptualization of Empiricism and Rationalism [Santillana and Zilsel, 1941], which anticipated the formation of Logical Empiricism between the two World Wars in the 20th century. The context for this innovation was the so-called "Second Scientific Revolution" in science, with Einstein's Relativity Theory and Quantum Physics around Bohr and Schrödinger and the input of modern symbolic logic and set theory with Frege, Russell and

Handbook of the Philosophy of Science: General Philosophy of Science - Focal Issues, pp. 577–658.
Volume editor: Theo Kuipers. General editors: Dov M. Gabbay, Paul Thagard and John Woods.

Whitehead, Wittgenstein and Gödel. Generally, we find therein a permanent tension between a normative philosophy of science (methodology) on the one hand, and a descriptive history of science (theory dynamics) on the other, which indicates the later introduced distinction of the context of justification and the context of discovery as a main issue in this "Rise of Scientific Philosophy" [Reichenbach, 1938/1951].

It is not surprising, that most textbooks and the few handbooks or encyclopedias in the philosophy of science do not explicitly deal with the history of its own discipline, or are restricted regarding themes and time periods. The important comprehensive historical study on philosophy of science as a monograph from Greek philosophy up to contemporary issues is also limited to the natural sciences and its methodologies [Losee, 1972/2001]. Another restriction — with some exceptions [Serres, 1989; Collins, 1998] — is a strong European perspective disregarding the important contribution of the Chinese and Islamic world to the sciences and their philosophy. And this bias is re-inforced by the missing gender perspective, although in the meantime there are valuable contributions to the problem of women in philosophy and in the philosophy of science, especially feminist philosophy (of science) [Fricker and Hornsby, 2000].

Given this status quo of the fragmentary historiography and research and with reference to related contributions in this volume, the subsequent "History of Philosophy of Science" is restricted to a paradigmatic case study of transfer, transformation and institutionalization of philosophy of science from Europe to America, in the period from 1930 to 1960. This is done with reference to the standard volume on *The Intellectual Migration. Europe and America, 1930–1960* [Fleming and Bailyn, 1969], and to volumes on the origins and influence of Logical Empiricism in America and the history of the Vienna Circle leading up to 1938 [Stadler, 1997/2001; Giere and Richardson, 1996; Hardcastle and Richardson, 2004; Richardson and Uebel, 2006].

The point of departure of the following account is the rise of Logical Empiricism in Central Europe before the forced migration, and the end of this intellectual and institutional history is the placement of "The *Wiener Kreis* in America" [Feigl, 1969] in the Cold War period. The further development is characterized by the criticism of the so called "received view" and the re-transfer of a modified philosophy of science — as a mostly analytic and normative methodology — back to Europe, which was dealt with only in the last years. And these currents since Quine's "Two Dogmas" [1951] are a remarkable return of a hidden agenda in the described history of philosophy of science, namely the pragmatic and historical turn. These new insights allow us to reasonably speak of a re-union of the history *and* philosophy of science, or of a (cultural) history of philosophy of science, which does not privilege the context in comparison with the (meta-)theoretical dynamics.

1 THE EMERGENCE OF PHILOSOPHY OF SCIENCE: "WISSENSCHAFTSLOGIK" (LOGIC OF SCIENCE) BEFORE 1938

The emergence of the discipline known today as "philosophy of science" can be seen as converging with the process of the increasingly scientific status of philosophy, the so-called "rise of scientific philosophy" (Reichenbach 1951) in the inter-war years. Already in the programmatic text of the Vienna Circle (*Wissenschaftliche Weltauffassung. Der Wiener Kreis,* 1929), the autonomous regal discipline of philosophy had given way to an antimetaphysical, physicalist unified science. This idea was systematically elaborated in the thirties, most notably in Rudolf Carnap's writings. In the manifesto, reference had primarily been made to his *Logical Structure of the World* [Carnap, 1928] — as a constitutive system based on experience with logical analysis. A few years later the position he took in his *Logical Syntax* [1934a] found acceptance. The task of "Wissenschaftslogik" [1934b] is seen as lying in the study of science as a whole or in its disciplines:

> "The concepts, propositions, proofs, theories appearing in the various realms of science are analyzed — less from the perspective of the historical development of science or of the sociological and psychological conditions of its functioning, but more from a logical perspective. This field of work for which no generic term has been able to gain acceptance, could be called theory of science or to be more precise logic of science. Science is understood as referring to the totality of accepted propositions. This does not just include the statements made by scholars but also those of everyday life. There is no clear boundary line drawn between these two areas." [Carnap, 1934b, p. 5]

Here the distancing from traditional philosophy becomes highly salient, even if the role and function of a scientific *philosophy*, as linguistic analysis in Wittgenstein's sense, is not called into question. This new discipline is not so interested in propositions on the external world as the realm of the empirical disciplines (thing language), as in "science itself as an orderly structure of propositions", known as object language (ibid., p. 6) – accordingly, in the "sense" of the propositions and the "meaning" of concepts from a logical point of view. The realm of these concepts is limited either to the analytic propositions of logic/mathematics or to the empirical propositions of the sciences. This culminates in the view:

> "that the propositions of the logic of science are propositions of the logical syntax of language. Thus these propositions lie within the boundaries drawn by Hume, for logical syntax is ... nothing other than mathematics of language." (ibid.)

Before the logic of science as a "Wissenschaftslehre" (theory of knowledge or theory of science) was promulgated in the 19th century, for instance by Johann G. Fichte, Bernard Bolzano and Ernst Mach, the term "theory of science" was in circulation as an alternative to classical philosophy besides empirical disciplines

[Losee, 1980]. Nevertheless, it was the first time here that the so-called "overcoming of metaphysics through the logical analysis of language" [Carnap, 1931a] was propagated.

Carnap had combined the elaboration of this program of unified science in his *Logical Syntax of Language* [1934a] with its promulgation. As part of the internationalization of the Vienna Circle under way since 1929, two small books appeared almost at the same time in England, i.e., *The Unity of Science* [1934c] and *Philosophy and Logical Syntax* [1935] in the series "Psyche Miniatures" published by Kegan Paul. The former was an edition of the German article on physical language [Carnap, 1931b], reworked by the author and translated by Max Black. The latter united three lectures that Carnap had given at the University of London in October of 1934: "The Rejection of Metaphysics", "Logical Syntax of Language", "Syntax as the Method of Philosophy". These attempts to popularize "Logic of Science" in the Anglo-Saxon world were continued by the translation of *Logical Syntax* which appeared in 1937 in a expanded edition at the same English publisher [Carnap, 1937].

It is known that already in his *Logical Syntax* Carnap had been influenced by Polish and American logicians and philosophers of science (Tarski, Quine and Morris) to further develop the possible field of "Logic of Science". In addition to the syntactic dimension, he cited the semantic and pragmatic dimensions as future fields of work. Accordingly, he described the logic of science in his preface to the second edition as the "Analysis and Theory of the Language of Science":

> According to the present view, this theory comprises, in addition to logical syntax, mainly two further fields, i.e., semantics and pragmatics. Whereas syntax is purely formal, i.e., only studies the structure of linguistic expression, semantics studies the semantic relationship between expressions and objects or concepts; ... Pragmatics also studies the psychological and sociological relations between persons using the language and the expressions. [Carnap, 1968, VII]

With this new conceptualization of the logic of science, which already took place before the transfer of these ideas to the United States, we have also outlined the logical space for the philosophy of science as well as the terminological structure for the Unity of Science movement [1934ff]. Of course, Logical Empiricism before 1938 had no codified understanding of "logic of science" in relation to philosophy. Here, however, only the paradigmatic elements have been indicated which proved to be relevant later in the Anglo-American realm. In this context, I cannot dwell on the controversial protocol statement debate within the Vienna Circle in which various positions on the basic issue of knowledge were unearthed, e.g., in Edgar Zilsel's "Bemerkungen zur Wissenschaftslogik" (Notes on the Logic of Science) [1932/33]. This eventually led to a heated debate on fundamental questions in the epistemology of that time [Uebel, 1992].

The fact that there was, on both sides, a strong reception of European positivism centered around Ernst Mach and of American pragmatism focusing on William

James clearly shows that the trans-Atlantic process of communication did not suddenly begin in the 20th century. Rather, there had been a continuous process of international exchanges between related intellectual movements (positivism and pragmatism, operationalism and behaviorism) which became manifest in classical Logical Empiricism.

The direct contacts between Mach, William James and Paul Carus — the editor of the journals *Monist* and *Open Court* — paved the way for a strong convergence of Logical Empiricism and neo-pragmatism with Otto Neurath and Rudolf Carnap, on the one hand, and John Dewey and Charles Morris, on the other. The positive reception of Percy Bridgman's *The Logic of Modern Physics* [1927] has often been described, most notably by Philipp Frank [1949], as a milestone in this theoretical rapprochement. Even in psychology, there was direct cooperation, facilitated by Edward Tolman, Egon Brunswik and the Bühler-School, which led to a transfer of individual scholars and ideas [Fischer and Stadler, 1997; Smith, 1986]. (At that time, however, this did not mean that behaviorism predominated, since within the context of the *Encyclopedia of Unified Science*, psychoanalysis or cognitive psychology was seen as being at least equal, as Egon and Else Frenkel-Brunswik's attempts to integrate them show.)

The reception of philosophical ideas between the old continent and the United States is meanwhile well documented [Giere and Richardson, 1996; Hardcastle and Richardson, 2003].

The historian of science Gerald Holton, who played a seminal role in the forties in the Unity of Science Institute and as an assistant of Philipp Frank, has given a very apt reconstruction of these cognitive parallels and this transfer of knowledge in his "From the Vienna Circle to Harvard Square: The Americanization of a European World Conception" [1993]. This history of ideas, which also includes Quine, describes a growing internationalization best illustrated by the International Congresses of the *Encyclopedia of Unified Science* and the "Unity of Science Institute" founded by Frank. Holton characterizes the favourable conditions for Logical Empiricism in the United States from 1940 to 1969 metaphorically as an "ecological niche" in the New World and depicted these developments as an osmotic success story. Another (auto-)biographical study of Holton on this phenomenon already alluded in the title — "On the Vienna Circle in Exile" [1995[— to the possibility of the Viennese philosophy of science finding a dynamic (and transitory) context in American intellectual life. It should only be added here that these theoretical developments in the history of science called into question the dominant idea of Unified Science, both intensifying the inherent contradictions and overcoming them — which went almost unnoticed within this setting. If one takes into account the parallel reciprocal tie between positivism and pragmatism since the turn of the century leading up to the synthesis in today's analytic philosophy [Dahms, 1987/88; 1994], the non-linear and self-organizational theoretical dynamic becomes evident in all its historical and systematic complexity. If one wishes to make a qualitative assessment of the transfer of science through intellectual emigration against the background of the discourse of the scientific community, the actual contacts must be studied irrespective of the history of reception related to emigration history. This would include, for instance, Moritz Schlick's early lecture trips and visiting professorships. His two sojourns in the United States left traces as well as his presentation at the Seventh International Congress of Philosophy in Oxford in 1930 where he addressed the programmatic, linguistic-analytic turn in philosophy in his "The Future of Philosophy" lecture. Here he advocated the dissolution of the classical philosophical canon by drawing a functional distinction between scientific philosophy on the one hand and the related scientific theorizing on the other:

> "There will always be men who are especially fitted for analyzing the ultimate meaning of scientific theories, but who may not be skilful in handling the methods by which their truth or falsehood is ascertained. These will be the men to study and to teach philosophizing, but of course they would have to know the theories just as well as the scientist who invents them. Otherwise they would not be able to take a single step, they would have no object on which to work. A philosopher, therefore, who knew nothing except philosophy would be a knife without blade and handle. Nowadays a professor of philosophy very often is a man who is not able to make anything clearer, that means he does not really philosophize at all, he just talks about philosophy or writes a book about it. This will be impossible in the future. The result of philosophizing will be that no more book will be written about philosophy, but all books will be written in a philosophical manner." [Schlick, 1931, p. 116]

Here this necessary merging of philosophizing with the results of the so-called empirical sciences documents — in Wittgenstein's language — the transformation of philosophy as a regal discipline into the "maiden of sciences". This was the result of calling into question the existence of an autonomous area. It also comes as no surprise when one considers Schlick's allegiance to British empiricism and to the scientific orientation of philosophy since the turn of the century. His visiting professorsips in Stanford [1929] and Berkeley [1931/32] reinforced this anglophile leaning, but also paved the way for a gradual shift toward the United States which Herbert Feigl had already begun in 1930. Schlick reported on his impressions in his lecture "On the Scientific World View in the U.S.A." which he gave at the "Verein Ernst Mach" (Ernst Mach Society) in Vienna in 1930. Here he drew attention to the fact that there was a well-developed everyday rationalism, accompanied by favorable conditions for the scientific world view thanks to empirical psychology and John Dewey's pragmatism. [Schlick, 1930/31, p. 76]. It is thus no coincidence that Schlick already figured on the Advisory Board in the first issue of the quarterly *Philosophy of Science*, which had been founded in 1934, with his student Herbert Feigl in the Editorial Board — together with Rudolf Carnap. The institution behind this journal, edited by William M. Malisoff, was the, still existing, "Philosophy of Science Association" (PSA) as an

> "organized expression of the will of a fairly large body of intellectually competent individuals whose basic interest is in both science and philosophy, and particularly in their union... The Association humbly undertakes the task of uniting in one body the scattered elements available for this enterprise." (Announcement of the Publisher)

It should also be added that Schlick's assistant in Stanford was the young Paul Arthur Schilpp, who served as the editor of the important and influential book series "The Library of Living Philosophers" (including volumes on Carnap, Dewey, Lewis, Quine, Popper, Russell, Einstein, among others). With this series he contributed greatly to the continuation of the discussions in analytic as well as pragmatist philosophy of science.

Proceeding from the early forties as the beginning of the specific American philosophy of science, it is possible to reconstruct the intellectual conditions of the convergent development of Central European and US-American philosophy of science [Stadler, 2004, 227ff].

In the contemporary *Dictionary of Philosophy* [Runes, 1944] we find the relevant discussions of that time presented in various short entries. Here it becomes clear that the central contributions to the philosophy of science were written by Rudolf Carnap, Carl G. Hempel and Heinrich Gomperz. Carnap presents philosophy of science as

> "that philosophic discipline which is the systematic study of the nature of science, especially of its methods, its concepts and presuppositions, and its place in the general scheme of intellectual disciplines. No very precise definition of the term is possible since the discipline shades imperceptibly into science, on the one hand, and into philosophy in general, on the other. A working division of its subject-matter into three fields is helpful in specifying its problems, though the three fields should not be too sharply differentiated or separated." [Carnap, 1944, p. 284]

According to Carnap the three fields addressed here are the following:

1. A critical study of the method or methods of the sciences, of the nature of scientific symbols, and of the logical structure of scientific symbolic terms...
2. The attempted clarification of the basic concepts, presuppositions and postulates of the sciences, and the revelation of the empirical, rational, or pragmatic grounds upon which they are presumed to rest. ...
3. A highly composite and diverse study which attempts to ascertain the limits of the special sciences, to disclose their interrelations one with another, and to examine their implications so far as these contribute to a theory either of the universe as a whole or of some aspect of it. [Carnap, 1944, 284f.]

In a preceding section, Carnap had already subsumed today's science studies under "science of science" as "the analysis and description of science from various points of view, including logic, methodology, sociology, and history of science" [ibid]. In this connection he referred to his entry on "Scientific Empiricism" and the "Unity of Science" as "a wider movement, comprising besides Logical Empiricism other groups and individuals with related views in various countries" [ibid., p. 286].

The Unity of Science was also identified with internationalization. "Scientific Empiricism" was introduced as a transformation of Logical Empiricism. With this self-understanding, the institutionalization and further differentiation of philosophy of science took place — a development which had been anticipated by two decades of intellectual exchange between Europe and America. But in parallel to this transatlantic movement the European philosophy of science is to be addressed in more detail.

2 THE *WIENER KREIS* IN GREAT BRITAIN: EMIGRATION AND INTERACTION

2.1 Prologue

In 1968 the Austro-American philosopher of science Herbert Feigl (1902–1988) published a remarkable, largely autobiographical essay on "The *Wiener Kreis* in America". This historical and theoretical account of the Vienna Circle's emigration story was first anthologized in the second volume on *Perspectives in American History* [1968], edited by the "Charles Warren Center for Studies in American History" and then included in the standard volume on *The Intellectual Migration. Europe and America, 1930–1960*, published in 1969 by Harvard University Press — distributed in Great Britain by Oxford University Press. A last reprint can be found in Feigl's *Inquiries and Provocations. Selected Writings 1929–1974* [Feigl, 1981].

Together with his influential article on "Logical Positivism" (co-authored with Albert Blumberg in the *Journal of Philosophy)* that already appeared in 1931, the year of his definitive emigration to the USA — the above-mentioned publication marked a watershed in the historiography of Logical Empiricism as a paradigmatic intellectual history of forced migration — in addition to the autobiographical reports of Philipp Frank [1949] and Rudolf Carnap [1963] inter alia:

> It was this article, I believe, that affixed this internationally accepted label to our Viennese outlook ... Blumberg and I felt we had a mission in America, and the response to our efforts seemed to support us in this. We had, indeed, 'started the ball rolling', and for at least twenty years Logical Positivism was one of the major subjects of discussion, dispute, and controversy in United States philosophy. [Feigl, 1968, p. 646f.].

In his essay on one of the most influential philosophical movements in the field of philosophy of science coming from Central Europe to the USA Feigl reconstructed the intellectual and institutional trajectory of the Vienna Circle from a personal and professional perspective in what could best be described as a sort of philosophical "oral history". Starting with the origins and development in Vienna, Feigl describes the early contacts with American philosophers (Schlick 1929/1931 in Stanford and Berkeley) and the beginnings of the Vienna Circle's migration from the 1930s onwards. His own contacts with Dickinson Miller and Charles A. Strong (a son-in-law of John D. Rockefeller) enabled the Schlick student and gifted young philosopher of science to embark upon his brilliant academic career in Harvard with a Rockefeller Fellowship.

Because Carnap, Frank and most of the other members had to emigrate to the USA, we still lack a complementary account — a sort of "*Wiener Kreis* in Great Britain". Moreover, the most influential history of this transfer and intellectual transformation came from Alfred J. Ayer, who had attended the Vienna Circle in 1932–33, with his publications on the history and influence of the Viennese philosophy, especially with his booklet *Language, Truth and Logic* [1936a] — a publication that was influential into the postwar period. This was reinforced by Ayer's "The Vienna Circle" (in *The Revolution in Philosophy*, 1956) and his textbook volume on *Logical Positivism* [1959]. I do not want to deal with all the factors influencing the intellectual migration and cultural transformation of the Vienna Circle to Great Britain but only provide some significant material in order to criticize what I will call the *standard view* of "Logical Positivism" in England.

This widespread position has been challenged over the last decade in studies in the history and philosophy of science but we still lack a critical reconstruction analogous to the better researched topic of *Origins of Logical Empiricism* [Giere and Richardson, 1996] and *Logical Empiricism in North America* [Hardcastle and Richardson, 2003].

This *standard view* is determined on the one hand by Ayer's role as most important mediator and interpreter, and, on the other hand, by the extensive research on Ludwig Wittgenstein's impact on English analytic philosophy before and after World War II. The essence of this traditional historiographical account is that via Logical Positivism of the Vienna Circle (partly with Popper) postwar philosophy of science in Great Britain was directly influenced, whereas analytic philosophy, especially "ordinary language philosophy", was mostly motivated by Wittgenstein's late philosophy of the *Philosophische Untersuchungen/Philosophical Investigations* (published posthumously in 1953).

In the following account I want to show — in an admittedly cursory fashion — that

1. this traditional image of Logical Empiricism has shortcomings and is highly selective.

2. this distinction between two different currents (Vienna Circle vs. Wittgenstein) is rather artificial.

3. there was a flourishing communication among the dominant figures I have already mentioned, which seems to me at least an important correction of the usual history of reception.

From a *biographical* point of view this means that the players in this complex intellectual history have to be extended on both sides. Accordingly, we are dealing with a typical example of networking in the period from the 1930s to the 1960s with regard to “The *Wiener Kreis* in Great Britain”. My intention is to show that, parallel to the American story there was another interconnected development and that it is futile to argue that “Continental”, “British” and “American” branches all existed as separate movements.

2.2 Intersections and Interventions: Philosophy of Science and Analytic Philosophy between Central Europe and Great Britain

The process of interaction between the Continental and the British tradition in analytic/scientific philosophy did not suddenly begin in the 1930s: The key figure is without doubt Bertrand Russell, together with Einstein and Wittgenstein — “the leading representatives of the scientific world-conception” mentioned in the Vienna Circle’s manifesto in 1929. Besides this context his early dispute with Alexius Meinong [1905] and his lifelong conflict-ridden involvement with the life and work of Ludwig Wittgenstein are well known, even if strongly contested in its interpretation.

It also seems plausible, that Russell was one of the most important partners for “Austrian philosophy” from the turn of the century until his American days, if we reconstruct this coincidence by overcoming the dominant stories on his relations to the Vienna Circle — and Wittgenstein, too (especially with the reception of his book *Our Knowledge of the External World*, [1914]).

I am not referring to the prehistory of the bilateral scientific-philosophical reflection as becomes manifest in the correspondence of Ernst Mach with Pearson, Whewell or the reception of John Stuart Mill with Theodor and Heinrich Gomperz ([Stadler, 2001, Ch. 3] and with reference to the US: [Holton, 1993]). This intellectual impact can be illustrated more precisely by the internationalization of Logical Empiricism in Europe and America, especially with a focus on the six “International Congresses for the Unity of Science” 1934–1941 organized in Paris (twice), Copenhagen, Cambridge (in England at Girton College), Harvard and Chicago:

At the latest in Paris in 1935 we encounter the first significant presentation of Logical Empiricism in an international context with an increasing overlap into the Anglo-Saxon world [Stadler, 2001, 363–371]. Russell gave a widely acclaimed opening address in which he presented “scientific philosophy” as a synthesis of logic and empiricism [Russell, 1936]. And for the American delegates Charles Morris had stressed the international cooperation of scientists, which indicated the growing Austro-American relations. But what about the British connection?

Already in January 1935 Neurath had organized an informal meeting on Logical Positivism in London (held at Belsize Park) with A. J. Ayer, G. E. Moore, Max Black, and Carl G. Hempel, which resulted in a still rather skeptical statement on common points in Vienna and Cambridge [Stebbing, 1935].

Afterwards in Paris, the young Ayer delivered a paper on "The Analytic Movement in Contemporary Philosophy" [1936b] referring to the analogous anti-metaphysical movements in Vienna and England since the turn of the century. He expressed his hope for a stronger interpenetration of science and philosophy — as opposed to the often deplored "Scientism" as "Infatuation of Philosophy with Science" [Sorell, 1991].

Otto Neurath gave a very positive account on the Paris congress, conveying his "impression that there was in fact something like a scholar's republic of logical empiricism" [Neurath, 1936, 377]. This, by the way, seems to be congruent with what was happening at the time: Robert Musil tried to get an invitation for the congress, the "Frankfurt School" sent Walter Benjamin, and Bert Brecht expressed his interest in cooperating with Neurath. The meeting of American pragmatists (Charles Morris), the English analytic philosophers (Susan Stebbing), the Polish logicians (Kasimir Ajdukiewicz) and of Italian scientific philosophers (Federigo Enriques) illustrated this tendency towards a unification of empiricism and rationalism, but also the international setting of this ambitious project.

A year later Russell [1936, 10f.], who had delivered his laudatio of Gottlob Frege in German (!), wrote, that "The congress of Scientific Philosophy in Paris in September 1935, was a remarkable occasion, and, for lovers of rationality a very encouraging one ...". Russell's review perfectly illustrates the international context of the already exiled "Vienna School", and particularly its rational-empiricist interpretation in the spirit of Galileo and Leibniz [ibid., 11]:

> Modern science arose from the marriage of mathematics and empiricism; three centuries later, the same union is giving birth to a second child, scientific philosophy, which is perhaps destined to as great a career. For it alone can provide the intellectual temper in which it is possible to find a cure for the diseases of the modern world.

Whether scientific philosophy was able to fulfill Russell's hopes is not for us to decide. Still his statement remains an impressive document of an atmosphere of optimism and awakening — one of the last ones before the war in Europe amidst a constantly growing tendency towards totalitarian "(final) solutions".

A large international committee for the International Congresses for the Unity of Science was formed, including the English members C. K. Ogden, Bertrand Russell, Susan Stebbing and Joseph H. Woodger. All of them will play a significant role in the convergence and divergence of Logical Empiricism and British philosophy (of science). The case of Woodger as biologist of the Unity of Science movement and the "Theoretical Biology Club" in Oxford is one of the issues which deserves further investigation.

The "Fourth International Congress for Unity of Science" on the main topic "Scientific Language" in Cambridge (UK, Girton College) was the last European meeting of the community in scientific philosophy, which took place in the framework of a larger Enyclopedia-oriented program some months after Austria's occupation by Nazi-Germany.

It also documents the high level of the dialogue between British and Central European proponents of scientific and analytic philosophy. At the same time it also provided a forum for constituting the international committee for the forthcoming congresses and the organizational committee for the *International Encyclopedia of Unified Science* [Carnap *et al.*, 1938ff].

In his inaugural address G.E. Moore focused on the historical reference point of Cambridge philosophy, i.e., Russell's and Whitehead's *Principia Mathematica,* but surprisingly without mentioning Wittgenstein, who was not present at the congress. Oxford Philosophy was represented by Gilbert Ryle, who discussed the practical and theoretical reasons for the "disunity of sciences".

Finally, one of the most important figures of this dialogue, namely Susan Stebbing (1885–1943), host and initiator of the congress, spoke about "Language and Misleading Questions", apparently in the spirit of Wittgenstein, but with a remarkable preference for Carnap's alternative:

> Since the conference is meeting in Cambridge and since its topic is 'Scientific Language', it seems to me not inappropriate to take for this inaugural address 'Language and Misleading Questions'. For it is, perhaps, to Wittgenstein more than to any other philosopher that the conception of philosophy as 'the critique of language' is due. His influence has, so I understand, now permeated Cambridge students of philosophy that to the outsider all their discussions appear to be concerned with investigation of language ... I have learnt even more from studying Carnap's writings. I have felt the attraction of the view that: 'an die Stelle des unentwirrbaren Problemgemenges, das man Philosophie nennt, tritt die Wissenschaftslogik.' [Stebbing, 1939–40, p. 1]

Despite this professed affiliation with the "Logic of Science" Stebbing concludes her paper with reference to Heinrich Hertz's *Principles of Mechanics*, which contains a linguistic critique of (metaphysical) questions and answers, with another (early) Wittgensteinian thought:

> "We want an answer to a question we have not asked. Our minds cease to be vexed when we find that the question is illegitimate; we no longer seek for an answer for there is no longer a question to be asked" [ibid., 6].

As can be seen from the congress report, however, the program focused on logical-analytical questions, with many special contributions on "scientific language". Inter alia, Otto Neurath who later should have to flee from the Netherlands (The Hague) to England in 1940, postulated many small scientific units as a logical starting point for the development of a future unified science, once again directing polemical attacks against one privileged "system" — as preference of 'Encyclopedism' vs. hierarchical , Pyramidism' [Neurath, 1983, esp. chapters 8–23].

Among the printed contributions of this congress the paper by the British-American philosopher Max Black on the "Relations between Logical Positivism and the Cambridge School of Analysis" is of special interest, because it offers a profound discussion from a British point of view of what Wittgenstein, the Vienna Circle and the Cambridge School have in common and what separates them (cf. [Skorupski, 1993]):

> ... the development of the analytical movement in England and of Logical Positivism are found to have much in common. They have had, roughly speaking, the same friends and the same enemies. The teachings of Wittgenstein, Russell, Moore and the earlier English empiricists have been among the most important formative influences of both. If Logical Positivists have proclaimed their attachment to the advance of science more loudly, the English movement, (...), has to some extent been permeated with the same values. There should be room for further fruitful interchange of opinions between the two movements. [Black, 1939/40, 33f.]

With this argument the translator of Frege and Carnap once more described the background for the relationships between the German-speaking and the Anglo-Saxon world in the philosophical field. But this convergence was interrupted for several years by the War, which became manifest in 1941, when the European participants of the "Sixth Congress for the Unity in Science" at the University of Chicago who had already registered met instead for a small conference on "Terminology" on October 3–5, 1941 at Linton Road in Oxford. This event was once again organized by Neurath who was living in exile, together with J. A. Lauwerys and Susan Stebbing, shortly after his release from the internment at the Isle of Man.

We do not possess proceedings, but Neurath's publications from this period on "Universal Jargon and Terminology" [1941] and "The Danger of Careless Terminology" (1941) seem to indicate the motivation and orientation of this joint activity – which ended significantly with another variation of his famous boat metaphor [Uebel, 1992; 2000]. This contribution was partly evoked by the critical remarks of Bertrand Russell in his William James Lectures *An Inquiry into Meaning and Truth* [1940], where the author writes in the preface: "I am, as regards method, more in sympathy with the logical positivists than with any other existing school". In order to avoid so-called 'ontological misinterpretations' as a consequence of a correspondence theory of truth Neurath accordingly directs his proposals towards the history and sociology of sciences (American neo-pragmatism), because "no judge is in the air who says which of us has the TRUTH" [Neurath, 1983, p. 229]. By the way, this nonfoundationalist attitude corresponds to his radical criticism of Schlick and Popper, thereby rejecting verificationism *and* falsificationism as absolute philosophies of science [Neurath, 1935].

In 1941 the Chicago congress united Americans and European emigrants as well as "Contributors from Europe" whose papers were presented in the absence of their authors — among them Friedrich Waismann and Martin Strauss.

From 1938 on we can note a renewed convergence of the *International Encyclopedia of Unified Science* (with Neurath, Morris and Carnap as main editors) and the transformation of the journal *Erkenntnis* [1930ff.], (ed. by Carnap and Reichenbach) into an English edition as *Journal of Unified Science* with the eighth and last volume. Since 1933 it had come under increasing pressure after the Nazis seized power. The first volume of the *International Encyclopedia*, with contributions by Neurath, Niels Bohr, John Dewey, Carnap and Russell again marked the beginning of the (uncompleted) project with the University of Chicago Press.

Here Russell wrote "On the Importance of Logical Form" [1938, p. 41] based on the instrument of mathematical logic for pure mathematics and the empirical sciences with the (somewhat Popperian) conclusion that

> "the unity of science ... is essentially a unity of method, and the method is one upon which modern logic throws much new light. It may be hoped that the Encyclopedia will do much to bring about an awareness of this unity".

When we look at the authors of the *Foundations of the Unity of Science* and also consider the members of the "Advisory Committee", we can detect a strong UK/US-Austrian dominance. A first analysis of the contents of the 8 volumes of the *Erkenntnis* shows a similar development as the six Congresses for the Unity of Science: among the English-language speaking authors of the journal from 1934 on — apart from the printed congress contributions — we find, for instance, Ayer, Black, Stebbing. The reviews on these publications reflect the increasing reception of the philosophy of science. From the first issue on, works by Ayer, Church, Lewis, Nagel, Quine and Woodger were presented within the context of what was still Central European "Wissenschaftslogik", the Logic of Science according to Carnap.

2.3 Transfer and Transformation: Circles and Networks Continued

Apart from these rather well known developments of the *Wiener Kreis* in the Anglo-Saxon world we can reconstruct another story of international relations in philosophy of science between the wars, which is also representative for the scientific communication within Europe and between Europe and America. Let me illustrate this phenomenon by describing these features with the intellectual networking before and after the forced migration:

It is worth noting that it was, above all, Moritz Schlick (1882–1936) (married to an American citizen Blanche Hardy) the founder and head of the Vienna Circle, who fostered early intellectual contacts with the English-speaking world: he visited England at least twice in the late 1920s as can be shown by his correspondence with Frank P. Ramsey [1927/28], who stood in close personal contact with both Wittgenstein and Schlick. Ramsey, who invited Schlick to the famous "Moral Sciences Club", discussed his controversy with Wittgenstein on the Philosophy of Mathematics:

> "I had a letter the other day from Mr Wittgenstein criticising my paper 'The Foundations of Mathematics' and suggesting that I should answer not to him but to you. I should perhaps explain what you may have gathered from him, that last time we didn't part on very friendly terms, at least I thought he was very annoyed with me (for reasons not connected with logic), so that I did not even venture to send him a copy of my paper. I now hope very much that I have exaggerated this, and that he may perhaps be willing to discuss various questions about which I should like to consult him. But from the tone of his letter and the fact that he gave no address I am inclined to doubt it." (Ramsey to Schlick, July 22, 1927).

And in one of his last letters before his premature death he reports to Schlick on Cambridge philosophy:

> "It is a great thing for us to have Wittgenstein here, he is such a great stimulus and has been doing most excellent work, quite destroying my notions on the Foundations of Mathematics. Apart from that I think the school of philosophy here is severing a little; there are more and better pupils, and a distinct improvement from the very low level we were at when you visited us." (Ramsey to Schlick, probably Dec.1929).

The *Journal of Philosophy* which has existed since the turn of the century dealt directly and indirectly, especially in the inter-war years, with the work of John Dewey and functioned in America as a moderate forum for the development from scientific philosophy to philosophy of science, as becomes clear in the contributions of Ernest Nagel, Willard Van Orman Quine, Carl G. Hempel or Nelson Goodman. This trend may be illustrated e.g., with Nagel's informative 'Impressions and Appraisals of Analytic Philosophy' (1936). His reports from Cambridge, Vienna, Prague, Warsaw and Lwow of the early thirties is a document of the advanced stage of internationalization in Europe and between Europe and America. Thus he correctly observes that in the 'Wiener Kreis' "significant shifts in positions taken have been made by some of its members..." [Nagel, 1936, 216ff.]. And the leftist American student of philosophy concludes,

> "in the first place, the men with whom I have talked are impatient with philosophic systems built in the traditionally grand manner. Their preoccupation is philosophy as analysis ... The intellectual temper cultivated by these men is that of ethical and political neutrality within the domain of philosophic analysis proper, however much they may be moved by the moral and social chaos which threatens to swallow the few extant intellectual oases they stand.

> In the second place, as a consequence of this conception of the task of philosophy, concern with formulating the method of philosophic analysis dominates all these places...without 'dogmatism and intellectual intolerance'... In the third place, students whose primary interest is the history of ideas will find that, with some important exceptions, they will profit little from talking of these men ... In the fourth place, what pertains to a common doctrine, the men to whom I refer subscribe to a common-sense naturalism".

Nagel who attended Schlick's lectures commented as follows on the sociological background of Vienna,

> "that although I was in a city foundering economically, at a time when social reaction was in the saddle, the views presented so persuasively from the Katheder were a potent intellectual explosive. I wondered how much longer such doctrines would be tolerated in Vienna...
>
> Analytic philosophy has thus a double function: it provides quiet green pastures for intellectual analysis, wherein its practitioners can find refuge from a troubled world...; and it is also a keen, shining sword helping to dispel irrational beliefs and to make evident the structure of ideas... it aims to make as clear as possible what it is we really know." (Ibid.).

And with special focus on the Cambridge philosophy around Moore and Wittgenstein, he admits the significance of the latter, "in spite of the esoteric atmosphere which surrounds" him.

Let me return to the transfer and transformation of Central European philosophy of science to the Anglo-Saxon world, where Great Britain has been featured rather unjustified primarily as a transition country. It is true that Carnap, through his contacts with Charles Morris from 1934 on, gradually found entry into American universities, but it is important to note that his books had already been translated and read in England before then.

On the invitation of Susan Stebbing, Carnap came from Prague to London where he delivered three lectures at the University of London in October 1934. Here he came into contact with Russell, Ogden, Woodger, Braithwaite – and, significantly, with the young philosophy student Max Black. The latter wrote his PhD. thesis on "The Theories of Logical Positivism" under the influence of Moore and Ramsey.

In his introduction Black dealt with the origins of the Viennese Circle, its relations to Wittgenstein, and the central semantic notion of meaning. This looks like an anticipation of Ayer's best-selling book in 1936 on Logical Positivism: *Language, Truth and Logic.* And in Carnap's introductory notes we read already in a clear and distinct diction that "we are not a philosophical school and that we put forward no philosophical theses whatsoever ... for we pursue Logical Analysis, but no Philosophy" [Carnap, 1934, pp. 21 and 29].

The second booklet in this series *Philosophy and Logical Syntax*, with the content of the three mentioned lectures was the first popularization of Carnap's *Logical Syntax* period since the beginning of the 1930s in Great Britain:

> "My endeavour in these pages is to explain the main features of the method of philosophising which we, the Vienna Circle, use, and, by using try to develop further. It is the method of logical analysis of science, or more precisely, of the syntactical analysis of scientific language. Only the method itself is here directly dealt with; our special views, resulting from its use, appear rather in the form of examples (for instance our empiricist and anti-metaphysical position in the first chapter, our physicalist position in the last)." [Carnap, 1935, p. 7]

Max Black, born 1909 in Baku (Russia), had studied mathematics and philosophy in Cambridge, Göttingen and London, and published the book *The Nature of Mathematics* in 1933 (advised by Moore and Ramsey). Later on he emigrated to the United States in 1940 where he began his career, first at the University of Illinois and as of 1946 as Professor of Philosophy at Cornell University. He became a leading figure in (British-American) analytic philosophy and was an important mediator between Logical Empiricism and the British tradition in philosophy of science – apart from Ayer's popularization and idiosyncratic interpretation of the Vienna Circle from 1936 on. (cf. the co-authored book with Ernst Gombrich and Julian Hochberg on *Art, Perception and Reality* in 1972). This function can be detected in his article "Relations between Logical Positivism and the Cambridge School of Analysis" [1939/40]. In describing on the common ground between analytical and common sense philosophy in England associated with Moore and Russell in the Vienna tradition, Black explicitly refers to Wittgenstein's influence in the 1930s:

> "During the last eight years Wittgenstein's influence upon younger English philosophers has been comparable with that exerted by Morris. In this the Tractatus has played less part than his lectures ... and oral discussions based upon Wittgenstein's later and more radical views." [Black, 1938/39, p. 32]

And he notes that this influence could be more closely linked with Schlick and Waismann than with Carnap and Neurath. This corroborates the thesis that in England analytic philosophy (with the subfield of ordinary language philosophy) were better able to gain acceptance than the philosophy of science related to the *International Encyclopedia* project. This development is confirmed by Friedrich Waismann's work after his immigration to the UK in 1937. His influence was, compared to Wittgenstein, unspectacular but continuous, if one takes into account the publication of his oeuvre. (cf. Waismann's Ayer-critical *Principles of Linguistic Philosophy*, 1965).

Notwithstanding all differences, Black underlined the convergence of both movements at that time, still hoping that there would be further productive cooperation:

"... There should be room for further fruitful interchange of opinions between the two movements." [Black, 1938/39, 34].

On the basis of this description it is not surprising that we can reconstruct a bilateral (and intercontinental) exchange of ideas also on the level of institutions and periodicals: The name of a philosophical journal published since 1933 in Oxford (with Basil Blackwell) indicates the program as such: *Analysis,* ed. by A. E. Duncan-Jones with the cooperation of Susan Stebbing and G. Ryle, issued six times a year, was founded under the influence of Moore, Russell, and Wittgenstein, followed by an "Analysis Society" in 1936. Besides Alfred Ayer and Max Black, also Carnap, Hempel and Schlick contributed to the early issues. After the war (the journal was suspended 1940–1947) we find amongst the authors, e.g., Friedrich Waismann (with 6 articles on 'Analytic-Synthetic'), and Karl Popper (on the Mind-Body-Problem).

Incidentally, together with another international project, the still existing journal *Synthese* 1936ff. published in the Netherlands and the already mentioned *Philosophy of Science,* in addition to the 1940 established *Philosophy and Phenomenological Research,* formed an extended international forum for the communication between the poles of (language-critical) analytic philosophy and philosophy of science.

In the context of these activities Black — although always maintaining a critical distance (a "friendly critic" in his own words) — paved the way for greater receptivity of the Vienna Circle — influenced and inspired mainly by the already mentioned ("Lizzy") Susan Stebbing: she contributed significantly to the intellectual acculturation of Logical Empiricism in Great Britain, but because of her early death in 1943 she regrettably fell into oblivion. She studied at Girton College, the University of London, before her teaching period at King's College in London (1913–1915), Bedford College (1915–1920), University of London (1920–1924), where she became the first woman professor of philosophy in Great Britain (1933–1943). While her colleagues remember her as being a passionate teacher, her philosophical writings document a highly profound knowledge of the empiricist and analytical tradition of Continental and English thinking: as president of the "Aristotelian Society" in 1933 she reinforced her presence on an institutional level, since she was also acquainted with Russell, Moore and Whitehead. As a supporter of Carnap, Neurath and Popper, and given her friendly relations with Wittgenstein she played the role of a go-between and mover and shaker in analytic philosophy of her time, as became clear in the philosophical lecture she gave to the British Academy on *Logical Positivism and Analysis* [1933]. In this lecture, she investigated the language-critical approach of the Vienna Circle and the *Tractatus* and compared it with the English empiricist tradition (from Russell, Moore to Ramsey). She argued that while all philosophy is concerned with language, she was rather skeptical that all philosophical problems are linguistic ones. Neurath welcomed one of her last lectures on "Men and the Moral Principles" [1944, 18f.] as follows in his notes in the complimentary copy, where Stebbing states:

> "Moral philosophy, I repeat, is not a science", but "Whatever maybe the case with politicians making weekend speeches in time of war, philosophers cannot afford to ignore the conditions of the problems set by the situations in which we live."

This appeal is consonant with the defense of democracy during World War II in her last book *Ideals and Illusions* [1941], which obviously impressed again the exiled Neurath in Oxford besides their philosophical affinity.

Only on the basis of this scientific communication and relationships can we fully appreciate the specific contribution of Alfred Jules Ayer as *the* chief interpreter and protagonist of the Vienna Circle and Wittgenstein I in England, esp. with his book *Language, Truth and Logic* (1936a). His primarily anti-metaphysical position already became manifest with his appearance at the Paris Congress of 1935, where he refers in his paper on "The Analytic Movement in Contemporary Philosophy" [1936b] to the analogous movement in Vienna and England since the turn of the century. The success of his book also influenced all other networks: still in 1955, the 11th imprinting of the second, enlarged (and critically revised) edition of this bestseller appeared. The 8 chapters addressed the following issues: I. The Elimination of Metaphysics, II. The (new) Function of Philosophy, III. The Nature of Philosophical Analysis, VI. The A Priori, V. Truth and Probability, VI. Critique of Ethics and Theology, VII. The Self and the Common World, and VIII. Solutions of Outstanding Philosophical Disputes.

In the Preface to the first edition we read that "*the views which are put forward in this treatise derive from doctrines of Bertrand Russell and Wittgenstein, which are themselves the logical outcome of the empiricism of Berkeley and Hume ...*" employing a modified verification principle for empirical hypotheses [Ayer, 1955, p. 31]. And he goes on to contextualize these assertions to the effect that, philosophizing is an activity of analysis which is associated in England with the work of G. E. Moore and his disciples, but, 'the philosophers whom I am in the closest agreement are those who compose the 'Vienna Circle' ... and of these I owe most to Rudolf Carnap.' [ibid., p. 32]. Additionally, Ayer not only expressed his indebtness to Gilbert Ryle and Isaiah Berlin, but also alluded to philosophical differences. Ultimately, he says: "*we must recognise that it is necessary for a philosopher to become a scientist, ..., if he is to make any substantial contribution towards the growth of human knowledge.*" [ibid., p. 153].

Was this really a part of *The Revolution in Philosophy* [1956], as Ayer later maintained in a collection on Bradley, Frege, Moore, Russell, Wittgenstein and the Analysis of the Ordinary Language Philosophy with reference to the Vienna Circle? Although his judgment is apparently ambivalent (e.g., he saw the Vienna Circle as a movement, as a thing of the past, to a certain extent, but on the contrary, he declared many of its ideas to be living on). Ayer remained a critic of the late Wittgenstein — and, by the way, of Popper too.

In summary, we can say that there was a lively culture of scholarly dialogue of Central European and English philosophers — with a stronger focus on analysis, as compared to the turn from "Wissenschaftslogik" to philosophy of science in the USA. But there were also mutual contacts since about 1900, which cannot be separated from what has been referred to as the Anglo-Saxon "sea change" [Hughes, 1975] proper. What we have here is a dynamic network on different levels (like personal contacts, publications, societies, conferences and institutions) with distinct convergences and divergences of ideas and theories. Moreover, it is a network that reflected the intellectual preoccupation with several philosophical and methodological disputes between thinkers from different countries: from the Austro-German *Methodenstreit*, the Positivism disputes to the foundational debates in mathematics and logic since the 1920s. But the style and form of theorizing changed under different social conditions and a new intellectual setting in the immigration countries and triggered a self-organizing set of innovation and academic exchange.

This hypothesis could be exemplified by case studies on the Bloomsbury Group, Wittgenstein's Cambridge and Neurath's Oxford — and, last but not least, Hayek's and Popper's London, which I want to briefly describe in the subsequent passages.

2.4 On the Neurath Connection (Oxford)

A striking example for such overlapping networking is the Neurath connection in England: The main promoter and proponent of the Unity of Science movement from 1934 on since his exile in the Netherlands re-established in the few years 1940-45 in England good old contacts with a remarkable intellectual and practical manifestation — and one is inclined to ask in the sense of counterfactual history: What would have been, if Neurath had survived the World War II period?

First of all, in Oxford he initiated Central European disputes on plan vs. market or socialism vs. liberalism (with Hayek) and philosophical relativism vs. absolutism (with Popper) again in the new context of the envisioned liberated postwar society. Therefore their relationships, significantly emerged already in the Viennese years, can be described as more or less conflict-ridden communication between family resemblance and distance. Given the limited space, I will only allude to Neurath vs. Popper before turning to Hayek and Neurath (cf. [Stadler, 2001, chapters 10.3–10.5]):

Besides sharing a rejection of Platonic social philosophy (*Republic*), seen as a legitimation for authoritarian and totalitarian ideas, including the *Führerkult* (which according to Hayek's Plato interpretation was also part of a specific English controversy) there was the controversial encounter of both personalities that began in the early twenties. Neurath immediately criticized Popper after the publication of his *Logik der Forschung* (1934) accusing him of being an advocate of an absolutist "pseudorationalism" [Neurath, 1935] — as he, incidentally, also rejected the verification endorsed by the 'Wittgenstein camp'. The options of 'unity of science' or 'unity of method' (which was also rejected by Hayek) appeared as main alternative. But there are also uncontested familiarities: it is Popper's appeal to planning for institutions in his exile publications on the *Open Society* (1945) and

Poverty of Historicism (1944/45), which highlight the differences with Hayek — marginalized by Popper given Hayek's total opposition towards any form of planning theory and practice.

As regards the relation of Neurath and Hayek, we can detect a truly unbridgeable gap in political and economical matters, but a remarkable convergence in methodology and even in epistemology.

The adamant Scottish liberal Hayek, although in earlier times fond of Mach's epistemology and Schlick's *General Theory of Knowledge* (1918/1925), distanced himself – together with Ludwig von Mises and Karl Popper — from the so called "positivist economics", esp. from all variations of planned economy (in kind) from the 1920s on. All of this opposition that had originated in Vienna culminated in a short dispute in the 1940s in the common English exile (although I admit Hayek coming to London in 1931 — as Wittgenstein returning 1929 to Cambridge – was not a typical emigrant). Hayek's articles on "Scientism and the Study of Society" (1942–1944) in *Economica* and the subsequent publication of his *Road to Serfdom* (1944) was the starting point for a renewed *Methodenstreit*: two cultures (natural science vs. the humanities) were the main options in scientific enterprise: Neurath published a remarkably moderate review of *Road to Serfdom*, by showing that Logical Empiricism was providing a "through and through" pluralist view towards "Planning for Freedom" [Neurath, 1942]. But Neurath's personal annotations in his own copy of *Road to Serfdom* are much more critical: "*His technique: Overstate a case, create car(r)ricature of it, then fight it and then kill it is either German or immoral etc.*"

The Central European social reformer is fighting against the extreme liberal economist. Nevertheless it is remarkable how these two different former Austrian intellectuals entered into a controversy abroad — with the common background of the tension between liberalism (as laissez-faire capitalism) and socialism (social democratic position). In short: it is the option between a liberalist international like the "Mont Pelerin Society" or the so-called "third way" between communism and capitalism which Popper preferred because of his Viennese progressive social liberalism [Hacohen, 2000]. Although Neurath tried to initiate a continuous discussion with Hayek between Oxford and Cambridge (where the London School of Economics had its wartime address), the latter refused to enter into details, even if agreeing "entirely with what you say on Plato. He is certainly the arch-totalitarian." (Hayek to Neurath, Febr. 2, 1945). Hayek was busy lecturing abroad and only moved back to London later. He was convinced that he had already dealt exhaustively with the issue of "Scientism". Two weeks after Hayek's last hesitant letter, Neurath died unexpectedly of a heart attack on Dec. 22, 1945: the dialogue between the adherents of plan and market shimmered through in the ensuing Hayek-Popper communication.

Mainstream historiography obscures the difference of these "ambivalent brothers in mind": first, Poppers's insistence on the unity of method for natural and social sciences, second, his preference for a limited planning for institutions, third, his adherence to a socially oriented welfare economy in the tradition of Austrian social reform.

Probably because of his personal indebtedness to, and acquaintance with, Hayek who essentially mediated Popper‘s engagement at LSE, Popper himself played down his differences with Hayek's social philosophy in New Zealand. This conclusion can be drawn on the basis of the published comments on Hayek's "Scientism". And although both directly/indirectly argue against Neurath under the (Cold War) labels of "objectivism", "collectivism" and "historicism", there can be no doubt about the shortcomings of the equation "scientism = historicism". In this field we lack further studies on the renewed interaction and relation of the Austrian School and Vienna Circle after their emigration, taking into account the controversial issues in "Red Vienna" between the wars.

But there is another hidden story to be revealed: it is the life and work of Otto Neurath and Marie Reidemeister (Neurath), who in their second exile systematically continued scientific relations during the 1930s. The unpublished memories of Marie illustrate their untiring efforts for the cause of the Encyclopedic movement and enlightened adult education via his *Isotype*-movement (*I*nternational *S*ystem *o*f *Ty*pographic *P*icture *E*ducation), and also his initiatives to continue the housing and settlement movement in his Vienna days. Neurath renewed all contacts (with Friedrich Waismann, Rose Rand, Friedrich Hayek, Susan Stebbing) from the continent and the US, and for a short time pursued his academic ambitions with significant publications and research projects that remained uncompleted. (Nemeth/Stadler 1996). A further look at his exile research library (which is now deposited at the Vienna Circle Institute in Vienna), once again established in England 1940ff., signals the continuity to develop the interwar plans including his experience with the fascist decade: books on 'International Planning for Freedom', 'Visual Education' and on 'Persecution and Toleration' were again on the agenda. And Hitler-Germany (and postwar German) education focusing on Plato and Kant — a topic which continued to preoccupy all participants. We realize Neurath's cooperation in a newly founded Fabian Society, production of Isotype-films with Paul Rotha for the anti-Nazi education and *Isotype* for visualization of public affairs as a contribution to the fight against totalitarianism.

Neurath's lectures at Oxford University (where also Ernst Cassirer and Friedrich Waismann taught in 1941) are forgotten, but his work on visual education and on modern social and economic museums has been further developed and integrated in the standards of today's work: at the University of Reading (Department of Graphic Communication and Typography) and partly at the British Natural History Museum.

2.5 The Unended Poker Story: Wittgenstein vs. Popper in Cambridge

Hayek serves as a link to this much more investigated and contested research topic because it was mainly "Wittgensteinians" who provided us with an alternative account of the 'genius' and his impact in Cambridge philosophy [Monk, 1990;

Hacker, 1996].

Hayek, a distant relative (cousin) to Wittgenstein, wrote a to date unpublished "Biographical Sketch" based on Russell's letters to Wittgenstein. (According to Hayek Wittgenstein's heirs refused any publication). Here I only want to indicate the essentials of my interpretation of 'Wittgenstein's Cambridge':

First, I do not share the inclusion of Wittgenstein in the traditional forced migration movement — as presented by H. Stuart Hughes (1975) focusing on the philosophical prologue with the transformation of Wittgenstein in Vienna into Wittgenstein in Cambridge on the basis of British idiosyncracies. Wittgenstein was not a typical emigrant of the interwar period, although he never would have attained any adequate academic position in Austria for philosophical and so-called "racial" reasons.

Second, even in England Wittgenstein appeared more as an apolitical and ahistorical philosopher (cf. [Sluga, 1999]) than an immigrated intellectual. He refused to become involved in Austrian exile organizations (as is reported by Engelbert Broda and Joseph Peter Stern).

Third, the philosophical importance of Friedrich Waismann independent from the old Viennese connection (Schlick, Wittgenstein, Waismann) should be made clear. This may be confirmed by the fact that Wittgenstein did not continue in England his former fruitful communication with Waismann, and refused to reestablish contact with his former partner and interpreter to the Vienna Circle. While Karl Popper's remembrances of this tragic relation are probably somewhat exaggerated, they indicate a typical feature of Wittgenstein's life and work.

Fourth, Wittgenstein was not a solitary thinker, even in England; as in his Vienna days he was once again active in intellectual circles and networks.

This brings me to another postwar Austro-English success story together with a philosophical hoax which also has its roots in the mid-thirties: The young Karl Popper was already welcomed around the Vienna Circle as an enrichment of the theoretical discussion — even if criticized by Neurath as 'official opposition' to Logical Empiricism. Like so many other figures of the movement, and for the same reasons, Popper had no chance of getting an adequate position at an Austrian university.

Therefore, he looked for a position abroad, preferably in the Anglo-Saxon world. Fortunately, once again, Susan Stebbing came to the rescue: on her invitation Popper delivered two lectures at Bedford College in 1935 on Alfred Tarski, thereby — in his own words — triggering Joseph Henry Woodger's interest in the Polish logician. In 1935/36 papers on probability followed this first presentation at Imperial College and further ones in Cambridge (with Moore) and Oxford (with Ayer, Ryle and Berlin). The decisive appearance was no doubt his talk on "The Poverty of Historicism" in Hayek's seminar at the LSE (inter alia with Ernst Gombrich who became his life-long friend and supporter). (By the way, in Oxford he also met the Austrian physicist and Nobel Laureate Erwin Schrödinger). During one session of the "Aristotelian Society" in 1936, to which Popper had been invited by Ayer, Russell spoke on "The Limits of Empiricism" on the basis of an inductivist approach. He appealed to a principle of induction, followed, as could be expected, by a controversial discussion with the adamant English neo-empiricists. (This is expressed again in Ayer's critique of Popper's falsificationism, to be found in his 1982 *Philosophy in the Twentieth Century*). After these contacts Woodger proposed that Popper apply for a lectureship in New Zealand, whereas Hayek via the "Academic Assistance Council", envisioned a position at LSE. As is well known, Popper decided for New Zealand, and recommended Friedrich Waismann for England. Regrettably, in his autobiography, Popper forgot the decisive positive role of Felix Kaufmann in establishing his English connections [Popper, 1974].

After Popper's second return to London when he accepted a position at LSE in 1946, with the help, in particular, of Hayek and Gombrich, he delivered a lecture on a "philosophical puzzle", organized by the "Moral Sciences Club" in Cambridge. Wittgenstein was present at this lecture. It was true that Popper had criticized Wittgenstein's concept of philosophy as the action of clearing up statements, claiming that all philosophical problems are essentially linguistic ones in his writings before 1945. Thus, it cannot be a surprise that his lecture with the rhetorical title "Are there philosophical problems?" provoked Wittgenstein to storm out of the room angrily, after having threatened Popper with a poker. This is only one side of the unended "poker story", reported primarily by the *agent provocateur* himself [Edmonds and Eidinow, 2001].

Let me close the thematic circle: what is striking here is, first, the fact that this topic is an old, classic Viennese one, and second, this stage two of the dispute was contextualized with the English philosophical peculiarities and thus transformed, which is confirmed in Popper's own triumphant words in relation to another lecture in Oxford — with a striking reference to World War II, and the Cold War discourse:

> One of the things which in those days I found difficult to understand was the tendency of English philosophers to flirt with nonrealistic epistemologies: phenomenalism, positivism, Berkeleyan or Machian idealism ('neutral monism'), sensationalism, pragmatism — these playthings of philosophers were in those days still more popular than realism. After a cruel war lasting for six years this attitude was surprising, and I admit that I felt that it was a bit 'out of date' (to use a historicist phrase). Thus, being invited in 1946–47 to read a paper in Oxford, I read one under the title 'A Refutation of Phenomenalism, Positivism, Idealism, and Subjectivism'. In the discussion, the defence of the views which I had attacked was so feeble that it made little impression. However, the fruits of this victory (if any) were gathered by the philosophers of ordinary language, since language philosophy soon came to support common sense. [Popper, 1974, 99f]

As is well known, Popper attacked Wittgenstein for his alleged immoral behaviour towards his former adherent and close partner Waismann — a story which has to be investigated in full length elsewhere. Notwithstanding all these oral histories of the communication in exile and with Popper finally settling in England, these controversies between Neurath, Hayek, Popper and Wittgenstein with a Viennese background of the classical Vienna Circle seems to me highly significant for the dynamic (in content and form) of intellectual migration — which was only partly determined by the unique external events of fascism, Nazism and the Holocaust.

If this is true, we can obtain a broader, subtler understanding of this period; which could also shed some light on the present state of intellectual life in Europe.

2.6 Frank P. Ramsey: Between Wittgenstein and the Vienna Circle

Frank Plumpton Ramsey (born February 22, 1903 in Cambridge, England, died in London on the 19th of January 1930) was certainly one of the most important and promising philosophers of the 20th century. Only his early and unexpected death at the age of 27 probably prevented him from becoming one of the leading figures in the philosophy of science and analytic philosophy — perhaps at par with Ludwig Wittgenstein, his lifelong close friend and also intellectual adversary.

It is well known that in his short life Ramsey immensely enriched philosophy and science with some profound and highly topical findings: the gifted student at Trinity College, Fellow at King's College and Lecturer at Cambridge University at least influenced Wittgenstein, Russell and Keynes as well as the Vienna Circle with his contributions on the foundations of mathematics, logic, and economics. Especially his significance for philosophy with its focus on the notions of truth, decision theory, belief and probability is worth mentioning. The intellectual context of Ramsey's thinking can also be illustrated with the famous Bloomsbury Group [Hintikka and Puhl, 1995].

Especially the period Ramsey spent in Vienna in 1924 and his contacts with the mathematician Hans Hahn, the physicist Felix Ehrenhaft, inter alia, draws attention to Ramsey's connection with the early Vienna Circle [Stadler, 2001]. Already in 1929, Ramsey was listed in the manifesto of the Vienna Circle and given credit for attempting to further develop Russell's logicism and cited as an author related to the Vienna Circle. There are references to his articles on "Universals" (1925), "Foundations of Mathematics" (1926), and "Facts and Propositions" (1927). And the proceedings of the "First Meeting on the Epistemology of the Exact Sciences in Prague" (September 15–17, 1929), mention Ramsey as one of the "authors closely associated with the speakers and discussions", together with Albert Einstein, Kurt Gödel, Eino Kaila, Viktor Kraft, Karl Menger, Kurt Reidemeister, Bertrand Russell, Moritz Schlick and Ludwig Wittgenstein. (*Erkenntnis* 1/1930–31, pp. 311, 329). But looking at the earlier communication of Ramsey with Wittgenstein and the Vienna Circle these references are not really surprising: whereas it is rather well known that Ramsey visited Wittgenstein in 1923 and 1924, his communication with Schlick and his probable participation in the Schlick Circle have not been fully appreciated.

Carnap's notes on the discussion in the Schlick Circle include Ramsey's definition of identity, the foundations of mathematics and probability. With reference to

July 7, 1927 we can read: "Discussion by Carnap and Hahn about Carnap's arithmetic and Wittgenstein's objection to Ramsey's definition of identity" [Stadler, 2001, pp. 238f]. Accordingly, Carnap reported on an earlier discussion (June 20, 1927) in the Wittgenstein group with Schlick and Waismann, in which the great "genius" (= Wittgenstein) also objected to Ramsey's notion of identity. Precisely this issue was on the agenda again 4 years later when Wittgenstein met Schlick and Waismann alone. (December 9, 1931, Ibid., p.441). His lifelong dealings with Ramsey is documented later on in Carnap's *Philosophical Foundation of Physics* (1966) with its special focus on the Ramsey sentence.

Another reference is worth mentioning here: commenting retrospectively on his article "The Role of Uncertainty in Economics" (1934) the mathematician Karl Menger, member of the Vienna Circle and the founder of the famous "Mathematical Colloquium", recognised the relevance of Ramsey's paper "Truth and Probability" (1931) — unknown to him at the time — for his own research, although distancing his own contribution from this study [Menger, 1979, p. 260]:

> "But the von Neumann-Morgenstern axioms as well as Ramsey's were based on the traditional concept of mathematical expectation and on the assumption that a chance which offers a higher mathematical expectation is always preferred to one for which the mathematical expectation is smaller. My study was not".

In connection with his stay in Vienna, there is another fact of Ramsey's life that merits attention: he underwent a (supposedly successful) psychoanalytic therapy with the lay psychoanalyst and historian of literature *Theodor Reik* (1888–1969), who, by the way, also gave him a book by the theoretical physicist Hans Thirring.

Ramsey, who invited Schlick to the "Moral Sciences Club" in Cambridge, discussed his personal controversy with Wittgenstein, which was triggered by his article "The Foundation of Mathematics" (1925). His description is also confirmed by Wittgenstein's critical and ambivalent comments on Ramsey in his *Diaries* (April 26, 1930) [Wittgenstein, 1999, S.20f].

These contacts continued, and in one of his last letters before his death, Ramsey reported to Schlick on Wittgensteins's impact on his own philosophy (namely in the sense that it "quite destroyed my notions on the Foundations of Mathematics") as well as on Cambridge philosophy in general. (Ramsey to Schlick, December 10, year not indicated).

It is no coincidence that Black many years later described Ramsey in the *Encyclopedia of Philosophy* (ed. by Paul Edwards), in the following terms, as

> "one of the most brilliant men of his generation; his highly original papers on the foundation of mathematics, the nature of scientific theory, probability, and epistemology are still widely studied. He also wrote two studies in economics, the second of which was described by J.M. Keynes as 'one of the most remarkable contributions to mathematical economics ever made'. Ramsey's earlier work led to radical

> criticisms of A.N. Whitehead and Bertrand Russell's Principia Mathematica, some of which were incorporated in the second edition of the Principia. Ramsey was one of the first to expound the early teachings of Wittgenstein, by whom he was greatly influenced. In his last papers he was moving toward a modified and sophisticated pragmatism." [Black, 1939/40, Vol.7/8, p.65].

2.7 Between Unity and Disunity of Science: Family Resemblance and Distance with Otto Neurath, Friedrich A. von Hayek, and Karl Popper

Between the two World Wars, Austria was the center for two major scientific movements with international influence and recognition: on the one hand the Vienna Circle of Logical Empiricism in philosophy and methodology of science, and the Austrian School of Economics in social sciences, on the other. Although these two renowned traditions have of course been studied both historically and systematically, esp. over the last decade, we still lack research on both together, on their similarities and differences, mutual influences and interaction in the course of the development of science.

Thus the question arises (even it is only rhetorical) if we can view *The Philosophy of the Austrian School* [Cubeddu, 1993] or (the Philosophy) of the Vienna Circle [Stadler, 2001] as homogenous fields of research. Given the complexity and dynamic character of these so called "schools" it is necessary to examine their common ground, their similarities and differences with regard to socio-cultural background, theoretical development and methodological orientation.

At first sight they both are a manifestation of a typical "delayed" enlightenment. They share the fate of marginalization as a result of clerico-conservative, later on fascist and national-socialist repression. Conceived as an intellectual network it seems legitimate to describe both developments as interrelated scientific phenomena. Furthermore to unearth new aspects on the mutual interaction and influences between the two groups, which can be mirrored as overlapping intellectual circles.

Even if the opposition of Ludwig v. Mises and Friedrich August v. Hayek to Otto Neurath regarding the dualism of planned vs. market economy is well known, there remain many aspects of subtle common and differing features, whereas, on the other hand, the Hayek-Popper exchange of ideas cannot be characterized as an example of theoretical agreement. To show this, I will focus on the central methodological notion of "scientism". There was no consensus on this notion, as opposed to the critical view of Plato shared by these three thinkers.

The positions of mediating figures such as Felix Kaufmann and Richard von Mises show that this conflict was an immensely more complex debate. Furthermore this is documented by the strong influences on mathematical economy, esp. game and decision theory by the young Karl Menger's "Mathematisches Kolloquium" — even if we admit that this branch of the Austrian School only figured rather peripherally at that time.

The Vienna Circle undoubtedly contributed to the foundations of mathematics and logic and played a prominent role in the rise of John von Neumann's and Oskar Morgenstern's expanding economies and game theory, Abraham Wald's equilibrium theory, culminating in the formation of econometrics with Carnap's adherent Gerhard Tintner, who later applied Carnap's probability and logic.

A list of the publications relevant to (methodology of) economics and social science and the implicit/explicit treatment of the value problem in connection with the is/ought relation would already reveal the most neglected issues of research — apart from some recent studies (by Cubeddu [1993] and Robert Leonard [1998] on the "Menger-connection". In the research on the Vienna Circle we can detect only in the last few years guiding publications [Köhler and Leinfellner, 1997]). It is here that we can find the key concepts for a comparative study: normativity vs. descriptivity, theory vs. experience, reason and action, explanation vs. intuition (*Erklären* vs. *Verstehen*), foundations of natural and social sciences with psychology vs. *Logic of Scientific Discovery* (K. Popper). All these topics are based on an evolutionary and/or probabilistic approach.

In Karl Popper's sense we can state that the two main problems of epistemology, namely induction/deduction and the delineation of science from nonscience or metaphysics, form the heuristic and theoretical background for the theories of rationality and action addressed — as are also addressed in Schlick's "On the Foundation of Knowledge" [1934]. It's not a coincidence that still two decades later Fritz Machlup — as well as Morgenstern — was still reflecting on "The Problem of Verification in Economics" [1955] with reference to Felix Kaufmann's decision making "rules of procedure" in his *Methodology of Social Sciences* [1944].

Much less surprising is the revival of Hayekian cognitive science following the publication of *The Sensory Order* [1952], which has its roots in his early strong reception of Mach and Schlick, and which was (not officially) criticized by Popper because of its causal theory of mind [Birner, 1998]. This topic can also be seen as a sort of variation of the theme that is the all-embracing "Methodenstreit" since the beginning of the 20th century. And there remains the main question if there is a general "Theory of Valuation" [Dewey, 1939] that bridges the gap between theoretical concepts and empirical science or normative and descriptive aspects of human action. This would amount to the validity of the crucial meta-theoretical position called "methodological individualism" and part of the "Duhem-Neurath-Quine-thesis". The methodological tension of holism and individualism (with its inherent deontic logic) has overshadowed all substantial discussions in philosophy of science in general and specifically in the social sciences.

Since we can relate scientific world conceptions to all these methodological positions it also seems legitimate to reconstruct the background of *Weltanschaungen* for the (manifest) *Methodenstreit* in order to explore all the ideas underlying the so-called "progressive liberalism" (as attributed to Popper by [Hacohen, 1997]) and (Austro-Marxist) socialism very much inspired by the social reform concepts of Popper-Lynkeus combined with Mach's epistemology.

One misconception of Logical Empiricism to be destroyed is that of the disregard of ethics and value statements in the Vienna Circle. Thus Wolfgang Stegmüller, advancing Quine's influential "Two Dogmas of Empiricism" [1951] — namely the absolute distinction of analytic and synthetic statements and the empirical reduction of theoretical concepts — was critical of the replacement of ethics by meta-ethics (meta-ethical noncognitivism) as a consequence of the meaning criterion. On the contrary, the late Carnap himself (together with Richard Jeffrey) furthered the development "From Logical Empiricism to Radical Probabilism" [Jeffrey, 1993] through his probabilistic foundation of decision theory. But even before World War II, notwithstanding the verification or falsification criteria, different conceptions of ethical discourses were discussed — even if admittedly not at the center of Logical Empiricism. One may refer to the relevant publications of Felix Kaufmann — which also included purely economical writings — and in addition Viktor Kraft, Karl Menger, Richard von Mises, Otto Neurath, Josef Schächter, Moritz Schlick and Friedrich Waismann. Schlick for instance (*Problems of Ethics*, 1930) did consider ethics/aesthetics as philosophical and scientific sub-disciplines, presenting a naturalist ethics of imperative values in combination with an empirical psychological description of moral behavior and the meta-ethical analysis of concepts and statements. Viktor Kraft's *Foundations for a Scientific Analysis of Value* (first published in German 1937) also presented ethics as a scientific discipline, e.g. by analyzing value concepts normatively and factually, allowing value judgments to enter into relations of logical inference among each other and together with factual statements. Schlick's book — by the way — was an incentive for Karl Menger's *Moral, Wille, Weltgestaltung/ Morality, Decision and Social Organization* (1934). The central idea was the view that moral attitudes are based simply on decisions. Menger applied logico-mathematical thinking to human relations and associations resulting from diverse and even incompatible attitudes by completely avoiding evaluations. In this sense Menger deployed formal decision theory and a game-theoretic logic of groups — a kind of "socio-logic" — against the then predominant (Neo-Kantian) ethics of value and duty. This led to an empiristically "externalised ethics of decision". Menger's meta-theoretical "Principle of Tolerance" regarding the use of logics and scientific languages was one more cornerstone in the application of Logical Empiricism in the modern social sciences — far away from the dogmatic de-historization of the *Received View* of scientific theories between *Positivismusstreit* and analytic philosophy of science. This fits very well with Neurath's theory of social science best characterized by his boat metaphor directed against absolutist and dualist epistemologies — popularized in the Anglo-Saxon world by W. V. O. Quine:

> "Imagine sailors who, far out at sea, transform the shape of their clumsy vessel from a more circular to a more fishlike one. They make use of some drifting timber, besides the timber of the old structure, to modify the skeleton and the hull of their vessel. But they cannot put the ship in dock in order to start from scratch. During their work they stay on the old structure and deal with heavy gales and thundering

waves. In transforming their ship they take care that dangerous leakages do not occur. A new ship grows out of the old one, step by step — and while they are building, the sailors may already be thinking of a new structure, and they will not always agree with one another. The whole business will go on in a way we cannot even anticipate today. That is our fate." [Neurath, 1939, 47].

2.8 Neurath and Popper: Relativism vs. Absolutism

Popper was aware of Neurath's life and work after World War I, as he himself reported in his autobiographical remarks [1974]. He remembered Neurath's involvement in the Bavarian revolution [1919/20] in connection with a planned economy based on full socialization and with reference to the semi-socialization program of Josef Popper-Lynkeus. Popper was inclined to sympathize with the latter's (utopian) project. Apart from differences in personality and mentality, on the one hand, the Marxist dissenter and politically oriented encyclopedist, and the critical rationalist philosopher on the other, Popper accused Neurath of having succumbed to utopianism, historicism and scientism as represented by the Vienna Circle and the *Ernst Mach Society*.

In his own words, Popper was very flattered that Neurath published his criticism as "Pseudorationalism of Falsification" [1935] and was not unpleased ("nicht unzufrieden") with this honorable attack, but surprisingly he never replied in a systematic way. Maybe because of Neurath's critique (in light of his methodological holism): "the absolutism of falsification ...is in many ways a counterpart against the absolutism of verification which Popper attacks". Popper's attempt to characterize Neurath's *Empirische Soziologie* (1931) in the context of historical prophecy fails to assess the author's foundation of social science. From the outset Neurath remained very skeptical of explanations on the basis of one method and one image of science without pragmatically relativizing the field of "Prediction and Induction" (1946): "Unity of Science" as represented in the ambitious project of the *International Encyclopedia of Unified Science* <u>or</u> "Unity of Method" — by the way contrary to Hayek — as explicated from Popper's *Logic of Scientific Discovery* to the *Open Society* and the *Poverty of Historicism* seemed to be alternative approaches in the history and philosophy of science. In Popper's own words:

> "Neurath and I had disagreed deeply on many and important matters, historical, political, and philosophical; in fact on almost all matters which interested us both except one — the view that the theory of knowledge was important for an understanding of history and political problems." [Popper, 1974, p. 56]

The direct confrontation of both opponents in their still unpublished correspondence shows a high level discussion in philosophy of science. At the same time one might wonder whether Popper did exaggerate the real differences between him and the so-called "positivists" [ter Hark, 2004] — a designation which Neurath so strongly opposed as a cliché — and underestimate any form of scientific cooperation between the new "encyclopedists". In this connection it is, indeed, surprising that also Critical Rationalism can be — counter-intuitively — a suitable tool for a planning methodology as Andreas Faludi tried to show in his book [1986]. And it is this aspect — namely Popper's appeal to planning for freedom or institutions (in his *Open Society* and *Poverty of Historicism*) — which highlights the differences marginalized by Popper due to Hayek's total opposition towards each form of planning theory and practice.

2.9 Neurath and Hayek: The Unbridgeable Gap (in Economics)

Hayek felt himself from the beginning of the 1920s on to be Neurath's opponent regarding economics: inspired by the Carl Menger's "conception of the spontaneous generation of institutions", by Ludwig von Mises' *Gemeinwirtschaft* and by Popper's anti-inductivist *Logik der Forschung* [1935] he lobbied against the so called "positivist economics" — with Neurath (1919: *Through War Economy to Economy in Kind = Durch die Kriegswirtschaft zur Naturalwirtschaft*) as the most suitable target.

Formerly, Hayek was impressed by Mach and Moritz Schlick, before he then began to oppose the Vienna Circle because its social science was dominated by Neurath. The intellectual divorce centered around planning theory, economy in kind, and generally speaking on the concept of value, which Hayek — erroneously — missed in Neurath's social science. All these indirect oppositions culminated in a short dispute in the 1940s in English exile. This conflict sheds more light on the alternative conceptions of social science and its methodology:

Neurath took the initiative, as of 1945, in communicating by correspondence:

> "Enclosing I am sending you a review of your book. I tried to discover, what we have in common - unfortunately you are rather "absolute" in your EITHER-OR attitude. On Plato you may find some remarks in the article enclosed." (11. Jan., 1945).

He defended Logical Empiricism as "through and through" PLURALIST — whereas accusing Plato of a totalitarian practice DIRECTLY — but did not really impress the reserved Hayek. Although we can reread Neurath's review of *Road to Serfdom* in *The London Quaterly of World Affairs* [1945, 121f] as a surprisingly moderate assessment of this anti-totalitarian pamphlet, it did at the same time offer a sophisticated justification of a special sort of "Planning for Freedom" [Neurath, 1942]. The latter proposed several possible solutions with a skeptical empiricist approach to find a third way between market economics and fascism — with happiness and prosperity as guiding notions.

Much more instructive seem to be Neurath's notations in Hayek's book:

> "There is some danger that planning as a fashion [may] be used by totalitarian groups for weakening the democratic behavior, which implies — muddle. Democracy — muddle — and victory. But that is not

> the muddle of slums, distressed wars, depressions etc. but multiplicity of decisions, freedom of societies, local authorities. The fascists try to discredit muddle and to praise order, unification, subordination as such, otherwise they cannot run the show!!
>
> Therefore we need an analysis of planning with reckoning in kind plus muddle!
>
> That lacks — therefore danger."

Scientifically more relevant is the controversy over the concept of "scientism" between Hayek and Neurath, which later also included Popper. There, Hayek dealt extensively with the applicability of the methods of natural science and "social engineering" to the problems of man and society (as directed against Mannheim, Neurath and maybe also in some sense against Popper).

In his "Scientism and the Study of Society" [1942–1944] — re-published in *The Counter-revolution in Science* [1952] — Hayek condemned the appraisal of natural science methodology as the only "scientific method". This does not hold, because "facts" in the natural and social sciences are totally different: on the one hand causally explicable, on the other they are mere unobservable "opinions" of the actors producing their "objects". Common sense via analogy is the central key for understanding in social science. These essays distanced Hayek from all forms of "objectivism", "behaviorism", directly referring to Neurath's "physicalism", accusing him of supporting *in natura* calculation (instead of calculation in terms of price and value) and taking "naively for granted that what appears alike to us will also appear alike to other people" (p.35). What surprises us at first sight is the omission of a critique of language and the merging of the theoretical and meta-theoretical levels of speaking about the external world. Despite of all these misunderstandings regarding "methods of science", Neurath was ultimately willing to agree with Hayek's conclusion, quoting Morris R. Cohen that *"the great lesson of humility which science teaches us, that we can never be omnipotent or omniscient, is the same as that of all great religions: man is not and never will be the god before whom he must bow down."* (p.39).

Although Neurath tried to start a discussion in which he referred to theoretical contributions on social science, Hayek refused to enter into a detailed discussion. From Cambridge (where LSE had its wartime address) to Oxford, Hayek wrote that he was *"by no means so much opposed to 'Logical Positivism' as you appear to think and with some members of your former group, particularly with Karl Popper, I find myself in complete agreement"* and — alluding to physicalism and *in natura* calculation — he continued to articulate his skeptical position towards Neurath, at the same time agreeing *"entirely with what you say on Plato. He certainly was the arch-totalitarian"*. (Hayek to Neurath, February 2, 1945).

2.10 Popper and Hayek: The Ambivalent Brothers in Mind

The deductive-hypothetical methodology was essentially directed against inductivism and/or apriorism, which were characteristic for the *Philosophy of the Austrian School* [Cubbeddu, 1993]. This move was promoted by Neurath, although there was no evidence of holism, collectivism or inductivism with any kind of prophecy in social science method. By the same token, we can reconstruct in the Vienna Circle's social philosophy with Kaufmann, Menger, Neurath and Richard von Mises (cautious) conceptions of empiricist methodologies with conventionalist tools. But these issues have not been investigated sufficiently so that we lack studies on the Hayek-Popper interaction. Already the critical review of Hayek's *Counter-Revolution* by Ernest Nagel (*Journal of Philosophy* 1952) should have been given greater recognition and should have provoked further discussions. (By the way, it is noteworthy that a structurally similar criticism has been raised by the Frankfurt School in the context of the "Positivismusstreit" from the thirties to the sixties — but this is a completely different story (cf. [Dahms, 1994; Uebel, 2000]).

It can be assumed that the motivational background of the scientism-controversy lies in the topicality of socialist planning theory between the wars (cf. Hayek's edition of *Collectivist Planning* in 1935). For insiders therefore the account of Richard von Mises seems adequate and more representative when he remarked in his book *Positivism* [1951] already finished in 1939 — a well-balanced re-evaluation of the Vienna Circle story from Mach to the high tide of Logical Empiricism — in his chapter on social sciences that

> "... neither the practical impossibility of experiments in the narrower sense nor the comparatively limited application of meta-mathematical methods is a specific feature of this field." (p.246)

Relating to classical economics he alludes to the shortcomings of terminology implying "eternal laws". In agreeing with Felix Kaufmann's *Methodenlehre der Sozialwissenschaften* [1936], Richard v. Mises acknowledged promising starting points for a rational treatment of economic problems in the theory of marginal utility, specifically in von Neumann's "economic games". Neurath's *Empirische Soziologie* [1931] — an alternative to the "polemic of neoliberalism against collectivist theories of economics" of Ludwig v. Mises' *Human Action* [1949] and Hayek's *Collectivist Economic Planning* [1935] — is presented by R. v. Mises as a sociology, which can not be separated from history. This account is in accordance with the prominent role of the Karl Menger jr. and his "Mathematical Colloquium" for the social sciences in interwar Vienna.

Already Morgenstern's article "Logistik und die Sozialwissenschaften" (*Zeitschrift für Nationalökonomie*, 1936) directed attention to the rich potential of the "new logic" (Karl Menger and Kurt Gödel) or "Logistics" (Russell and Whitehead) for economic research. There, Morgenstern explicitly endorsed the theory of types (Russell), axiomatics (Hilbert) as well as the use of an exact scientific language, of a so-called *Wissenschaftslogik* (logic of science) in Carnap's sense. He concludes his article by referring to the relevance of these methods for the social sciences as well, esp. for theoretical economics and political economy. To this end, he summed up the main ideas of Karl Menger's book *Morality, Decision and Social Organization* [1934]. Thus it is not surprising that we find Morgenstern's later collaborator John von Neumann participating at the Vienna Circle congress in Königsberg [1930] and in Karl Menger's "Mathematical Colloquium" in the twenties and thirties. Here we find some intellectual roots (issues such as experience and rationality, chance and determinism) for today's decision and game theories — extending to John Harsanyi's work on social theory as well as ethics.

The Vienna Circle's contribution to the foundation of probability calculation and theory (with Richard von Mises, but also Carnap's later inductive logic) and subsequent controversies between Karl Popper and Hans Reichenbach also inspired Abraham Wald's mathematical support of Richard von Mises' concept of probability. The latter also furthered mathematical economics by improving the equations of Walras and Cassel. Together with Morgenstern's first input-output model [1937] and von Neumann's equilibrium theory for expanding economies, the relevance of the "Mathematical Colloquium" is documented [Dierker and Sigmund, 1998]. This is also due to the fact that Menger himself was concerned with methodology of social science and "The Role of Uncertainty in Economics" as one of his articles in 1934 is entitled. And we fully agree with a contextualized analysis of Menger's significance for interwar social science [Leonard, 1998/99].

We should also add that Popper's participation in this context of mathematical economics (Hahn, Menger, Morgenstern, Tarski) is much stronger than in the paradigm of social science after 1945 as endorsed by Hayek, who still in 1937 ("Economics and Knowledge") was strongly interested in the mathematical foundations of social science. By the way, the evolutionary view of science can be traced back to the work of Mach and Boltzmann, but also to Popper's later position of *Objective Knowledge. An Evolutionary Approach* [1972].

All these influences are indirectly related to the above-mentioned dispute on "scientism", in which the central question was raised as to whether philosophy or philosophical foundation is needed for the methodology of the humanities and social sciences. Or alluding to the subtitle of Sorell's book *Scientism* [1991], one might ask whether scientism is the regrettable "infatuation of philosophy with science." At least it conveys the options of one, two or more scientific cultures with the inherent utopia of an unified methodology and theory or a science of man and nature.

For a better understanding of the lasting *Methodenstreit* in the 1930s and 1940s represented by the triangle Hayek-Popper-Neurath we first have to reconstruct the discussions in their socio-historical context, second to examine unpublished sources, third to distance ourselves from clichés about schools of thought and, finally to confront these results with today's research. In doing so we could fully appreciate the historical background together with internal theory dynamics without producing myths of partisanship. This would provide a rational option, namely a pluralist way for positioning this unsolved debate in an evolutionary context of theoretical fields.

3 VIENNA, PARIS AND THE "FRENCH CONNECTION": CONVENTIONALISM

In the last decade, a number of new publications on the transatlantic exchange of ideas have been presented by international scholars studying the Vienna Circle in the English-speaking countries.[1] Now it is high time to focus on the neglected "French Connection" in the philosophy of science.[2] That this connection has been somewhat neglected until now is all the more remarkable since we have known about the existence of close ties between Viennese and the Parisian intellectuals since the Fin de Siècle. Their exchange of ideas — studied by Ernst Mach — is most evident in the strong reception of Henri Poincaré and Pierre Duhem in the so-called "First Vienna Circle". This bilateral development in the Enlightenment discourse of the modern theory of science was already described by Philipp Frank in his book *Modern Science and its Philosophy* [1949], in which he underlined the threefold roots of logical empiricism in a more modern guise, making reference to English empiricism, French rationalism and American (neo-)pragmatism.[3]

What then evolved was, more specifically, a synthesis of empiricism and symbolic logic, with the Machian theory of science being refined by French conventionalism, along with an attempt to counter Lenin's critique of "empirico-criticism". Here Abel Rey's book *La théorie physique chez les physiciens contemporains* [1907] also sought to overcome mechanistic physics. It was ultimately Poincaré who tried to mediate between empirical description and analytic axiomatics of scientific terminology[4]:

According to Mach, the general principles of science are abbreviated economic descriptions of observed facts. For Poincaré, they are free creations of the human mind which say absolutely nothing about observed facts. The attempt to integrate both concepts in a coherent system was the origin of what was later to become known as logical empiricism.

This goal was attained with the help of Hilbert's axiomatics of geometry as a conventionalist system of "implicit definitions". This way Mach's philosophy could be integrated in the "new positivism" espoused by Henri Poincaré, Abel Rey and Pierre Duhem. The link between the new positivism and the old teachings of Kant and Comte consists in the demand that all abstract expressions of science — such as power, energy, mass — be interpreted as sense observations.[5]

As early as 1907, Pierre Duhem wrote the following in *La théorie physique, son objet et sa structure* (Aim and Structure of Physics), striking a similar chord as Mach :

---

[1]Cf. the most recent publications [Giere and Richardson, 1996; Hardcastle and Richardson, 2004.

[2]Anastasios Brenner, "The French Connection Conventionalism and the Vienna Circle", in: Michael Heidelberger / Friedrich Stadler (eds.), *History of Philosophy of Science. New Trends and Perspectives.* Dordrecht-Boston-London: Kluwer 2002, pp. 277–286.

[3]Philipp Frank, [1949].

[4]Frank, "Der historische Hintergrund", in: *ibid.,* p. 256.

[5]Frank, Ibid., p. 258f.

> "A physical theory is not an explanation. It is a system of mathematical propositions which can be derived from a small number of principles that serve to precisely depict a coherent group of experimental laws in a both simple and complete way."

This is followed by a crucial insight for the encyclopedia project: "The *experimentum crucis* is impossible in physics".[6] In spite of Duhem's metaphysical leanings his teachings became a framework of reference for further discussions between science and religion and, more generally, between science and ideologies. Reflecting further reciprocal influences, Frank compared Louis Rougier's publications with those of Moritz Schlick:

> "He proceeded from Poincaré, trying to incorporate Einstein in the 'new positivism' and wrote the best comprehensive critique of school philosophy ... the 'paralogisms of rationalism'." [Paris: Alcan 1920][7]

The physicist Marcel Boll translated writings by Rudolf Carnap, Hans Reichenbach, Moritz Schlick and Philipp Frank into French. The original influence exerted by Duhem was now to be reversed:

> "The French general Vouillemin (cf. C.E. Vouillemin, La logique de la science et l'Ecole de Vienne (Paris: Hermann 1935) recommended our group since we replaced the spelling "Science" by the more modest "science". ... The French neo-Thomists ... saw in logical positivism the destroyer of idealist and materialist metaphysics which for them were the most dangerous enemies of Thomism. To organize this international cooperation, a preliminary conference was held in Prague in 1934, in which Charles Morris and L. Rougier participated. The cornerstone was thus laid for the annual international congresses on the "Unity of Science."[8]

social science

With the heyday of the Vienna Circle in the interwar period, these European and transatlantic exchanges were increasingly consolidated while at the same time intellectual life in Germany and Austria had been disintegrating since 1930. Most significantly, however, direct recourse was taken both theoretically and practically to the French Encyclopédie of the 18th century in connection with the *International Encyclopedia of Unified Science* of the Logical Empiricists. Here it was mainly Otto Neurath who untiringly drew attention to the French intellectual precursors of his Unity of Science movement and was able to effectively implement this intellectual exchange before the outbreak of World War 2 at two international congresses in Paris (1935 and 1937) as late-enlightenment collective projects.[9] This comes as no

[6] Ibid., p. 259.

[7] Ibid., p. 291.

[8] Ibid., p. 291f.

[9] Stadler, *The Vienna Circle. Studies in the Origins, Development, and Influence of Logical Empiricism.* Vienna–New York: Springer 2001, pp. 363ff. and 377ff.

real surprise in view of the references to Comte, Poincaré and Duhem as precursors of the "scientific world view" in the programmatic manifesto of the Vienna Circle [1929].[10]

The "1st Congress for the Unity of Science in Paris in 1935", the "Congrès International de Philosophie Scientifique" held at the Sorbonne September 16-21, marked the first highpoint of the new philosophy of science of the Vienna Circle in exile. Already in late 1933 Neurath had conducted preliminary negotiations in Paris with Marcel Boll and Louis Rougier, still as representatives of the "Verein Ernst Mach" which disbanded in Vienna in 1934. These talks were then continued at a preliminary conference in Prague in 1934. Neurath's summary of the Paris conference reads like an extremely optimistic prognosis of the future of the "Republic of Scholars of the Logical Empiricism" and the "philosophique scientifique".

The first of the international congresses for the unity of science... was a success for logical empiricism vis-à-vis a larger public. The title "philosophie scientifique", which is so popular in France, aroused interest. The press constantly reported on the congress. Newspapers and journals dealt with it in sketches and interviews. This was all the more remarkable in view of the fact that, as Rougier and Russell had underlined in their introductory words, it was a conference whose task was to focus on a science without emotions. Some 170 persons from more than twenty countries had appeared and had shown a high degree of willingness to commit themselves to continuous cooperation. With their addresses at the opening of the congress in the rooms of the Institute for Intellectual Cooperation Rougier, Russell, Enriques, Frank, Reichenbach, Ajdukiewicz, Morris generated a living impression that there was such a thing as a republic of scholars of logical empiricism.[11]

The French institutions which co-organized the congress included the "L'institut International de Coopération Intellectuelle", the "Comité d'Organisation de L'Encyclopédie Francaise", the "Cité des Sciences", the "Institut d'Histoire des Sciences et des Techniques" as well as the "Centre International de Synthèse". The congress was documented in eight journals in the series "Actualités scientifiques et industrielles" published by the Parisian publisher Hermann & co (1936) with a number of French contributions. Bertrand Russell, who gave an appraisal of Frege in German in his opening address, remembered in retrospect a manifestation of rational-empirical thought in the tradition of Leibniz.[12] "The Congress of Scientific Philosophy in Paris in September 1935, was a remarkable occasion, and, for lovers of rationality a very encouraging one..." Neurath seconded this, arguing that "the individual sciences (should be) arranged next to each other by directly showing concrete relations and not indirectly by referring all of them to a common blurred conceptual system."[13]

The Congress unanimously committed itself to supporting the project of the

---

[10] *Wissenschaftliche Weltauffassung. Der Wiener Kreis*. Ed. Verein Ernst Mach. Vienna: Artur Wolf Verlag 1929. Reprint in: Fischer (ed.), loc.cit., p. 125 – 171.

[11] *Erkenntnis* 5, 1935, p. 377.

[12] Bertrand Russell, in: *Actes du Congrès International de Philosophie Scientifique*. Sorbonne, Paris 1935. Paris : Hermmann & co. 1936, p. 10.

[13] *Erkenntnis* 5, 1935, p. 381.

Encyclopedia of Unified Science, which had been organized by the Mundaneum Institute run by Neurath in The Hague. The committee of 37 included the French scholars Marcel Boll, H. Bonnet, E. Cartan, Maurice Frechet, J. Hadamard, P. Janet, A. Lalande, P. Langevin, C. Nicolle, Perrin, A. Rey and L. Rougier.

This event, which was also viewed as a testimony of the anti-fascist intellectuals, as can be seen in the interest shown by Robert Musil, Walter Benjamin and Bert Brecht, ultimately formed the basis for a stronger transcontinental development towards the cooperation of the German-, English- and French-speaking community of scholars who were primarily promoted by Neurath from his Dutch exile.[14]

The second round of the encyclopaedic renaissance was planned at the end of July 1937 to also take place in Paris as the "Third International Congress for the Unity of Science", after the organisational committee (Carnap, Frank, Joergensen, Morris, Neurath, Rougier) succeeded in obtaining a publisher's contract with the University of Chicago Press for the first two volumes of the "International Encyclopedia of Unified Science" (IEUS).

Moreover, a separate section on the "Unity of Science" (L'Unité de la Science: la Méthode et les méthodes) was organised in connection with the contemporaneous "Ninth International Congress of Philosophy" (Neuvième Congrès International de Philosophie — Congrès Descartes). Notwithstanding the theoretical differences on the conception of the "New Encyclopedia" between Carnap and Neurath (in particular on the notion of truth and probability), Neurath presented modern empiricism there as a type of heuristic puzzle aiming at a "mosaic of the sciences" in the following way[15]:

> "We can start out from the 'encyclopedia' as our model, and now observe how much we can achieve by a way of interconnection and logical construction and elimination of contradictions and unclarities. The synopsis of logical empiricism will be then the order of the day."

The main goal was thus to show "in addition to the existing great encyclopedias the logical framework of modern science"[16] with the construction of a sort of onion around a core consisting of 2 volumes with 20 introductory monographs, bringing forth a further 260 monographs, of which only 19 monographs were to appear in total due to the war."[17]

The complete realization of this project would have yielded 26 volumes with 260 monographs in English and French, supplemented by a ten-volume picture statistical "visual thesaurus" with global overviews in the spirit of Diderot and d'Alembert. The influence of history and sociology on the philosophy of science

[14] Cf. also Antonia Soulez. "The Vienna Circle in France", in: Friedrich Stadler (ed.), *Scientific Philosophy: Origins and Developments*. Dordrecht-Boston-London: Kluwer 1993, pp. 95 -112.

[15] Otto Neurath, "The New encyclopedia', in *Unified Science*, ed by Brian McGuinness. Dordrecht: Reidel, 1987, p. 136f.

[16] Ibid.

[17] Otto Neurath / Rudolf Carnap / Charles Morris (eds.), *Foundations of the Unity of Science. Toward an International Encyclopedia of Unified Science*. 2 vols. Chicago and London: The University of Chicago Press, 1971.

was also a sort of "science in context" meant to circumvent a strongly formalistic "scientism".

This encyclopedism was not intended to provide an absolute foundation of epistemology or "system" of sciences (neither with verification nor falsification as methodical instruments) but was to be more based on a broad everyday experience as the point of departure against the backdrop of uncertainty and indeterminacy, namely with the

> "Basic idea that one does not have any solid foundation, any system, that one has to keep trying on the basis of research and can experience the most unexpected surprises on later verification of many basic views that are used, is characteristic of the outlook that might be described as "encyclopedism"... As empiricists we will always proceed from our everyday expressions, and as empiricists we will use them again and again to verify our theories and hypotheses. These broad propositions with their many indeterminacies are the point of departure and the final point of our science."[18]

Now the question arises as to why it came to the rupture of this productive Austro-French cooperation and why this exchange was forgotten after this relatively successful history. Here I can only touch upon some of the reasons:

1. World War 2 destroyed a Central European late-Enlightenment culture of science, in particular in "Red Vienna".[19]

2. The ideologization in the wake of the second positivism debate (Horkheimer vs. Neurath) and the central focus of the project after 1945 being Louis Rougier who was a controversial figure in France prevented a reintegration in the community of scholars.[20]

3. Emigration, exile and the transfer of science to the Anglo-American world of scholars and the prevention of the return of ideas, reinforced by the third positivism debate within the context of the Cold War and the dominant *Dialectic of Enlightenment* (Horkheimer/Adorno) as well as the Marxist and the structuralist philosophers reinforced this rupture after 1938.

4. The preference for "German philosophy" of idealism and existentialism of the post-war years with a cliché of "positivism" and the belated research of the buried tradition of the philosophy of science since the turn of the century in France contributed to the rupture of a flourishing bilateral communication.

5. The integration of the 2nd Republic of Austria in the intellectual life of the West with a focus on the Anglo-American intellectual world additionally marginalized the "French Connection" after World War 2.

[18] Neurath, loc. cit., p. 213.
[19] *Wien und der Wiener Kreis. Orte einer unvollendeten Moderne. Ein Begleitbuch.* Ed. Volker Thurm-Nemeth and Elisabeth Nemeth, Vienna: WUV-Verlag 2003.
[20] Cf. Hans-Joachim Dahms, *Positivismusstreit*. Frankfurt/M.: Suhrkamp 1994.

An attempt has been made to help reduce the lacunae in the research and to study the common intellectual past with a view on its innovative potential for today's research.[21]

4 THE *WIENER KREIS* IN AMERICA: LOGICAL EMPIRICISM AND (NEO-)PRAGMATISM

4.1 Herbert Feigl and the Minnesota Center for the Philosophy of Science (MCPS) in Minneapolis

As the earliest immigrant of the Vienna Circle and a participant of the Unity of Science Movement at Harvard associated with Philipp Frank and, later, at the Boston Colloquium organized by Robert S. Cohen, *Herbert Feigl* (1902–1988) played a pivotal role in the transfer and development of "logical positivism" in the United States. With his research and teaching activities at the University of Iowa (1931–1940) and at the University of Minnesota in Minneapolis (from 1940), as well as his many visiting professorships on the East and West Coast, his functions as president of the "American Philosophical Association" and vice-president of the "American Academy of Arts and Science", Feigl became one of the most influential figures of the second generation of the Vienna Circle in the United States. The *Minnesota Center for the Philosophy of Science (MCPS),* was founded in 1953 by Feigl, has published 18 volumes of the *Minnesota Studies in the Philosophy of Science* (MSPS) since 1956 and became a sort of training center for the history and philosophy of science. It is therefore all the more surprising that there has been hardly any research on the life, work and reception of this original thinker who has been associated with the Vienna, Harvard and Minneapolis circles. We only have the scanty autobiographical fragments and indirect references to publications to go by, which give us some idea of Feigl's great transatlantic impact [Feigl, 1981; Haller, 2003].

In his own reminiscences, Feigl first speaks of his revered teacher and founder of the Vienna Circle, Moritz Schlick:

> "Several members of the Circle had a reading knowledge of English, but Schlick, whose wife was American, spoke English perfectly. Some of the conversations at Schlick's house were in English, notably with such visitors as Roger Money-Kyrle but occasionally with Wittgenstein who was also fluent in English. Schlick was the first of our group to

[21]Two examples: The Moritz Schlick edition project at the Institute Vienna Circle: `http://www.univie.ac.at/Schlick-Projekt/` as well as the foundation of the "Austrian-French Society for Cultural and Scientific Cooperation // Societé franco-autrichienne pour la cooperation culturelle et scientifique". These activities follow the tradition of the Austrian (late) Enlightenment which has been marginalized in research. Cf. on this: Kurt Blaukopf, "Kunstforschung als exakte Wissenschaft. Von Diderot zur Enzyklopädie des Wiener Kreises", in: Friedrich Stadler (ed.), *Elemente moderner Wissenschaftstheorie. Zur Interaktion von Philosophie, Geschichte und Theorie der Wissenschaften*. Vienna-New York: Springer 2000, pp. 177-211.

> be invited to the United States. ... Schlick enjoyed his sojourn at Stanford, made many friends, and was promptly invited to another visiting professorship, this time (in 1931) to the University of California at Berkeley. Thus it came about that Schlick was the first to spread the Vienna 'gospel' (with a strong emphasis on Wittgenstein's ideas) in America. My own first journey to the United States occurred in September 1930, when I was fortunate to obtain an International Rockefeller Research Fellowship. This allowed me to work at Harvard University for about nine months." [Feigl, 1968, p. 643]

Feigl's own transit was finally made possible through his acquaintance with two American philosophers — Dickinson S. Miller and Charles A. Strong — as well as through his contact with the American student Albert Blumberg, who had written his dissertation under Schlick. All these contacts led Feigl to the conclusion that:

> "'Over there', I felt was a Zeitgeist thoroughly congenial to our Viennese position. It was also in 1929 that, I think through Blumberg's suggestion we became acquainted with Percy W. Bridgman's *Logic of Modern Physics* (1927). Bridgman's operational analysis of the meaning of physical concepts was especially close to the positivistic view of Carnap, Frank, and von Mises, and even to certain strands of Wittgenstein's thought." (Ibid., 645)

Feigl in his autobiographical account [1968] reported just as enthusiastically on his first impressions of his New World contacts in Harvard (including C. I. Lewis, Henry Sheffer, A. N. Whitehead, Susanne K. Langer, Paul Weiss, W. V. O. Quine — and also, once again, Karl Menger) as he had on his Bauhaus experience in Dessau. After his article "Logical Positivism", written together with Blumberg in 1931, Feigl saw a debate unfold which was going to last twenty years — a sort of "succés de scandale" in philosophy. (ibid., 647). As one of the first authors in the *Philosophy of Science,* "a periodical, for whose initiation I was in small part responsible" (ibid.), Feigl was already a part of the relevant discussions at a very early stage. Through Felix Kaufmann's intervention, Feigl was also able to teach at the "New School for Social Research", which strengthened the New York philosophy of science scene to which Carl G. Hempel also belonged. Communication also flourished on the West Coast. Feigl's contacts with the philosophers there — Hans Reichenbach, W. R. Dennes, Paul Marhenke, David Rynin and even Else Frenkel-Brunswik and Egon Brunswik (with a Unity of Science meeting in Berkeley in 1953) — made the *Rise of Scientific Philosophy* (Reichenbach 1951) a country-wide issue.

In Iowa, Feigl offered a course on philosophy of science for the first time. In some of the leading anthologies, e.g., *Philosophical Analysis* (ed. 1949) and *Readings in the Philosophy of Science* (ed. 1953), he prepared the ground for the reception of these ideas. The journal *Philosophical Studies,* founded as a counterpart to the British *Analysis,* was a successful parallel initiative, which is still being continued today. A private foundation enabled the MCPS to be established:

> "For a few years in the late forties and early fifties, Sellars and I, together with May Brodbeck, John Hospers, Paul Meehl, and D.B. Terrell, made up a discussion group in which occasionally visitors from other universities would participate. Gradually we came to think about organizing a more official center for research in the philosophy of science. Encouraged by the generous financial support of Louis W. and Maud Hill Family Foundation in St. Paul, the Minnesota Center for the Philosophy of Science was established in 1953. During the first few years the local staff members were Paul E. Meehl ..., Wilfrid Sellars ... and Michael Scriven ... In the fourteen years of its activities, the Center has enjoyed visits of various durations by many outstanding American, European, and Australian and New Zealand scholars. Our major publications (Minnesota Studies in the Philosophy of Science and Current Issues in the Philosophy of Science) have aroused considerable interest. Several younger generation philosophers have been our visitors, among whom have been Scriven, Adolf Grünbaum (Pittsburgh), Hilary Putnam (Harvard), N.R. Hanson (Yale), Wesley Salmon (Indiana), Karl R. Popper (London), Paul Feyerabend (Berkeley), Bruce Aunne (University of Massachusetts), Henryk Mehlberg (Chicago), George Schlesinger (Australia, now North Carolina) and Arthur Pap (Yale)." [Feigl, 1968, 664f]

This account underplays the important role of the MCPS, founded as Research Department at the College of Liberal Arts of the University of Minnesota, in the propagation of the philosophy of science, which Feigl only alludes to when referring to the influence of similar centers at Indiana University or the University of Pittsburgh. This, combined with the teaching activities at these universities exerted an influence on a number of student generations. With regard to the influence on Austrian thinkers, we should mention the stimulating activities of Arthur Pap and Paul Feyerabend who, together with Grover Maxwell, co-edited the only Feigl Festschrift to date, which appeared in 1966: *Mind, Matter and Method. Essays in Philosophy of Science in Honor of Herbert Feigl.*

There, in his "Biographical Sketch", we find one of the few tributes to Feigl and his center. Feyerabend had met Feigl for the first time in a Vienna coffee house in 1954 (at that time he was an assistant of Arthur Pap). This encounter was seen by Viktor Kraft's group as being a particularly enriching one. Feyerabend continued to praise Feigl's style of philosophizing, which, at the MCPS, reflected the high level of discussion. Referring to the internal life of the center, Feyerabend wrote the following:

> "The atmosphere at the Center, and especially Feigl's own attitude, his humor, his eagerness to advance philosophy and to get at least a glimpse at the truth, and his quite incredible modesty, made impossible from the very beginning that subjective tension that occasionally accompanies debate and that is liable to turn individual contributions

> into proclamations of faith rather than into answers to the question chosen. The critical attitude was not absent; on the contrary, one now felt free to voice basic disagreement in clear, sharp, straightforward fashion. The discussions were, and still are, in many respects similar to the earlier discussions in the Vienna Circle. The differences are that things are seen now to be much more complex than was originally thought and that there is much less confidence that a single, comprehensive empirical philosophy might one day emerge." [Feyerabend, 1966, p. 9]

The discussions that took place in this atmosphere were thus mainly devoted to the analysis of scientific theories with a strong focus on current research. In an epistemological sense, Feigl's personal preference for a critical realism found acceptance, resulting in a sort of "anti-Copenhagen mood" which prevented metaphysics from appearing obsolete a priori. (A similar development could be detected in the Harvard group associated with Frank). The list of participants at the MCSP is certainly impressive, since it practically covered the entire horizon of philosophy of science. In the publications named this broad spectrum is largely documented in the individual contributions. Feigl describes his acquaintance with the philosophical rebel in the following words:

> "I met Feyerabend on my first visit to Vienna after the war (my last previous visit was in 1935). This was in the summer of 1954 when Arthur Pap was a visiting professor at the University of Vienna. Feyerabend had been working as an assistant to Pap. Immediately, during my first conversation with Feyerabend, I recognized his competence and brilliance. He is, perhaps, the most unorthodox philosopher of science I have ever known. We have often discussed our differences publicly. Although the audiences usually sided with my more conservative views, it may well be that Feyerabend is right, and I am wrong." [Feigl, 1968, p. 668]

Against this background it is not surprising that Feyerabend published one of his first critical studies in the fourth volume in 1970 of the MCPS "Against Method: Outline of an Anarchistic Theory of Knowledge" which was a highly controversial attack on of the standard version of the philosophy of science. As opposed to the institutions in Harvard and Boston, the MCPS presented an almost complementary trend to psychology and the social sciences along with a fundamental analysis to be found, for instance, in Feigl's influential article "The 'Mental' and the 'Physical'" (1958). But the focus on the general status of scientific theories and on the "cognitive turn" was also characteristic of the field of study. The original self-description in the *Minnesota Studies*, which appeared in the volume *The Foundations of Science and the Concepts of Psychology and Psychoanalysis* [Feigl and Scriven, 1956] confirm this assessment:

> "The first volume of Minnesota Studies in the Philosophy of Science

> presents some of the relatively more consolidated research of the Minnesota Center for the Philosophy of Science and its collaborators. Established in the autumn of 1953 by a generous grant from the Hill Family Foundation, the Center has so far been devoted largely, but not exclusively, to the philosophical, logical, and methodological problems of psychology." [Feigl, 1956, V]

Fourty years later a new assessment provides a broader image of the MCPS's fields of study. In its 16th volume, *Origins of Logical Empiricism* [Giere and Richardson, 1996] a striking survey is given:

> "Minnesota Studies in Philosophy of Science is the world's longest running and best known series devoted exclusively to the philosophy of science. Edited by members of the Minnesota Center for the Philosophy of Science ... (MCPS) since 1956, the series brings together original articles by leading workers in the philosophy of science. The ... existing volumes cover topics ranging from philosophy of psychology and the structure of space and time to the nature of scientific theories and scientific explanation." (MCPS in Internet)

The various thematic volumes of the MSPS largely confirm this. The first volume includes studies on Logical Empiricism (Feigl), the methodological status of theoretical concepts (Carnap), a critique of psychoanalysis (Skinner), an account on radical behaviorism (Scriven), the principles of psychoanalysis (Ellis), motives and the unconscious (Flew), psychological tests (Crobach/Meehl), ego-psychology (Meehl), the logic of general behavioral system theory (Buck), the concept of emergence (Meehl/Sellars), empiricism and cognitive philosophy (Sellars) and the human sciences within the canon of science (Scriven). The critical reevaluation of philosophy of science leads Feigl to a modified image of its development since the Vienna Circle and at the same time indicated a qualitative transformation in terms of pluralization and relativism:

> "I have tried to convey my impression that the philosophy of science of logical empiricism, after twenty-five years of development, compares favorably with the earlier logical positivism, in that it is, firstly, more logical ... Secondly, it is more positive, i.e., less negativistic ... Thirdly, logical empiricism today is more empirical, in that it refrains from ruling out by decree ontologies or cosmologies which do not harmonize with the preconceptions of classical positivism. Alternative and mutually supplementary logical reconstructions of the meaning of cognitive terms, statements, and theories have come increasingly to replace the dogmatic attempts at unique reconstructions. Logical empiricism has grown beyond its adolescent phase. It is rapidly maturing, it is coming of age..." [Feigl, 1956, p. 34]

This statement on the task of a differentiated approach to philosophical research in the theory pool of the original Logical Empiricism gives us further evidence of an interdependent evolution of theories since the beginning of the thirties. This, together with an exchange of ideas and cooperation between Austria and America, transcend the classical boundaries of a traditional history of reception. A further area of study will emerge, if the as yet unpublished archival materials and the numerous activities of the MCPS are analyzed and interpreted in this context. Such a project has to focus on the Austro-American transfer, transformation and retransfer of the philosophy of science in the period from 1930 to 1960.

4.2 Rudolf Carnap, Karl Menger and the "Chicago Circle"

With his trips to America in 1929 and 1931/32, Schlick paved the way for his student Herbert Feigl, who decided very early in 1931 to emigrate to the US because of the increasing anti-Semitism in Austria. With his networking a door to America was opened for Rudolf Carnap, the most influential thinker of the Vienna Circle. On an organizational level, the Unity of Science movement became significantly international at the first congress in Paris 1935. Neo-pragmatism and Logical Empiricism found a platform thanks to Otto Neurath's untiring efforts together with Morris (who had become the driving force behind the semantically oriented synthesis of Logical Empiricism and pragmatism) and Carnap to launch the joint publication project of the "International Encyclopedia of Unified Science" (from 1938). The three philosophers were involved in the planning of the six "International Congresses for the Unity of Science" in Paris (1935 and 1937), Copenhagen (1936), Cambridge (1938), and already during World War II in Harvard (1939) and Chicago (1941).

The contact with Morris enabled Carnap to gradually emigrate to the United States. After a stay in London in 1934, Carnap traveled to the United States for the first time in December 1935 — this move was also motivated by the increasingly unbearable political atmosphere in Prague. Already the year before he had met Willard Van Orman Quine (Harvard) in Vienna and Prague, which was followed by an intense dialogue and continuous contact following his emigration to Chicago in 1936.

At the University of Chicago, Carnap and Morris held a regular colloquium, known as the "Chicago Circle", on methodological and interdisciplinary issues, even if the knowledge of modern logic was somewhat limited there.

With this development, Carnap broke with the original conception of the logic of science ("Wissenschaftslogik") understood as a logical syntax of language. Influenced by the work of Alfred Tarski, who immigrated from Warsaw in 1939, Carnap had undergone a "semantic turn" in the US by the time his *Introduction to Semantics* (1942) appeared. And the discussion of Quine's "Two Dogmas of Empiricism" (1951) drew from the early moment on Carnap's sensitivity to the question of the analytic/synthetic or theoretical/empirical dualism.

Carnap writes about this new circle in exile:

> "In Chicago Charles Morris was closest to my philosophical position. He tried to combine ideas of pragmatism and logical empiricism.

> Through him I gained a better understanding of the Pragmatic philosophy, especially of Mead and Dewey.
>
> For several years in Chicago we had a colloquium, founded by Morris, in which we discussed questions of methodology from scientists from various fields of science and tried to achieve a better understanding among representatives of different disciplines and greater clarity on the essential characteristics of the scientific method. We had many stimulating lectures; but, on the whole, the productivity of the discussions was somewhat limited by the fact that most of the participants ... were not sufficiently acquainted with logical and methodological techniques. It seems to me an important task for the future to see to it that young scientists, during their graduate education, learn to think about these problems both from a systematic and from an historical point of view." [Carnap, 1963, p 34f]

Referring to these meetings the editors of Karl Menger's *Reminiscences* add the following remarks:

> "Carnap and Morris had organized a discussion group, inevitably called the 'Chicago Circle', which met irregularly on Saturday mornings at the University of Chicago. As often as he could Menger came from South Bend and participated in the sessions of the Circle.
>
> The one tangible accomplishment of the Chicago Circle was to get some of its participants to write, and the University of Chicago Press to publish, the first monographs in the series called the International Encyclopedia of Unified Science. Apart from this the Circle suffered from an early series of blows from which, although it continued to meet in an desultory fashion until the 70's, it never fully recovered. The first of these was the departure of the noted linguist Leonard Bloomfield from the University of Chicago to become Sterling professor at Yale ... The next major and practically fatal blow ... was the war, which in the United States began in 1941, and which disrupted academic life in general." [Menger, 1994, pp. XIII f]

With respect to the early transatlantic communication on logical and mathematical issues, the role played by Karl Menger cannot be overemphasized. Through his journeys abroad and his publications in the thirties, both the results of his Viennese "Mathematical Colloqium" (1928–1936) as well his own works on scientific logic became known internationally.

Already in 1930 he lectured at Harvard, where he came into contact with philosophers like Percy Bridgman and Paul Weiss, who later founded the *Journal of Symbolic Logic* (from 1936 on), in which he and his famous disciple Kurt Gödel as well as Carnap published their work in the following years. Just as Bridgman was influenced by Mach's ideas, this was also true, surprisingly enough, of the Austrian economist Joseph Schumpeter, who joined Philipp Frank discussion group in Harvard.

It was Menger who lectured on Kurt Gödel's revolutionary work on completeness and consistency in the US before he decided to emigrate in 1937 from Vienna to Indiana (University of Notre Dame) because of the depressing political situation in Austria, especially after the murder of Moritz Schlick in June 22, 1936. From 1946 until his retirement in 1971, he taught at the Illinois Institute of Technology in Chicago, where he continued his Viennese project as the organizer and editor of the (less successful) *Reports of a Mathematical Colloquium.*

The last two "Congresses for the Unity of Science" in Harvard and Chicago functioned as a forum for the transfer of knowledge and the transformation of philosophy of science into the international Unity of Science Movement:

> "Quine wrote simply: 'Basically this was the Vienna Circle, with accretions, in international exile'. One might say that Mach's spirit had found a resting place in the New World at long last, and that the advance of the Vienna Circle had arrived at Harvard Square." [Holton, 1993, p. 62]

In summary, it is clear that a basis for dialogue between Vienna and Chicago in philosophy of science had been created on various levels already prior to the outbreak of World War 2 and the preceding cultural exodus from Austria. A path had been paved for the actual transfer of knowledge in the context of (direct and indirect) contacts, congresses and journals: the International Congresses for the Unity of Science (1935–1941) and the *International Encyclopedia of Unified Science* (1938–1962) provided a framework and forum for this scientific communication.

The final congress at the University of Chicago (September 2–6, 1941) ushered in the phase of internationally established philosophy of science — in spite of the fact that the program had been reduced as a consequence of the war. The Chicago congress united Americans and emigrants including "Contributions from Europeans" whose papers where presented in the absence of their authors. The discussions focused on "The Task of the Unification of Science", "Logic and Mathematics', "Psychology and Scientific Method', and brought together scholars from the encyclopedia project (like Morris, Neurath, Feigl, Carnap, Brunswik, Reichenbach, and Hans Kelsen) as well as new scientists such as Alfred Korzybski, Lewis Feuer, and Charles Stevenson.

From 1938 on, publications on these activities were edited by Neurath, Carnap and Morris as part of the *International Encyclopedia of Unified Science* (IEUS), a modernist project that extended into the 1960s but was to remain uncompleted.

At the same time, the Journal *Erkenntnis* edited by Carnap and Reichenbach, became international with the eighth (and last) volume as *Journal of Unified Science*, after it had come under pressure by the Nazi regime in 1933 [Spohn, 1991]. In 1938, the first volume of the IEUS, with contributions by Neurath, Niels Bohr, John Dewey, Bertrand Russell and Carnap, marked the beginning of the uncompleted project with 19 instead of 260 projected monographs published with the University of Chicago Press (Reprint of all 19 monographs: Neurath/Carnap/Morris 1970/71).

Even though the editors had very different ideas about the unification of the sciences, the project was continued also after the war although the death of Neurath (1945) and the Cold War resulted in the deterioriation of the whole enterprise of "late Enlightenment". The last path breaking contribution by Thomas Kuhn on *The Structure of Scientific Revolutions* (1962) can be seen as reflecting a change in the philosophy of science, characterized as a pragmatic or sociological turn with the philosophy of science embedded in a historical context.

It is really remarkable that Carnap expressed his appreciation of Kuhn's article in two letters (April 12, 1960 and April 28, 1962) as part of the *Encyclopedia* — a fact, which was obscured by subsequent historiography for many reasons to be discussed in a different context. (cf. [Reisch, 2004]).

Given the prehistory it is no surprise that the Advisory Committee of the IEUS documents a strong UK/US-Austrian bias. Accordingly, it becomes difficult to speak of an input-output or loss-gain transfer caused by the forced *Cultural Exodus from Austria* from 1938 [Stadler and Weibel, 1995]. We are rather dealing with a multilateral dynamic of science as transfer, transformation from Central Europe to Great Britain and America, which can be described as a parallel process of disintegration and internationalization.

We now come to a third important person, who was also an emigrant from Austria, the social scientist Hans Zeisel (1905–1992):

Zeisel, well known as a co-author of the study *Die Arbeitslosen von Marienthal* in 1933 (English edition: *Marienthal: The Sociography of an Unemployed Community,* 1971) studied law and political science at the University of Vienna before he was forced for political and "racial" reasons to leave Vienna for the U.S. after the *Anschluss* in 1938. In "Red Vienna" he had cooperated with Paul Lazarsfeld, Marie Jahoda as a member of the "Wirtschaftspsychologische Forschungsstelle" at the Institute of Psychology run by Charlotte and Karl Bühler of the University of Vienna. In New York he met Lazarsfeld again and worked as manager of media research at the McCann Erickson advertising agency (1943–1950) and was also an executive at the Tea Council (1950–1953). During this time he published the textbook *Say it with Figures* (1947) which became a standard reference book in social research with six editions and countless translations.

In 1953 Zeisel was appointed professor at the University of Chicago Law School, where he applied empirical social research to the field of (sociology of) law.

Together with Harry Kalven, Jr. he worked on the American jury system, esp. provided by the book *The American Jury* (1966). His *Prove it with Figures: Empirical Methods in Law and Litigation* (1997) is a sort of summary of his lifelong focus on social research as a tool for tackling legal problems. Apart from this focus, he, like Carnap, opposed the Cold War legislation and capital punishment for a number of reasons.

Zeisel also had studied philosophy under the influence of the Vienna Circle (both methodologically and scholarly), especially as a participant of Carnap's lectures in

Vienna. One year before he died in Chicago, March 7, 1992, he participated in the founding conference of the Institut Wiener Kreis/Vienna Circle Institute in 1991 ("Wien–Berlin–Prag: Der Aufstieg der wissenschaftlichen Philosophie" on the occasion of the centenaries of Rudolf Carnap, Hans Reichenbach and Edgar Zilsel). In his posthumously published contribution to the proceedings (Haller/Stadler 1993) entitled "Erinnerungen an Rudolf Carnap" (Remembrances of Rudolf Carnap) he presented personal recollections of his beloved teacher and his admiration of Vienna Circle's philosophy and methodology in general. (By the way, Herbert Feigl was a distant relative of Zeisel's family; he also took issue with the break with this tradition at the University of Vienna after 1945). Especially the dualism of facts and values seemed to him most important for his own work, inspired by the analysis of language and the anti-metaphysical orientation.

And Zeisel confirmed Carnap's autobiographical allusion (1963) to the fact that he was not really happy in Chicago because of the administration and the situation of the so-called "continental" philosophy there, dominated by Richard McKeon and Mortimer Adler. And he reports that Carnap only had few students and was somewhat isolated from the academic life – despite the "Chicago Circle" with Charles Morris.

Although Carnap left Chicago for Los Angeles when Zeisel was appointed there, we can see the strong impact of the Vienna Circle and the appreciation felt by the younger sociologist for the philosophy of Logical Empiricism in exile, which was only a continuation of the Vienna connection from the 1920's and 1930's:

> " ... alles zusammen hat er einen bedeutsamen Einfluß auf mein Leben (ausgeübt), auf meine Arbeit, und ich glaube auch auf viele andere Menschen. Sein steigender Weltruf ist nicht nur berechtigt, sondern entwickelt sich mit großer Klarheit, denn er war einer der großen Philosophen dieses Jahrhunderts, und als den haben wir ihn empfunden und geschätzt." [Zeisel, 1993, p. 223]

Summarizing this short account one might ask if there wasn't a sort of re-transfer of the "Chicago Circle" to Europe after 1945? One could first answer: there was a remarkable hidden impact, which could be the subject of future research focusing on following aspects:

1. Through the Austrian philosopher Wolfgang Stegmüller, who came into contact with Carnap after World War 2 the modern philosophy of science was introduced in the German speaking academia as (modified) analytic "Wissenschaftstheorie". `http://www.univie.ac.at/ivc/stegmueller`.

2. Hans Zeisel was an important figure, serving as both initiator and *spiritus rector* of the still existing Vienna-based "Institut für Kriminalsoziologie" (Institute of Legal and Criminal Sociology), founded in 1973 as one manifestation of the legal reform under the former Austrian Minister of Justice Christian Broda.

4.3 Philipp Frank and the "Institute for the Unity of Science"

One of the most important figures for the transfer, transformation and the further development of the Central European philosophy of science was the physicist/philosopher Philipp Frank (1884–1966) whose work has continued to have an indirect influence through his students up to this day.

Frank was Einstein's successor in Prague where he worked as full professor and director of the Institute for Theoretical Physics from 1912 until he had to emigrate in 1938. As a leading member of the Vienna Circle he significantly contributed, together with Carnap and Neurath, to the dissemination of Logical Empiricism. (Frank 1949) After giving a series of lectures on modern physics at various American universities, he taught at Harvard University until 1953 — first as lecturer and later as Hooper Fellow. Given his age he was no longer able to obtain an adequate position, yet his charisma and skills as an intellectual and as an organizer allowed the younger Herbert Feigl to become a central figure of intercontinental philosophy of science. Up until his death he played a seminal role in the Unity of Science movement and in organizing an innovative interdisciplinary forum for discussion at Harvard. Through these activities and functions and through his students (including, among others, Robert S. Cohen, Gerald Holton, Ernest Nagel) he also influenced an entire generation of young philosophers of science. The latter, in turn, contributed significantly to a critical further development of the philosophy of science, which had changed significantly as a result of the immigration of intellectuals. They have continued to shape the scientific scene through their academic positions and publications. It was, above all, Philipp Frank's achievement that science became a field of discourse, next to philosophy and religion, in the war-ravaged intellectual scene - in the heyday of the McCarthy era (Reisch 2004). Together with the American Academy of Arts and Sciences, he succeeded in positioning philosophy of science in a number of events. Frank wrote a very enlightening account of the development from Vienna, Berlin and Prague to "Harvard Square" (Introduction to 1949). This intellectual history was illustrated by Gerald Holton (1993 and 1995) in connection with a European-American reception. On his first teaching job in the USA, Frank wrote the following:

> "Since the fall of 1939 I have had the privilege of teaching at Harvard University not only mathematical physics but also the philosophy of science. This teaching has been a great experience for me and has been of great influence on my philosophical writing. I started with an audience of about fifteen students. Since this was an unusual subject I did not quite know what to tell them. I began by presenting to them the logical structure of physical theories as envisaged by logical empiricism. But very soon I noticed this was not the right thing to do. The frequent discussion that I had with the students showed me what they really wanted to know. By a process of interaction, a program was finally worked out that was a compromise between what I wanted to tell the students and what they wanted to know." [Frank, 1949, p.

50]

Based on this interaction, Frank also developed the curricula for science courses at Harvard, giving philosophy of science a more contextual orientation and including the historical and sociological perspectives. His efforts were backed by Harvard president J. B. Conant, who had strongly advocated a new approach to teaching science which he presented in his book *On Understanding Science* (1947). In his book *Science and Common Sense* (1951), he further refined these considerations with reference to the interdisciplinary discussion circles at Harvard, which Thomas Kuhn and Bernard I. Cohen were part of. In this connection, it is interesting that Frank saw all his writings after 1940 as influenced by his involvement with teaching at Harvard:

> "My point was now that the philosophy of science should, on the one hand, give to the science a more profound understanding in his own field, and on the other hand, be for all students a link between the sciences and the humanities, thus filling a real gap in our educational system." (Ibid., 51)

Having become increasingly interested in the issue of metaphysical interpretations of modern science as a result of his US acculturation, Frank began to systematically analyse the various metaphysical understandings of science in a logico-empirical and socio-psychological way. Since 1940 he was invited by Harlow Shapley, the director of the Harvard College Observatory, to the annual "Conferences of Science, Philosophy and Religion" which viewed science from a critical perspective:

> "I addressed this group several times between 1940 and 1947. My contributions were mostly around the question of whether 'relativism' of modern science is actually harmful to the establishment of the objective values in human life. I made an argument to prove that 'the relativism of science' has also penetrated every argument about human behavior. 'Relativism' is not responsible for any deterioration of human conduct. What one calls 'relativism' is rather the attempt to get rid of empty slogans and to formulate the goals of human life sincerely and unambiguously.' (Ibid., 52)

With this proclamation of an anti-absolutist understanding of science, the Einstein biographer Frank described "Science in Context" along similar lines as the *Encyclopedia* project and documented it in his book *Relativity — A Richer Truth* (1950). In the German edition of this 1952 publication, where he distinguished clearly between arbitrariness and convention in science and ethics on empirical grounds, there is a programmatic preface written by Albert Einstein.

In Harvard a so-called "Science of Science Discussion Group" had been formed as early as 1940/41 in which Frank also participated. (Hardcastle 2003, 170-196). It was obviously to serve as a model for the "Interscientific Discussion Group" that began meeting in 1944:

> "In the Fall of 1940 around Harvard University, an invitation was distributed to thirty-eight scientists from various fields, together with some logicians and methodologists present this year in Harvard. 'As an effort in the direction of debabelization' it began, 'the undersigned committee is organizing a supper-and-discussion-group to consider topics in the Science of Science' ". [Hardcastle, 1996, p. 24]

The participants of this monthly transdisciplinary discussion forum on theory and study of science included, in addition to Frank, Rudolf Carnap, Herbert Feigl, Willard Van Orman Quine, Richard von Mises, Alfred Tarski, Nelson Goodman, George Polya, Percy Bridgman, as well as the psychologists E. G. Boring, and S. S. Stevens and economist Joseph A. Schumpeter. Stevens had organized this forum after talks with Carnap and Frank.

This platform, a clear sign of the cooperation of the Vienna Circle in exile and American philosophy of science, paved the way for the scientific communication that was to follow and represented a sort of proto-circle of the expanded "Interscientific Discussion Group" that lead up to the Institute for the Unity of Science.

> "Between roughly 1940 and the end of the 1950's, the movement for a scientific philosophy in the USA flourished, pushed forward especially by the influx of arrivals from Europe. The main direction of the movement brought over from Europe was now identified most often by the slogans 'Unity of Science' and 'Unified Science', versions of the old terms Einheitswissenschaft and Gesamtauffassung which had been prominent in the manifestos of 1911–12 and 1929 as well as Carnap's Aufbau — a concept that had roots in the phenomenalistic monism of Mach." [Holton, 1993, 62]

This movement was seen as based on the *Encyclopedia of Unified Science* and the *Institute for the Unity of Science (IUS)* under the auspices of the *American Academy of Arts and Sciences* (AAAS) — one of the manifestations of the intellectual symbiosis between related European and American movements.

What was this new institution and what role did it play in the transfer of science already sketched ? The published statutes of the IUS read as follows:

> "The purposes for which the corporation is formed are to encourage the integration of knowledge by scientific methods, to conduct research in the psychological and sociological backgrounds of science, to compile bibliographies and publish abstracts and other forms of literature with respect to the integration of scientific knowledge, to support the International Movement for the Unity of Science, and to serve as a center for the continuation of the publications of the Unity of Science Movement." (*Synthese* 1947, VI, 158f., cited after [Holton, 1993, p. 72]

The AAAS Proceedings were used as a basis for publications as well as its infrastructure for diverse conferences and symposia. The participants of these considerably expanded discussion rounds show how the forum increasingly became more open, providing a setting in which the sciences could be discussed within a cultural and social context.

Even if Quine documented these regular meetings as a "Vienna Circle in exile" [Holton, ibid., p. 63], there had been a considerable leap in terms of quality and pluralization as becomes clear in the wide range of topics and in the composition of the participants. The "academization" of *Wissenschaftslogik*, [Dahms, 1987], applied to the proponents, but not to the organizational form outside of the universities. Holton is right in asking how an "ecological niche" was created for this hybrid scientific movement within two decades. Apart from the anti-metaphysical and empiricist orientation of pragmatism, a number of related factors can also be mentioned here, e.g.,the preceding personal contacts on an university level, the private organizations promoting science such as the Rockefeller Foundation (e.g., for Feigl in Harvard) and the anti-Nazi outlook of the Scientific Community — all factors that should be evaluated more closely. The most important factor was ultimately the recognition of the high quality of émigré scholars, which could be exemplified by the integration of Philipp Frank as a consequence of Bridgman's and Conant's efforts.

Whether this intellectual osmosis also provides an adequate explanation of the qualitative transformation and paradigm shift that has taken place since the 'received view' remains to be seen. In any case, it is certain that this interaction significantly enriched philosophy both theoretically and methodologically.

The personal composition of the "Interscientific Discussion Group" (IDG), which met in Harvard since 1944, was an interesting mixture of older and younger scientists of American and European origin. Most of them spoke at both of the Unity of Science Congresses in Harvard and Chicago [Holton, 1995]. Personal invitation letters from the committee (Percy W. Bridgman, Walter Cannon, Philipp Frank, Philippe LeCorbeiller, Wassily W. Leontief, Harlow Shapley and George Uhlenbeck) were also sent to Karl W. Deutsch, Roman Jakobson, Willard Van Orman Quine, Charles Morris, Richard von Mises, Ernest Nagel, Giorgo de Santillana, Victor F. Weisskopf and Norbert Wiener, who were requested to participate regularly. The group described itself as follows:

> "Our group consists of persons in different fields who feel that the extreme specialization within science demands as its corrective an interest in the entire scientific edifice. We plan to hold meetings from time to time in which discussions of different topics will be led by competent scholars." (Inter-Scientific Discussion Group, December 30, 1944, cited after [Holton, 1995, p. 284]

In the program planned for 1945, three large areas, "Logic of Science", "Psychology" and "Sociology of Science", were announced along with further subtopics. This reflected both an internal and external perspective on philosophy of science and confirmed psychology and sociology of science anticipating Kuhn's work. Holton, at the time IDG secretary, saw this program in retrospect as continuing the tradition of the Vienna Circle, the Verein Ernst Mach and their programmatic manifesto of 1929. The speakers and the themes of these early meetings comprised the history of philosophy (Santillana), psychoanalysis and social science (Talcott Parsons), mathematics/statistics (Richard von Mises), cybernetics (Norbert Wiener), biology and social science (Georg Wald) and science in general (C. J. Ducasse) — in the context of Morris's conception of semiotics (syntax, semantics, pragmatics).

The osmotic process of discussion between the immigrant philosophers of the first wave since 1930, of the second wave from 1938 on and of the American scientific community became an experimental setting for a future history *and* philosophy of science. Sociology of science (Talcott Parsons) was represented along with philosophy of economics (Paul Samuelson, Gottfried Haberler, Joseph Schumpeter), cybernetics (Norbert Wiener), mathematics (Gustav Kuerti), and political science (Karl W. Deutsch), psychology (Gordon Allport), and history of science (I. B. Cohen, G. de Santillana). The fact that philosophy of science in a more narrow sense also had strong presence in American philosophy (e.g., C. I. Lewis, W. V. O. Quine) was decisive for a theory dynamics that involved both diffusion and confrontation of different movements. Lecture themes such as Simplicity of Science, What is Science?, Psychoanalysis and Social Sciences, Sense and Nonsense in Modern Statistics, Biology and Social Behavior, Living Organisms and the Second Principle of Thermodynamics, Stability and Flux in the Living Organism, Relation of Hypothesis and Experiment, served to stimulate the interdisciplinary dialogue. A highlight was certainly Oskar Morgenstern's presentation of his and John von Neumann's *Theory of Games and Economic Behavior* (1944) on February 28, 1944. This book had an immense influence on modern social sciences. Here a line of reception became manifest for which the ground had been laid in the Vienna years.

To complete this theoretical development the following could be added: After emigrating back to Austria, Morgenstern tried, after World War 2, to bring modern social scientific research back to Austria; together with Paul Lazarsfeld, he founded the Viennese "Institute for Advanced Study" in 1963.

This informal IDG circle gave rise to the need for a continuous, expanded forum, which resulted in the initiative for founding the "Institute for the Unity of Science" (IUS). Philipp Frank's efforts resonated particularly well with Harvard president James Conant, since they related to his General Education Program — a sort of survey course on the scientific disciplines. With the help of the Rockefeller Foundation, the IUS — a sort of international variant of the former Vienna Ernst Mach Society — was established in 1947 — in cooperation with the AAAS, officially based in Boston. The self-description reads as follows:

> "This Institute is a non-profit corporation which has offices in Ithaca, New York and Boston, Massachusetts. The Charter says, 'The purpose for which the corporation is formed, are to encourage the integration of knowledge by scientific methods, to conduct research in the psychological and sociological backgrounds of science, to compile bibliographies

> and publish abstracts and other forms of literature with respect to the integration of scientific knowledge, to support the international movement for the unity of science, and to serve as a center for the continuation of the publications of the unity of science movement.' The Institute attempts to stimulate interest in these issues among college students, college faculties, and among the public at large. The Institute has arranged an essay contest for college students and young college graduates. It is editing the Encyclopedia of Unified Science, published by the University of Chicago Press. It is starting research projects in the fields of semantics, logic of science, and sociology of science. It arranges discussion groups and meetings at several places in the United States. It is a part of the International Union for the Philosophy of Science. It cooperates with the International Society for Significs (psycholinguistic studies) in Amsterdam and is organizing, together with this society, an international meeting in Amsterdam. In cooperation with the European societies for the philosophy of science..., this Institute publishes communications in the international journal 'Synthese' which is published in Amsterdam and is the central organ of these groups...
>
> The Institute cooperates also with the movement for general education, which attempts to integrate the college curriculum and to break down the barriers between the departments. The Institute arranges lectures and courses at different places in the United States." (Announcement of the IUS, cited after [Holton, 1995, p. 288]

In this self-description there are at least three important factors. First, the international nature of the activities in the U.S. and of those between the American-European institutions, second, the interdisciplinary orientation and third, the strongly public-oriented and educational political motivation resembling the popularization efforts of the Vienna Circle.

In a theoretical sense, it is interesting to note the presence of *Wissenschaftslogik* ("logic of science"), together with the sociology of science, as well as the reference to the Dutch representatives of the Significs movement (Gerrit Mannoury), which Neurath worked together with in exile until 1940. This connection, which has hardly been taken into account until now was mentioned explicitly in the journal *Synthese* which was published from 1936 to 1939 and after World War, from 1946 on. The contributors to this organ included Gustav Bergmann, E. W. Beth, P. W. Bridgman, L. E. J. Brouwer, Rudolf Carnap, J. Clay, R.S. Cohen, Karl W. Deutsch, C.G. Hempel, G. Holton, W. McCulloch, Karl Menger, Charles Morris, Ernest Nagel, Otto Neurath, Karl Popper, W. V. O. Quine, N. Rashevsky, Moritz Schlick, Herbert A. Simon, Friedrich Waismann as well as Philipp Frank, who contributed to *Synthese* (6, 1947/48). And in the book series "Synthese Library", also launched in 1959, one finds an early *Festschrift* for Rudolf Carnap [Kazemier and Vuysje, 1962] and the first volume of the "Boston Studies in the Philosophy of Science" [Wartofsky, 1963].

The international orientation of the early Unity of Science movement was already reflected in a separate *Synthese* supplement which appeared as *Unity of Science Forum* in the period 1936–1939 under the auspices of the International Institute for the Unity of Science (headed by Frank, Morris and Neurath) at the Hague. For Neurath, this institution served as a platform for the *Encyclopedia* project, while at the same time it was an important organizational bridge in the difficult years in Dutch exile. In its embryonic stage, it was also the model for the IUS in the U.S.A. after European science had been ravaged by National Socialism. The small brochures also included a report by Neurath on the "Fourth International Congress for Unity of Science" in Cambridge (August 1938) and relevant articles or reviews on the Unity of Science Movement. The editors' board of *Synthese* had welcomed the *Unity of Science Forum* already in 1936.

This was only a brief digression to the Dutch background of the IUS, which ten years later was able to work under considerably more favorable conditions.

Already the composition of the Board of Trustees assured a more successful point of departure: Philipp Frank as president, Charles Morris and Ernst Nagel as vice-presidents, Milton R. Konvitz as treasurer and the further members Percy W. Bridgman, Egon Brunswik, Rudolf Carnap, Herbert Feigl, Carl G. Hempel, Hudson Hoagland, Roman Jakobson, Willard Van Orman Quine, Hans Reichenbach, Harlow Shapley and Stanley S. Stevens. This was a renowned group of older and younger generation scientists, both emigrants and American scholars from philosophy, natural and social sciences, who negotiated and coordinated lectures, meetings, conferences and publications from 1948 to 1966.

If we now analyse the related archival material, we note that a more consistent line was taken towards the themes addressed in both natural science and social science. This also meant that issues were dealt with also from a historical and sociological perspective: symposia titles such as "Science and Value", "Logic and the Sociology of Science", "Social Physics" or "Current Issues in the Philosophy of the Sciences" reflected this more open approach and self-reflection within the sciences. Of the individual lectures, the following merit mention: the one given by Roman Jakobson (linguistics), a series on the problem of meaning in the individual disciplines (W. V. O. Quine, P.W. Deutsch), information theory (D. Gabor). A central event was the Boston "Conference on the Validation of Scientific Theories" from December 27 to 30, 1953. The proceedings were edited by Philipp Frank in a book with the same title, at Beacon Press in Boston. The sections included — Acceptance of Scientific Theories, The Present State of Operationalism, Freud's Psychoanalytic Theory, Organism and Machine as well as Science as a Social and Historical Phenomenon — illustrate the cognitive process of transformation in the Philosophy of Science beyond a syntactic/semantic *"Wissenschaftslogik"*. In his introduction, Frank described the situation determined by an increasingly critical view of science as follows:

> "In order to obtain a basis from which one can pronounce sound judgment about this situation one should put the question: In what sense

> does science search for the 'truth' about the universe? ... What are the criteria under which we accept a hypothesis or a theory? If we put this or a similar question, we shall see soon that these criteria will contain, to a certain extent, the psychological and sociological characteristics of the scientist, because they are relevant for the acceptance of any doctrine. In other words, the validation of 'Theories' cannot be separated neatly from the values which the scientist accepts. This is true in all fields of science, over the whole spectrum ranging from geometry and mechanics to psychoanalysis." [Frank, 1956, VII f.]

Here the psycho-sociological turn can easily be recognized as a programmatic requirement for every future history and philosophy of science. The conference was sponsored by the Institute for the Unity of Science as an event of the American Association for the Advancement of Science (AAAS). The individual contributions reflect the pragmatic and operational dimension in the validation and corroboration of theories. Comparing the natural and human sciences from a cross-disciplinary perspective, Warren McCulloch noted:

> "Cybernetics has helped to pull down the wall between the great world of physics and the ghetto of the mind." [Ibid., X]

To elaborate these general principles in more concrete terms, historians of science and scholars were invited; among them: Henry Guerlac, Alexandre Koyré, Robert S. Cohen and E.G. Boring. The conference was headed by R. J. Seeger, H. Margenau, H. Feigl, G. Wald and G. Holton who together had directly or indirectly worked on a radical reform of the philosophy of science, which had already been in the making for three decades.

Against this background, Gerald Holton's argument must be reconsidered. Was it really only Alexandre Koyré's philosophy of science (epistemology) that brought about a loss of meaning in empiricist unified science? The cited publications from the fifties written by Frank and others around him — in connection with the AAAS publications (*Proceedings* and *Daedalus*) indicate that there was actually a trend towards a sort of "Science of Science". The thematic issue "Science and the Modern World View" of *Daedalus* (winter 1958) clearly reflects the complementary contribution made to the symposium volume mentioned above. This volume was presented by Holton himself in his role as editor-in-chief of the Academy on the occasion of the retirement of Bridgman and Frank. The *Weltanschauung* analyses demonstrating an internalist philosophy of science that had become relativized and pluralized are impressing documents of this evolutionary and multi-facetted transition from the received or standard view to the non standard view of scientific theories, or to put it differently, the transition from text to context that was to be "proclaimed " ten years later by a younger generation of scholars as a decisive event [Suppe, 1977].

The conferences organized by AAAS together with IUS from 1951 to 1954 and the proceedings published in the four volumes of *Contributions to the Analysis* *and Synthesis of Knowledge* show the broad spectrum of approaches in the recent discussion on the unity/diversity of sciences [Galison and Stump, 1997].

Alluding to the *Old Testament* Frank described the dangers of Babylonian linguistic confusion in the 20th century in his contribution on "Contemporary Science and the Contemporary World View":

> "As a matter of fact, the view that science is the product of abstraction from our rich and full experience is rather misleading. It has become more clear by the evolution of science in our century that the principles of science are not dehydrated abstractions but a system of symbols that is produced by the creative imagination of the scientist." [Frank, 1958, p. 59]

Referring to Richard von Mises' *Positivism. An Essay in Human Understanding* (1938/1951), Frank once again gave — in analogy to poetry — a non-foundationalist and relativistic account of science reflecting a certain continuity with the thirties:

> "We have seen that the main activity of science does not exist in producing abstractions from experience. It consists in the invention of symbols and in the building of a symbolic system from which our experience can be logically derived. This system is the work of creative imagination which acts on the basis of our experience."
>
> Subsequently he concludes with the late Wittgenstein: 'One might give the name 'philosophy' to what is possible before all new discoveries and inventions'." [Ibid., p. 65]

This elegant attempt to rehabilitate philosophy, science and philosophy of science in a critical phase governed by public skepticism vis-à-vis science and the controversy over *The Two Cultures* [Snow, 1959] provided crucial impulses for the numerous accounts of *Science and Antiscience* [Holton, 1993] and for today's historiography situated between modernism and postmodernism (cf. [Galison, 1996]).

It would thus be clearly problem-oriented if the eternal question as to the unity or disunity of science had already then been dealt with in various ways — parallel to and correlating with the *Encyclopedia of Unified Science* — be it in the guise of "social physics" (J. Q. Stewart), the unification of systems theory (N. Wiener) or sign theory (Ch. Morris). Holton's more recent publications, including *The Scientific Imagination* [1978], *The Advancement of Science, and its Burdens* [1986], *Science and Anti-Science* [1993b], are also manifestations of this problem located in the specific constellation of science, society and *Weltanschauung.*

It seems as if the external factors of this evolutionary, cognitive process slowly led to a convergence of the "Vienna Circle in America" [Feigl, 1968] with American philosophy of science. It would thus be inaccurate to say that the emergence of a younger generation (Quine, Kuhn, Hanson, Feyerabend, et al.) had resulted in the demise of a philosophical school

The fact that in recent years a new generation has joined ranks in the *History of Philosophy of Science* as well as the *History and Philosophy of Science* (cf.

HOPOS or IVC) documents this historico-pragmatic turn in the philosophy of science.

After Frank and Bridgman, Robert S. Cohen has been a seminal figure in the transmission and further development of this Austro-American philosophy of science. Together with Marx W. Wartofsky, he was active already in the fifties as scholar and secretary of the IDG and the IUS. He finally founded his own forum, the Boston Colloquium for the Philosophy of Science, where also the pioneers of this movement, Frank and others, appeared. Since 1959 this institution has contributed to the continuity and criticism of the philosophy of science. After the IUS was dissolved, the remaining funds were directed to the journal *Philosophy of Science* and the newly founded *Philosophy of Science Association* [Holton, 1995, p. 279]. Holton's own recollections as an immediate participant of this movement in the forties and fifties illustrate in an exemplary way what is generally described abstractly as history of reception:

> "As I fully appreciated at the time, for a young person, participating in these activities was immensely stimulating. Perhaps precisely because of the high density of superb intellectuals, the various leading members of the group, brought together by remnants of the Vienna Circle, made no effort to accept any uncomfortable agreements, but relished in the most wide-ranging debates. I never felt that I had to follow, or to struggle against, any doctrinaire master. When my own first historical studies convinced me of the need to add Thematic Analysis to the older tool-kit of the historian and the philosopher of science, I sensed only encouragement, instead of the kind of opposition one might have expected from rock-hard logical empiricists. If I had to characterize the members of that group in one sentence, I would focus on their unlimited curiosity and their generosity of spirit, a generosity which seemed founded on their ever-youthful self-confidence. When future historians study the philosophy of science during the middle part of this century, I hope they, too, will remember this." [Holton, 1995, p. 279]

4.4 Robert S. Cohen and the "Boston Center for the Philosophy and History of Science"

The physicist Robert S. Cohen was active at Columbia University, in the Division of War Research and on the Communications Board of the U.S. Joint Chiefs of Staff during World War 2. He met most of the members of the Vienna Circle personally after 1938, and, as already mentioned above, he was involved in the organizational work of IUS in its final phase. (On his life and work, see [Gavroglu *et al.*, 1995]).

As early as 1953 he participated in the conference "The Validation of Scientific Theories" where he contributed a paper on "Alternative Interpretations of the

History of Science" [1956]. Here his later, untiring work as organizer, editor and scholar is already alluded to:

> "I am neither a historian nor a sociologist, and at a symposium of the unity of science movement I can only join with those, who are deploring the lack of detailed studies in this history of the social relations of science. I can only regret that the sociology of knowledge, especially of science, has remained so long outside that movement's sweep, and so largely in the hands of metaphysically oriented phenomenologists and other speculative thinkers. The early death of Edgar Zilsel, a pioneer in the sociological treatment of science, left his work tragically incomplete." [Cohen, 1956, p. 219]

His references to Zilsel and also to H. Guerlac, E.G. Boring and A. Koyré stake off the area for Cohen's idea of a history and philosophy of science and for his *Boston Studies in the Philosophy of Science* [1963 ff.] book series which grew out of the Colloquium and is still being continued today. This series has presented a volume on Edgar Zilsel's various works on *The Social Origins of Modern Science* [Raven *et al.*, 2000]. It thus follows a line leading back to the early phase of the Boston Center (BCPS) and the Boston Studies (BSPS).

The first BSPS volume edited by Marx Wartofsky — published in the *Synthese Library* — presented the contributions of the initial phase of 1961/62. It was a very pluralist program ranging from the mind-body problem, scientific language and concept formation, logical foundations of physics (Philipp Frank, among others), modal logic (W. V. O. Quine, Saul Kripke, among others), quantum theory, falsificationism and holism in the philosophy of science (Adolf Grünbaum) to experience and language (Noam Chomsky). An interdisciplinary survey was presented along with comments reflecting the principles of the colloquium:

> "Initiated in 1960 as an inter-university interdisciplinary faculty group, the Colloquium is intended to foster creative and regular exchange of research and opinion, to provide a forum for professional discussion in the philosophy of science, and to stimulate the development of academic programs in philosophy of science in the colleges and universities of metropolitan Boston. The base of the Colloquium is our philosophic and scientific community, as broad and heterodox as the academic, cultural and technological complex in and about this city." [Cohen and Wartofsky, 1963, VII]

The second voluminous volume of the *Boston Studies*, that was edited as *Proceedings of the BCPS* of 1962–64 by Robert S. Cohen and Marx Wartofsky, and printed by Humanities Press in New York, was dedicated to Philipp Frank and appeared in 1965, one year before his death. In the preface, the editors paid tribute to Frank as an individual and describe the orientation of the colloquium in keeping with Frank's scientific life work:

> "Our Colloquium construes the philosophy of science broadly, as he has advised us to do. We try to discuss open problems in the foundations of science, and, wherever relevant, to bring material from the history and cultural relations of science to bear upon such problems. We try also to talk with each other across all boundaries of discipline, to include scholars from philosophy, logic and mathematics, the physical and biological sciences, history and the social sciences, and the humanities as well." [Cohen and Wartofsky, 1965, VII]

The subsequent contributions of Frank's students, including Peter G. Bergmann, Rudolf Carnap, R. Fürth, Gerald Holton, Edwin C. Kemble, Henry Margenau, Hilda von Mises, Ernest Nagel, Raymond Seeger (on behalf of the National Science Foundation) and Kurt Sitte, illustrate the range of his intellectual charisma, his life-long efforts to convey science, the merging of science and philosophy, as well as the contextualization of science – as he demonstrated convincingly in his Einstein biography [Frank, 1947]. Gerald Holton once again underscored the importance of Frank's anthology *Between Physics and Philosophy* [1941] as a link to European philosophy of science after Ernst Mach and Henri Poincaré. Seeger's comment is interesting in the sense that it refers to science politics. In connection with the "National Science Foundation Program on the History and Philosophy of Science" he took account of Frank's suggestions and advice.

The contributions to the Festschrift ran the whole gamut of the history and philosophy of science in keeping with this general orientation, and also included articles and comments. Among the advocates of this approach we already find "young wild" thinkers (Norwood R. Hanson, Paul Feyerabend, Hilary Putnam) who criticized the proposition-oriented theory of science. The volume also included a contribution of Herbert Marcuse "On Science and Phenomenology" on Husserl with a comment by Aron Gurwitsch.

When we examine the subsequent BSPS volumes — as a reflection of the philosophy of science in the USA since the beginning of the sixties — we see the programmatic text on the themes, authors and editors of the series confirmed:

> "The series Boston Studies in the Philosophy of Science was conceived in the broadest framework of interdisciplinary and international concerns. Natural scientists, mathematicians, social scientists and philosophers have contributed to the series as have historians and sociologists of science, linguists, psychologists, physicians, and literary critics. Along with the principal collaboration of Americans, the series has been able to include works by authors from many other countries around the world. As European science has become world science, philosophical, historical, and critical studies of that science have become of universal interest as well.
>
> The Editors believe that philosophy of science should itself be scientific, hypothetical as well as self-consciously critical, human as well as rational, skeptical and undogmatic while also receptive to discussion

> of first principles. One of the aims of the Boston Studies, therefore, is to develop collaboration among scientists and philosophers. However, because of this merging, not only has the neat structure of classical physics changed, but, also, a variety of wide-ranging questions have been encountered. As a result, philosophy of science has become epistemological and historical: once the identification of scientific method with that of physics had been queried, not only did biology and psychology come under scrutiny, but so did history and the social sciences, particularly economics, sociology, and anthropology.
>
> Boston Studies in the Philosophy of Science look into and reflect on all these interactions in an effort to understand the scientific enterprise from every viewpoint."

This text, taken from the cover of the 1985 anniversary volume (Cohen/Wartofsky, eds.) can be read as the intellectual legacy of the Harvard group around Frank. It also reformulates the demand for a view of the natural and social sciences encompassing in principle all the sciences, with a relativisation of the methodological and meta-theoretical dualism of the "two cultures". Seen in this light, the very direct, dialectical argument of Marx Wartofsky is plausible. In a 1994 lecture he spoke on the influence of the exiled Vienna Circle in the Boston region, on "Invariance through Transformation: The Boston Adventures of the 'Wiener Kreis', 1960–1994". [Gould and Cohen, 1994]

In our connection, the volumes of the BSPS, which to date number more than 200, are particularly informative. They were published as *Proceedings of the Boston Colloquium for the Philosophy of Science* (volumes 1-5, 13-14, 31) or dealt with the theories of Logical Empiricism and their reception (cf. volumes 6, 8, 1, 9, 3, 7, 39, 53, 76, 87, 118, 132, 133, 168). In addition, the current BCPS program figures significantly – next to the *Minnesota and Pittsburgh Centers for the Philosophy of Science* – as the only institution following the tradition of the *Encyclopedia of Unified Science* and the IUS.

In the cited volume marking the 25^{th} anniversary of the interdisciplinary work of the international BCPS 160-1985 (Cohen/Wartofsky, eds. 1985), this quarter of a century was summed up as follows:

> "The Boston Colloquium for the Philosophy of Science began 25 years ago as an interdisciplinary, interuniversity collaboration of friends and colleagues in philosophy, logic, the natural sciences and the social sciences, psychology, religious studies, arts and literature, and the often celebrated man-in-the-street. Boston University came to be the home base. Within a few years, proceedings were seen to be candidates for the journal Synthese within the Synthese Library, both from the D. Reidel Publishing Company of Dordrecht, then and now in Boston and Lancaster too. Our Colloquium was inheritor of the Institute for the Unity of Science, itself the American transplant of the Vienna Circle, and we were repeatedly honored by encouragement and participation

> of the Institute's central figure, Philipp Frank." [Cohen and Wartofsky, 1985, VII]

In addition to the *Proceedings* which were selected and reworked following the discussions at the Boston Center, the series also includes outside volumes (monographs and anthologies — a series that was first published by Reidel Publishing Company (later: Kluwer Academic Publishers, today: Springer) in Holland. It is a collective undertaking that dates back to the pre-war period. The selection made for the anniversary volume ("Invariance through Transformation") sheds interesting light on the self-image of the series: Adolf Grünbaum wrote on holism in the philosophy of science, Hilary Putnam on explanatory models in linguistics, Nelson Goodman on the epistemological argument, Stephen Toulmin on conceptual revolutions in natural science, Herbert Feigl on empiricism, Robert S. Cohen and Marx Wartofsky on the limitations of science and historical epistemology, respectively, Carl Hempel on values and objectivity in science, Abner Shimony on the philosophy of Bohr, Heisenberg and Schrödinger. The fact that texts by Herbert Marcuse and Noam Chomsky were also included once again reflects the openness of the project which did not adhere to an orthodox logico-empiricist line of research.

Here, we can only focus on the above-mentioned BSPS proceedings and a few volumes which more or less directly relate to the history of the reception or the transformation of the Vienna Circle in exile. The first five volumes document the intellectual spectrum of the Boston Colloquium in the years 1961–1968. (The third volume ("In Memory of Norwood Russell Hanson") focused on a scholar who, along with Kuhn, Toulmin and Feyerabend, was a staunch critic of the internalist philosophy of science. Dedicated to the re-evaluation of the work of Ernst Mach the physicist and philosopher, the sixth volume [Cohen and Seeger, 1970] shows the productive reception of Mach's work leading all the way up to Feyerabend. It also confirmed Holton's reconstruction (1993) of the intercontinental Mach/Boltzmann reception since the turn of the century and in so doing prepared the way for further more in-depth analyses of Mach research. Related studies can be found in volume 143 of the BSPC, in *Ernst Mach — A Deeper Look* [Blackmore, 1992] and in studies on Boltzmann [Blackmore, 1995].

This critical reassessment of the Central European "*Wissenschaftslogik*" and its reception can be found in the commemorative Carnap volume published in 1970 [Buck and Cohen, 1970]. Similarly, the volume of Helmholtz's epistemological writings originally published by Paul Hertz and Moritz Schlick [Cohen and Elkana, 1977], the volume in memory of Imre Lakatos [Cohen *et al.*, 1976], the two volumes by Herbert A. Simon (*Models of Discovery and Other Topics in the Methods of Science*, 1977) or Nelson Goodman's *The Structure of Appearance* [1977] that followed the line of reception of Carnap's work. Reference should also be made to those editions that document, as translations, the historical and sociological expansion of philosophy of science in the sense alluded to above: *Cognition and Fact. Materials on Ludwik Fleck* [Cohen and Schnelle, 1986], *Philosophy, History and Social Action. In Honor of Lewis Feuer* [Hook *et al.*, 1988], *Beyond Rea-*

son. Essays on the Philosophy of Paul Feyerabend [Munévar, 1991], *The Natural Sciences and the Social Sciences* [Cohen, 1994].

This history of reception has come full circle in the documentation of the rediscovery and further study of Central European philosophy of science over the past twenty years after it was interrupted by the forced emigration of the leading protagonists. Here *Rediscovering the Forgotten Vienna Circle. Austrian Studies on Otto Neurath and the Vienna Circle* [Uebel, 1991] deserves mention. The international orientation of the BSPS is also underscored by the many publications on the history and philosophy of science outside of the Anglo-American world. Such studies have helped to overcome the mental barriers between East and West and the North–South hierarchy and to cultivate a dialogue of the scientific community without political and socio-economic restrictions.

The *Festschrifts* for both of the two leading figures of the BCPS, Robert S. Cohen and Marx Wartofsky (1994 and 1995) reflect the personal aspects involved in this particular history of knowledge transfer and reception.

4.5 Felix Kaufmann, John Dewey and Edgar Zilsel: Between Phenomenology, Pragmatism and Sociology of Science

The *New School for Social Research* in New York, founded by Alvin Johnson in 1919, became a classical university of emigrants. In 1933, the University in Exile with its "Graduate Faculty of the Political and Social Science" offered a platform for German-speaking social scientists. In the following decades it also played an important role in the further development of the philosophy of science in New York. (On the history of the New School see [Rutkoff and Scott, 1986; Krohn, 1987]). A number of generations of American and Central European scientists had taught and studied on Fifth Avenue since the thirties, contributing to a project of modern social science with philosophical underpinnings [Srubar, 1988]. The Austrian contribution to this project, while limited to a few scholars, is significant for the transfer of scientific and philosophical ideas. Within the Unity of Science movement the history of science and social science were dealt with in an academic setting. The convergence of logical empiricism, phenomenology and neopragmatism became manifest in research, teaching and publications, most notably in the journals *Social Research* and *Philosophy and Phenomenological Research.* The Viennese mathematician, philosopher of law and social scientist Felix Kaufmann (1895–1949), the "phenomenologist of the Vienna Circle" played a seminal role in this history of reception which has received little attention to date. (On the life and work of Kaufmann [Zilian, 1990; Stadler, 1997a]).

Kaufmann had studied law and philosophy in Vienna. From 1922 to 1938 he was a lecturer of legal philosophy at the University of Vienna's School of Law, while at the same time he worked as a manager. He frequented a number of Viennese intellectual circles — from the Vienna Circle, the Kelsen School to the (Ludwig von) Mises Circle and the "Geist-Kreis" associated with F. A. von Hayek. Even before he emigrated, Kaufmann practiced interdisciplinary thought, mediating between various positions (e.g., between Husserl and the Vienna Circle or between understanding and explication). Because of his Jewish background and liberalism, he became part of the *Cultural Exodus* from Austria [Stadler and Weibel, 1995]. At the age of 43 he succeeded in securing an academic position at the New School in New York, together with his old Viennese friend Alfred Schütz: first as "associate" and from 1944 on as "full professor" for philosophy at the Graduate Faculty until his untimely death in 1949. He made an effort to come into contact with John Dewey and to discuss his ideas with him, but the latter did not cooperate as Kaufmann had hoped. He was also co-editor of the quarterly *Philosophy and Phenomenological Research*, an organ for interdisciplinary discussion, published from 1940 on, after the Dutch *Synthese* was discontinued at the outbreak of the war. The following members of the Editorial Board deserve mention since they were to a greater or lesser extent involved in this Unity of Science movement: C.J. Ducasse, Aron Gurwitsch, Charles Hartshorne, Wolfgang Köhler and Alexandre Koyré. This pluralism is also reflected in the names of the contributors to volume 6 of the journal (June 1946) whose articles were also related in a certain way: Gustav Bergmann, Rudolf Carnap, Horace M. Kallen, Felix Kaufmann, Alexandre Koyré, Richard von Mises, Ernest Nagel, Alfred Schütz and Donald Williams. The discussions ranged from induction and probability to the Unity of Science. Horace Kallen, dean and representative of "cultural pluralism" at the *New School*, contributed an impressive obituary on Otto Neurath [Kallen, 1946]. The volume also included shorter discussions and book reviews (e.g., on Kaufmann's *Methodology of the Social Science* by V. J. McGill or Reichenbach's *Philosophic Foundation of Quantum Mechanics* by Victor Lenzen) which rounded off this controversial discourse on philosophy in general and philosophy of science in particular. It comes as no surprise that the same volume includes Rudolf Carnap, Fritz Machlup, Ludwig von Mises, Günther Stern (= Günther Anders) among members of the "International Phenomenological Society" such as Felix Kaufmann and Alfred Schütz.

Austrian philosophers of science and social scientists were also represented in the official organ of the Graduate Faculty, *Social Research. An International Quarterly of Political and Social Science*, where they figured as authors and members of the Editorial Board (Felix Kaufmann and Ernst Karl Winter). There we find the *Graduate Faculty* lecture programs which have unfortunately not been considered up until now. The 1940/41 curriculum, for instance, includes a joint seminar on "Methodology of the Social Sciences" by Max Wertheimer, Gerhard Kolm, Kurt Riezler and Felix Kaufmann. The title of the seminar was later to become the title of a book published in 1944. In sociology, Kaufmann contributed "Forecast and Prediction in the Natural and Social Sciences", "Modern Philosophy and Value" and in philosophy, "The Logic of Pragmatism", "Analysis of Dewey's Logic". Later, Charles Morris and Otto Neurath's son Paul taught at the New School as visiting professors. Up until his death, Kaufmann published in the two journals named (as well as in the *Journal of Philosophy*), presenting above all his methodology of the social sciences within the context of a unified science. He made use of a methodology based on linguistic critique and phenomenological and analytical elements. In the year of his death, Kaufmann's article "The Issue of Ethical Neutrality in Political Science" appeared. This was followed, posthumously, by a long survey with the title "Basic Issues in Logical Positivism" [1950], which provided a sort of overview of the developent of the philosophy of science from Vienna to New York.

A short analysis of Kaufmann's lecture program in the forties — New School and Graduate Faculty — show that he tried to cover the areas of "Science and Philosophy", "History and Modern Theory of Knowledge", "Philosophical Introduction to Scientific Method" and even value theory. At the same time, the official philosophy figured centrally with Morris Cohen's skeptical contribution "Scientific Method" in the *Encyclopedia of the Social Sciences* for the *New School*, together with the contributions by Horace Kallen and Sidney Hook [Rutkoff and Scott 1986, 75ff].

> "Between them, Cohen, Kallen, and Hook made versions of pragmatism the unofficial philosophy of the New School, and the school's unofficial sponsorship of the Encyclopedia of the Social Sciences reinforced its advocacy. Together with Dewey, these three comprised the core of a distinguished group of New York philosophers who dominated American philosophy between the world wars." [Ibid., 78]

Here the (neo)pragmatic background of New York which was to play a significant role in the contact with Logical Empiricism is once again addressed. The subsequent turn to phenomenological social theory (lifeworld-oriented, "understanding" social science), represented most notably by Alfred Schütz, was not counted out here.

If one reads both accounts of the *New School*, one sees that Kaufmann's own contribution to the philosophy of science was marginal. At the same time, as a continental liberal in the tradition of German enlightenment (also as a student of Hans Kelsen), he was a central figure in the interdisciplinary discussion group on liberalism and democracy against the background of the Nazi catastrophe.

> "Felix Kaufmann, ... addressed these issues from a different perspective, taking issue with Horace Kallen's assertion that democracy could thrive only where there was a tradition of liberal politics. In his reexamination of Dewey's German Philosophy and Politics, Kaufmann discussed the ways in which he thought American philosophers, particularly Dewey and Santayana, had unfairly treated the tradition of German idealism. Kaufmann defended Kant's philosophical and ethical positions against Dewey, ... Kaufmann instead argued that Kant was the embodiment of the German Enlightenment, the philosopher of reason." [Rutkoff and Scott 1986, 137f.]

In spite of his early death, Kaufmann seems to have exerted a lasting influence on American academic life in the wake of the Schütz reception [Zilian, 1990].

A similar history of reception, but with a much more tragic turn, can be found in the case of another "scholar in exile", namely the Viennese mathematician, sociologist of science, and educator *Edgar Zilsel* (1891–1944).

Zilsel, one of the pioneers of an externalist history and philosophy of science was not able to find an adequate academic position after emigrating to the United States. He had to make ends meet under the most difficult circumstances at various colleges and with the help of insufficient grants (e.g., 1939–41, through Horkheimer's "Institute for Social Research"), before finally committing suicide out of desperation and exhaustion resulting from his work on his large project on the "Social Origins of Modern Science". (On his life and work cf. Haller/Stadler, eds. 1993, in particular the articles by Dahms and Fleck). Already in Vienna, Zilsel had started working on his ambitious project on *The Social Origins of Modern Science* [Krohn, 1976; Zilsel, 1990; Dahms, 1993]. The comparison of the original plan with its actual realization clearly shows how this innovative project was carried out. The existing parts of the study in German were expanded in English, the language in which they were then published [Raven *et al.*, 2000]. These fragments also include the study "Problems of Empiricism" [1941] which was integrated into the *Encyclopedia of Unified Science*. In view of his marginalized position, already evident in Vienna, but even more pronounced in the United States, Zilsel can certainly be described as a "case of failed transfer of knowledge" [Fleck, 1993], even if the reasons for this have yet to be analysed. American sociology of knowledge was not so much influenced by Zilsel's fundamental, fragmentary studies as by Robert K. Merton's work from the year of Zilsel's emigration (1938) on. This is true in spite of the fact that both Zilsel and Merton appeared at the "Fifth International Congress for the Unity of Science" at Harvard in 1939 and Zilsel published three years later his study on the problems of empiricism in the *Encyclopedia of Unified Science*. At the Harvard congress he summarized his studies as follows:

> "In the period from the end of the Middle Ages until 1600 the university scholars and the humanistic literati are rationally trained but they do not experiment as they despise manual labor. Many more or less plebian craftsmen experiment and invent but lack methodical rational training. About 1600, with the progress of technology, the experimental method is adopted by rationally trained scholars of the educated upper class. So the two components of scientific search are united at last: modern science is born. The whole process is embedded in the advance of early capitalist economy which weakens collective-mindedness, magical thinking, traditions, and the belief in authority, which furthers mundane, rational, and causal thinking, individualism and rational organization." [Zilsel, 1939]

It is only today that studies in the history of science have, in retrospect, shown the relevance of Zilsel's oeuvre in a larger context [Raven, 2000]:

> "In the early forties Edgar Zilsel published a number of important and well-known essays on the emergence of science. These essays have

> given rise to the so-called Zilsel thesis. But Zilsel published a couple of smaller and far less well-known essays. These essays are directed particularly against the efforts of South-West-German Neo-Kantianism..., Dilthey's philosophy of life, and interpretative sociology. ... His main argument is that philosophers from cultural science and the humanities proceed on the basis of a false understanding of natural science. It looks as though these two sets of essays do not seem to have much in common. Closer investigation of Zilsel's life and work reveals, however, that for Zilsel at least, there is an inner connection. The essays on the emergence of modern science are, in fact, a case study aimed at showing that law-like explanations in history are, indeed, possible; something that the other sets of essays argued in the abstract... We show how these two projects are not only complementary but in fact form part of an overarching motive of Zilsel: to argue the modernity of the socio-historical sciences."

Edgar Zilsel can be seen as a case of a slow, highly delayed transfer of knowledge through emigration. His findings have only been unearthed in contemporary history of science and science studies and now belong to the main stream together with Merton, Fleck, Kuhn and Feyerabend — even if one is often not aware of the background from which they emerged.

4.6 Epilogue: Continuity and Break in the Philosopy of Science — Boston, Pittsburgh, and Vienna

In North America the most important institutions continue to be the Philosopy of Science Association (PSA), the Boston Center and especially the Center for Philosophy of Science at the University of Pittsburgh, which was founded in 1960 by the philosopher of science Adolf Grünbaum, inspired by Feigl's Minnesota Center. (`http://www.pitt.edu/~pittcntr`).

Adolf Grünbaum (born 1923 in Köln, Germany) — currently the Andrew Mellon Professor of Philosopy of Science, Research Professor of Psychiatry, and Chairman of the Center for Philosophy of Science at the University of Pittsburgh — is one of the most distinguished and influential scholars working on philosophy of physics (space and time), the theory of scientific rationality, the philosophy of psychiatry, and the critique of theism in the tradition of Logical Empiricism (Carnap, Feigl, Hempel, and above all Hans Reichenbach). As the President of the International Union for the History and Philosophy of Science (2006/07) he continues his outstanding academic and scholarly career in this field [Cohen and Laudan 1983/1992; Earman *et al.*, 1993]. Because of its extraordinary importance the Pittsburgh Center will be desribed separately by the author on the occasion of the transfer of Adolf Günbaum's private papers to the Vienna Circle Institute, where the Robert S. Cohen collection is already located.

Regarding the long neglected return of philosopy of science back to Europe after World War II, research has been launched only in the recent years.

This is in contrast to the recent decades, where analytic philosophy and philosophy of science has become a paradigm for research and teaching in philosophy, also in the German speaking world. With the forced emigration (principally to the USA and UK) of the Vienna, Berlin and Prague Circles, representatives of Logical Empiricism disappeared almost entirely from Germany and Austria. None of them returned after the war, for there were no official invitations to do so. Nevertheless, it is possible to reconstruct some aspects of the transformation and belated return of the philosophy of science to the places of its Central European origins. The focal point of investigation is directed at some philosophers of science, who were mainly responsible for its transfer, transformation and retroactive development: Rudolf Carnap, Herbert Feigl, Wolfgang Stegmüller and the members of a Viennese post-war discussion circle around Viktor Kraft, with Paul Feyerabend and the US-Visiting Professor Arthur Pap. Feigl was the first member of the Vienna Circle to emigrate (in 1931) to the USA and to introduce Logical Empiricism into American academia. Through his contacts to European scholars after the war, he — together with Carnap — were of most importance for introducing philosophy of science in Austria and Germany. In parallel, the Viennese group around Viktor Kraft and Bela Juhos (the "Third Vienna Circle") was another attempt to revive the banished philosophy of science in the country of its origin. In the context of a hostile atmosphere Kraft tried to re-establish lost contacts and to take up international developments. Another proponent of the Kraft Circle, Wolfgang Stegmüller, succeeded because of his philosophical commitment not in Austria, but in Munich, where he founded a school of philosophy of science in close contact with Carnap and Feigl, a school which continues to be influential to this day.

We can speak of the forgotten "Third Vienna Circle", as a so far hidden story of the survival and return of philosophy of science in the Cold War phase. (Fischer/Stadler 2006).

This process was initiated by Viktor Kraft (1880–1975), who — after being dismissed by the Nazis in 1938 and working in inner emigration during the war — founded and led the so called "Kraft-Kreis" (Kraft-Circle) 1949-1953, and who contacted again some former members of the Vienna Circle (Herbert Feigl, Philipp Frank, Rudolf Carnap), and Karl Popper.

This discussion group at the Vienna based "Institut für Wissenschaft und Kunst" as well as the "Austrian College/Forum Alpbach", both still existing, was a remarkably short renaissance of the Viennese heritage in the Philosophy of Science. It exerted influence on the second wave of emigré philosophers of science after World War II — like Paul Feyerabend, who wrote his dissertation "Zur Theorie der Basissätze" under Kraft; Ernst Topitsch, who took over a chair at the University of Heidelberg; and Wolfgang Stegmüller, who succeeded at the University in Munich after being rejected by the Universities of Innsbruck and Vienna. The main results of this Circle are documented in the *Festschrift* of Kraft [1960]. Ten years later another re-transfer of philosophy of science took place at the "Institute for Advanced Study" in Vienna with Carnap, Feigl and Popper as visiting lecturers.

A decisive event in this context was Kraft's invitation of Arthur Pap as a visiting professor to Vienna in 1953/54, where he published the book *Analytische Erkenntnistheorie* [1955] with the assistance of Feyerabend, and dedicated to the Vienna Circle. Together with Feigl's article "Existential Hypotheses" [1950] this book formed the philosophical background to be debated controversially in the "Kraft Circle" (*inter alia* with Elisabeth Anscombe, Walter Hollitscher, Bela Juhos, and, by the way, with one appearance of Ludwig Wittgenstein). The main issue on the agenda was realism, especially the existence of an external world.

From a philosophical point of view the central debate was on realism vs. phenomenalism in the philosophy of science, to be continued at Feigl's "Minnesota Center for Philosophy of Science", and which obviously influenced the participating Feyerabend. From a broader perspective of the *Methodenstreit* we can identify the dualism of the hypothetico-deductive (critical or constructive) realism and inductive phenomenalism. On the one hand, the transfer of this controversy to England and America obscured its origins in the "Third Vienna Circle" and was later on overshadowed by Karl Popper's dominant preference for realism and objectivism. On the other hand, the return and modified transformation of Analytic Philosophy and Philosophy of Science back to the German speaking countries was realized by Wolfgang Stegmüller as a late consequence of the Forum Alpbach, and the Kraft Circle.

In the long run, the founding of the Vienna based "Institut Wiener Kreis" (Vienna Circle Institute) in 1991 is a late outcome of the emigration and return of the philosophy of science from the city of its origin. (`http://www.univie.ac.at/ivc`). And by the end of 2006 a "European Philosophy of Science Association" (EPSA) was founded in Vienna — as a sort of counterpart and partner of the Philosophy of Science Association (PSA), which has been the successful American institution for the philosophy of science since 1934.

5 CONCLUDING REMARKS

1. The transfer, transformation and impact of Central European, in particular Austrian, German and Polish philosophy of science in the period 1930–1960 did not take place abruptly. Rather, it involved a continuous brain drain which was reinforced by the mass exodus that set in around 1938. The early ties to Anglo-American philosophy of science prepared the ground for a pronounced convergence between "Wissenschafslogik" and the history and philosophy of science in the United States. Already in the twenties, there was a trend towards internationalization. With the dominance of neo-pragmatism/behaviorism in the context of American philosophy, the "*Wiener Kreis* in America" [Feigl, 1968] became relatively successful.

2. In spite of the fact that there was no significant intellectual remigration after 1945, there was a considerable re-transfer of knowledge primarily influenced by the contribution of former emigrants. The belated (re)discovery of the

philosophy of science in this context took place in Austria's Second Republic in connection with the paradigm of "Austrian philosophy". However, there has also been an increasing disciplinary specialization and autonomization of the *Wissenschaftstheorie* which no longer could be seen as reflecting the model of a comprehensive history and philosophy of science. The fact that the emigrated philosophers of science did not return to Central Europe was also related to the phenomenon that they had been completely uprooted from the German-speaking world.

3. From an intellectual perspective, it is striking that the transformation of the philosophy of science as a break with the so called *Received View* was already anticipated in the development of the history and philosophy of science in the thirties — long before Kuhn's path breaking book on *The Structure of Scientific Revolutions* [1962/1970]. Here one already finds a number of themes and methodological principles more or less anchored in the program of the *Encyclopedia of Unified Science,* from 1930–1960, as a result of the pragmatic, historical and naturalistic turns in the philosophy of science. This was accompanied by a development towards pluralism and relativism.

4. For historiography, these findings suggest a need to bring together exile and emigration studies with history of science research (including psychology and sociology of science) within the framework of contemporary history [Stadler, 1998/2001]. One of the most significant insights of this new perspective is the futility of linear cause-effect models given the fact that it is impossible to detect qualitative "units of impact" in the cognitive sphere. Science must be regarded as a largely complex, self-organizing project within a sociopolitical context. The above mentioned persons and institutions, along with a comparative account of the international history of the disciplines, provide the basic elements for a historical study of science and its philosophy. This research, however, cannot dispense with the theoretical core, i.e., representative scientific texts. The complementarity of text and context is thus postulated for historical studies.

5. In this sense, history of science can be regarded as a constitutive part of an interdisciplinary historiography. It must address from a common perspective disparate fields of human, social and natural sciences as cultural phenomena. If one draws conclusions from the *Cultural Exodus*, it becomes clear how positively factors such as migration, mobility and internationality influenced the development of philosophy and science. To do justice to history, the development from *Wissenschaftslogik*, via *Philosophy of Science*, to today's *Wissenschaftstheorie*, must be contextualized.

ACKNOWLEDGEMENT

This article is part of a research project on "The Banishment and Return of the Philosophy of Science", funded by the Austrian Science Fund (FWF), P18066-G04: `http://www.univie.ac.at/ivc/stegmueller`.

BIBLIOGRAPHY

[Achinsten and Barker, 1969] P. Achinstein and S. F. Barker, eds. *The Legacy of Logical Positivism.* Studies in the Philosophy of Science, Baltimore: The Johns Hopkins Press, 1969.

[Albert, 1994] H. Albert. Wissenschaft in Alpbach, in *Forum Alpbach*, pp. 17-22, 1994.

[Ash and Sellner, 1996] M. Ash and A. Sellner, eds. *Forced Migration and Scientific Change. Emigre German-Speaking Scientists and Scholars after 1933.* Cambridge: Cambridge University Press, 1996.

[Ash and Sellner, 1996b] M. Ash and A. Sellner. Forced Migration and Scientific Change after 1933, in: [Ash and Sellner, 1996, pp. 1–22].

[Auer *et al.*, 1994] A. Auer, Behrendt, Flora, and Knapp, eds. *Das Forum Alpbach 1945-1994. Die Darstellung einer Europäischen Zusammenarbeit.* Hrsg. von Alexander Auer. Wien: Ibera Verlag European University Press, 1994.

[Ayer, 1936a] A. J. Ayer. *Language, Truth and Logic*, London: Victor Gollancz Ltd, 1936. (15th impression 1955).

[Ayer, 1936b] A. J. Ayer. The Analytic Movement in Contemporary Philosophy, in: *Actes du Congres International de Philosophie Scientifique* VIII, Paris: Hermann & Cie, 1936.

[Ayer, 1959] A. J. Ayer.*Logical Positivism.*Glencoe, Ill.: The Free Press, 1959.

[Ayer, 1970] A. J. Ayer. *Sprache, Wahrheit und Logik.* Hrsg. van H. Herring. Stuttgart: Reclam, 1970.

[Ayer, 1982] A. J. Ayer. *Philosophy in the Twentieth Century*, London: Weidenfeld and Nicolson, 1982.

[Ayer *et al.*, 1956] A. Ayer, W. C. Kneale, G. A. Paul, D. F. Pears, P. F. Strawson, G. J. Warnock and R. A. Wollheim. . *The Revolution in Philosophy.* With an Introduction by Gilbert Ryle. London: Macmillan & Co. Ltd, 1956.

[Bernard and Stadler, 1997] J. Bernard and F. Stadler, (Hrsg). *Neurath: Semiotische Projekte und Diskurse.* Wien: OGS, 1997.

[Black, 1939/40] M. Black. Relations between Logical Positivism and the Cambridge School of Analysis, *Erkenntnis/Journal of Unified Science* VIII, 24-35, 1939/40.

[Black *et al.*, 1972] M. Black, E. H. Gombrich, and J. Hochberg. *Art, Perception, and Reality*, Baltimore-London: The Johns Hopkins Press, 1972.

[Blackmore, 1992] J. Blackmore, ed. *Ernst Mach - A Deeper Look. Documents and New Perspectives.* Dordrecht-Boston-London: Kluwer, 1992.

[Blumberg and Feigl, 1931] A. E. Blumberg and H. Feigl. Logical Positivism, *Journal of Philosophy* 28, 281-296, 1931.

[Brunswik, 1952] E. Brunswik. *The Conceptual Framework of Psychology*, in: Neurath/Carnap/Morris (Eds.) 1970, 655-760, 1952.

[Buck and Cohen, 1970] R. C. Buck and R. S. Cohen, eds. *PSA 1970. In Memory od Rudolf Carnap.* Dordrecht-Boston: Reidel, 1970.

[Carnap, 1932] R. Carnap. Oberwindung der Metaphysik durch logische Analyse der Sprache, in: *Erkenntnis* II, 91-105, 1932.

[Carnap, 1934] R. Carnap. *The Unity of Science.* Translated with an Introduction by M. Black, London: Kegan Paul, 1934.

[Carnap, 1934a/1968] R. Carnap. *Logische Syntax der Sprache.* Wien: Springer, 1934a/1968.

[Carnap, 1934b] R. Carnap. *Die Aufgabe der Wissenschaftslogik.* Wien: Gerold & Co, 1934.

[Carnap, 1934c] R. Carnap. *The Unity of Science.* London: Kegan Paul, 1934.

[Carnap, 1935] R. Carnap. *Philosophy and Logical Syntax*, London: Kegan Paul, 1935.

[Carnap, 1936] R. Carnap. Von der Erkenntnistheorie zur Wissenschaftslogik, in: *Actes de Congres International de Philosophie Scientifique* I. Paris: Hermann & Cie., 36-41, 1936.

[Carnap, 1937] R. Carnap. *The Logical Syntax of Language.* London: Kegan Paul, 1937.

[Carnap, 1942] R. Carnap. *Introduction to Semantics.* Cambridge, Mass.:Harvard University Press, 1942.

[Carnap, 1963] R. Carnap. Intellectual Autobiography, in *The Philosophy of Rudolf Carnap*, P. A. Schilpp, ed., pp. 1–84. La Salle, Ill.: Open Court, 1963.

[Carnap, 1993] R. Carnap. *Mein Weg in die Philosophie.* Übersetzt und mit einem Nachwort sowie einem Interview hrsg. von Willy Hochkeppel. Stuttgart: Reclam, 1993.

[Carnap *et al.*, 1938] R. Carnap, Ch. Morris, and O. Neurath, eds. *International Encyclopedia of Unified Science* 1-19, 1938. Reprint: *Foundations of the Unity of Science*, Chicago and London: University of Chicago Press, 2 Volumes, 1970-71.

[Cohen, 1994] I. B. Cohen. *Revolutionen in der Naturwissenschaft.* Frankfurt/M.: Suhrkamp, 1994.

[Cohen, 1956] R. S. Cohen. Alternative Interpretations of the History of Science, in: Frank (ed.), 218-132, 1956.

[Cohen, 1963a] R. S. Cohen, ed. *Boston Studies in the Philosophy of Science.* Dordrecht-Boston-London: Kluwer, 1963.

[Cohen and Laudan, 1983] R. S. Cohen and L. Laudan, eds. *Physics, Philosophy and Psychoanalysis. Essays in Honor of Adolf Grünbaum.* Dordrecht-Boston-Lancaster: D.Reidel, 1983.

[Cohen and Wartofsky, 1963b] R. S. Cohen and M. W. Wartofsky. Preface, in: Cohen (Ed.), Vllf, 1963.

[Cohen and Wartofsky, 1965a] R. S. Cohen and M. W. Wartofsky. Preface, in: Cohen and Wartofsky, VII, 1965b.

[Cohen and Wartofsky, 1965b] R. S. Cohen and M. W. Wartofsky, eds. *Boston Studies in the Philosophy of Science. Volume Two: In Honor of Philipp Frank.* New York: Humanities Press, 1965.

[Cohen and Wartofsky, 1985] R. S. Cohen and M. W. Wartofsky, eds. *A Portrait of Twenty-Five Years.* Boston Colloquium for the Philosophy of Science 1960-1985. Dordrecht-Boston-London: Reidel, 1985.

[Conant, 1951] J. B. Conant. *Science and Common Sense.* New Haven:Yale University Press, 1951.

[Coser, 1984] L. Coser. *Refugee Scholars in America. Their Impact and their Experiences.* New Haven-London: Yale University Press, 1984.

[Coser, 1988] L. Coser. Die 6sterreichische Emigration als Kulturtransfer Europa - Amerika, in: Stadler (Hrsg.), 93-101, 1988.

[Creath, 1990] R. Creath, ed. *Dear Carnap, Dear Van: The QuineCarnap Correspondence and Related Work.* Berkeley-Los AngelesLondon: University of California Press, 1990.

[Dahms, 1987] H.-J. Dahms. Die Emigration des Wiener Kreises, in: Stadler (Hrsg.), 66-122, 1987.

[Dahms, 1988] H.-J. Dahms. Die Bedeutung der Emigration des Wiener Kreises für die Entwicklung der Wissenschaftstheorie in: Stadler (Hrsg.), 155-168, 1988.

[Dahms, 1993] H.-J. Dahms. Edgar Zi1sels Projekt 'The Social Roots of Science' und seine Beziehungen zur Frankfurter Schule, in: Haller and Stadler (Hrsg.), 474-500, 1993.

[Dahms, 1994] H.-J. Dahms. *Positivismusstreit. Die Auseinandersetzungen der Frankfurter Schule mit dem logischen Positivismus, dem amerikanischen Pragmatismus und dem kritischen Rationalismus.* Frankfurt/M.: Suhrkamp, 1994.

[Dahms, 1997] H.-J. Dahms. Positivismus, Pragmatismus, Enzyk1opadieprojekt, Zeichentheorie, in: Bernard/Stadler (Hrsg.), 1997.

[Danneberg *et al.*, 1994] L. Danneberg, A. Kamiah and L. Schafer (Hrsg.). *Hans Reichenbach und die Berliner Gruppe.* Braunschweig: Vieweg, 1994.

[Dawson, 1997] J. W. Dawson. *Logical Dilemmas. The Life and Work of Kurt Gödel.* Wellesley, MA: A K Peters, 1997.

[De-Pauli-Schimanovich *et al.*, 1995] W. DePauli-Schimanovich, E. Kohler, and F. Stadler, eds. *The Foundational Debate. Complexity and Constructivity in Mathematics and Physics.* Dordrecht-Boston-London: Kluwer, 1995.

[Dewey, 1951] J. Dewey. *Theory of Valuation.* The University of Chicago Press. (= *International Encyclopedia of Unified Science* II/4), 1951.

[Earman *et al.*, 1993] J. Earman, A. I. Janis, G. J. Massey, and N. Rescher, eds. *Philosophical Problems of the Internal and External Worlds. Essays on the Philosophy of Adolf Grünbaum.* University of Pittsburgh Press and Universitätsverlag Konstanz, 1993.

[Edmonds and Eidinow, 2001] D. J. Edmonds and J. A. Eidinow. *Wittgenstein's Poker. The Story of a Ten-Minute Argument between two Great Philosophers*, London: Faber and Faber, 2001. German edition: *Wie Wittgenstein Karl Popper mit dem Feuerhaken drohte. Eine Ermittlung,* Stuttgart-München: DVA *Einheitswissenschaft* 1992. Hrsg. von Joachim Schulte und Brian McGuinness. Mit einer Einleitung von Rainer Hegselmann. Frankfurt/M.: Suhrkamp.

[Faludi, 1986] A. Faludi. *Critical Rationalism and Planning Methodology.* London: Pion Limited, 1986.

[Farber, 1950] M. Farber. *Philosophic Thought in France and the United states. Essays Representing Major Trends in Contemporary French and American Philosophy.* New York: University of Buffalo Publ, 1950.

[Feigl, 1936] H. Feigl. Sense and Nonsense in Scientific Realism, in: *Actes du Congres International de Philosophie Scientifique* III. Paris: Hermann & Cie., 50-56, 1936.

[Feigl, 1956a] H. Feigl. Preface, in: Feigl and Scriven (Eds.), V-VII, 1956.

[Feigl, 1956b] H. Feigl. Some Major Issues and Developments in the Philosophy of Science of Logical Empiricism, in: Feigl and Scriven (Eds.), 3-37, 1956.

[Feigl, 1968] H. Feigl. The *Wiener Kreis* in America, in *Charles Warren Center for Studies in American History*, ed., Perspectives in American History, Vol. II: Harvard University Press, 630-673, 1968.

[Feigl, 1969] H. Feigl. The *Wiener Kreis* in America, in Fleming, D. and Bailyn, B., eds., The *Intellectual Migration. Europe and America, 1930 – 1960.* Cambridge, Mass.: Harvard University Press, 630-673, 1969. Reprinted in Feigl 1981, 57-94.

[Feigl, 1981] H. Feigl. *Inquiries and Provocations. Selected Writings, 1929-1974.* Ed. by Robert S. Cohen. Dordrecht-Boston-London: Reidel, 1981.

[Feigl and Blumberg, 1931] H. Feigl and A. Blumberg. Logical Positivism. A New Movement in European Philosophy, in: *Journal of Philosophy* 28, 281-296, 1931.

[Feigl and Brodbeck, 1953] H. Feigl and M. Brodbeck, eds. *Readings in the Philosophy of Science.* New York: Appleton-Century-Crofts, 1953.

[Feigl and Scriven, 1956] H. Feigl and M. Scriven, eds. Minnesota Studies in the Philosophy of Science. Volume I: *The Foundations of Science and the Concepts of Psychology and Psychoanalysis.* Minneapolis: University of Minnesota Press, 1956.

[Feigl and Sellars, 1949] H. Feigl and W. Sellars, eds. *Readings in Philosophical Analysis.* New York: Appleton-Century-Crofts, 1949.

[Felderer, 1993] B. Felderer (Hrsg.) *Wirtschafts- und Sozialwissenschaften zwischen Theorie und Praxis. 30 Jahre Institut far Höhere Studien in Wien.* Heidelberg: Physica-Ver1ag, 1993.

[Felt *et al.*, 1995] U. Felt, H. Nowotny, and K. Taschwer. *Wissenschaftsforschung. Eine Einführung.* Frankfurt/M.-New York: Campus, 1995.

[Feyerabend, 1955] P. K. Feyerabend. Wittgenstein's 'Philosophical Investigations', in: The *Philosophical Review* 64, 449-483, 1955. Auch in: Feyerabend 1981, 293-325.

[Feyerabend, 1966] P. K. Feyerabend. Herbert Feigl: A Biographical Sketch, in: [Feyerabend and Maxwell, 1966, pp. 3–16].

[Feyerabend, 1978] P. K. Feyerabend. *Der wissenschaftstheoretische Realismus und die Autorität der Wissenschaften. Ausgewählte Schriften*, Band 1. Braunschweig-Wiesbaden, 1978.

[Feyerabend, 1981] P. K. Feyerabend. *Probleme des Empirismus. Schriften zur Theorie der Erklärung, der Quantentheorie und der Wissenschaftsgeschichte. Ausgewählte Schriften*, Band 2. Braunschweig-Wiesbaden: Vieweg, 1981.

[Feyerabend, 1995] P. K. Feyerabend. *Zeitverschwendung.* Frankfurt/M.: Suhrkamp, 1995.

[Feyerabend and Maxwell, 1966] P. K. Feyerabendand G. Maxwell, eds. *Mind, Matter, and Method. Essays in Philosophy and Science in Honor of Herbert Feigl.* Minneapolis: University of Minnesaota Press, 1966.

[Fischer, 1995] K. R. Fischer (Hrsg.). *Das goldene Zeitalter der Österreichischen Philosophie. Ein Lesebuch.* Wien: WUV-Verlag, 1995.

[Fisher and Stadler, 1997] K. R. Fischer and F. Stadler (Hrsg.). *'Wahrnehmung und Gegenstandswelt'. Zum Lebenswerk von Egon Brunswik (1903-1955).* Wien-New York: Springer, 1997.

[Fischer and Stadler, 2006] K. R. Fischer and F. Stadler (Hrsg.). *Paul Feyerabend – Ein Philosoph aus Wien.* Wien-New York: Springer, 2006.

[Fischer and Wimmer, 1993] K. R. Fischer and F. M. Wimmer (Hrsg.). *Der geistige Anschluß. Philosophie und Politik an der Universität Wien 19301950.* Wien: WUV-Verlag, 1993.

[Fleck, 1993] C. Fleck. Marxistische Kausalanalyse und funktionale Wissenschaftssozio1ogie. Ein Fall unterbliebenen Wissenstransfers, in: Haller/Stadler (Hrsg.), 501-524, 1993.

[Fleming and Bailyn, 1969] D. Fleming and B. Bailyn, eds. *The Intellectual Migration. Europe and America, 1930-1960.* Cambridge: Harvard University Press, 1969.

[Frank, 1947] P. Frank. *Einstein - His Life and Times.* New York: Knopf, 1947.

[Frank, 1949] P. Frank. *Modern Science and its Philosophy.* Harvard University Press, 1949.

[Frank, 1950] P. Frank. *Relativity – A Richer Truth.* Preface by Albert Einstein. Boston: Beacon Press, 1950. German edition: Zürich: Pan Verlag 1952.

[Frank, 1952] P. Frank. *Wahrheit - re1ativ oder abso1ut?* Mit einem Vorwort von Albert Einstein. ZUrich: Pan-Verlag, 1952.

[Frank, 1956] P. Frank, ed. *The Validation of Scientific Theories.* Boston: The Beacon Press, 1956.

[Frenkel-Brunswik, 1996] E. Frenkel-Brunswik. *Studien zur autoritaren Persönlichkeit.* Hrsg. und eingelitet von Dietmar Paier. Graz: Nausner & Nausner, 1996.

[Fuller, 2000] S. Fuller. Thomas Kuhn. *A Philosophical History for Our Time.* University of Chicago Press, 2000.

[Galison and Stump, 1996] P. Galison and D. Stump, eds. *The Disunity of Science.* Stanford University Press, 1996.

[Gavroglu *et al.*, 1995] K. Gavroglu, J. Stachel and M. W. Wartofsky, eds. *Physics, Philosophy and the Scientific Community; Science Politics and Social Practice; Science Mind and Art.* 3 Vlms. In Honor of Robert S. Cohen. Dordrecht-Boston-London: Kluwer, 1995.

[Giere, 1996] R. N. Giere. From Wissenschaftliche Philosophie to Philosophy of Science, in: [Giere and Richardson, 1996, pp. 335–354].

[Giere and Richardson, 1996] R. N. Giere and A. W. Richardson, eds. Minnesota Studies in the Philosophy of Science. Volume XVI: *Origins of Logical Empiricism.* Minneapolis-London: University of Minnesota Press, 1996.

[Gould and Cohen, 1994] C. C. Gould and R. S. Cohen, eds. *Artifacts, Representations and Social Practice. Essays for Marx Wartofsky.* Dordrecht-Boston-London: Kluwer, 1994.

[Gower, 1987] B. Gower, ed. *Logical Positivism in Perspective. Essays on Language, Truth and Logic.* London-Sydney: Croom Helm, 1987.

[Hacker, 1996] P. M. S. Hacker. *Wittgenstein's Place in Twentieth Century Analytic Philosophy,* Oxford: Basil Blackwell, 1996.

[Hacohen, 2000] M. Hacohen. *Karl Popper – The Formative Years, 1902 – 1945. Politics and Philosophy in Interwar Vienna,* Cambridge University Press, 2000.

[Haller, 1988] R. Haller. Die philosophische Entwicklung in Osterreich am Beginn der Zweiten Republik, in: [Stadler, 1987/88, pp. 157–180].

[Haller, 1993] R. Haller. *Neopositivismus. Eine historische Einführung in die Philosophie des Wiener Kreises.* Darmstadt: Wissenschaftliche Buchgemeinschaft, 1993.

[Haller and Stadler, 1988] R. Haller and F. Stadler (Hrsg.). *Ernst Mach - Werk und Wirkung.* Wien: Hölder-Pichler-Tempsky, 1988.

[Haller and Stadler, 1993] R. Haller and F. Stadler (Hrsg.). *Wien-Berlin-Prag. Der Aufstieg der wissenschaftlichen Philosophie.* Wien: Hölder-Pichler-Tempsky, 1993.

[Hardcastle and Richardson, 2004] G. Hardcastle and A. W. Richardson, eds. *Logical Empiricism in North America.* Minneapolis-London: University of Minnesota Press, 2004.

[Hark, 2004] M. ter Hark. *Popper, Otto Selz and the Rise of Evolutionary Epistemology.* Cambridge: Cambridge University Press, 2004.

[Hanisch, 1995] E. Hanisch. *Der lange Schatten des Staates. Österreichische Gesellschaftsgeschichte im 20. Jahrhundert.* Wien: Ueberreuter, 1995.

[Hayek, 1942] F. A. Hayek. Scientism and the Study of Society, *Economica*, IX-XI, 1942.

[Hayek, 1944] F. A. Hayek. *The Road to Serfdom*, London: Routledge, 1944. (Fiftieth Anniversary Edition: University of Chicago Press 1994).

[Hayek, 1945] F. A. Hayek and O. Neurath. Correspondence 1945. Vienna Circle Archives, Haarlem (NL).

[Hardcastle, 1996] G. Hardcastle. The Science Of Science Discussion Group at Harvard, 1940-41. Abstract. First International HOPOS Conference, Roanoke, Virgina, 19-21 April 1996, 24f.

[Hegselmann, 1988] R. Hegselmann. Alles nur Mißverständnisse? Zur Vertreibung des Logischen Empirismus aus Osterreich und Deutschland, in: Stadler (Hrsg.), 188-203, 1988.

[Heidelberger and Stadler, 2002] M. Heidelberger and F. Stadler, eds. *History of Philosophy of Science. New Trends and Perspectives*. Dordrecht-Boston-London: Kluwer 2002.

[Helbling and Wagnleitner, 1992] W. Helbling and R. Wagnleitner, eds. *The European Emigrant Experience in the U.S.A.* Tübingen: Gunter Narr, 1992.

[Hintikka and Puhl, 1995] J. Hintikka and K. Puhl, eds. *The British Tradition in 20th Century Philosophy*. Vienna: Holow-Pichler-Tempsky, 1995.

[Holton, 1993] G. Holton. From the Vienna Circle to Harvard Square: The Americanization of a European World Conception, in: [Stadler, 1993, pp. 47–74].

[Holton, 1993b] G. Holton. *Science ant Anti-Science*. Cambridge, Mass.: Harvard University Press, 1993.

[Holton, 1994] G. Holton. *Thematic Origins of Scientific Thought. Kepler to Einstein*. Revised Edition. Harvard University Press, 1994.

[Holton, 1995] G. Holton. On the Vienna Circle in Exile: An Eyewitness Report, in: [DePauli-Schimanovich *et al.*, 1995, pp. 269]-292].

[Hughes, 1975] H. S. Hughes. *The Sea Change: The Migration of Social Thought, 1930-1965.* New York, 1975

[Hughes, 1961] H. S. Hughes. *Consciousness and Society. The Reorientation of European Social Thought 1890 – 1930,* New York: Vintage Books, 1961.

[Hughes, 1983] H. S. Hughes. Social Theory in a New Context, in: [Jackman and Borden, 1983, pp. 111-122].

[Jackman and Borden, 1983] J. C. Jackman and C. M. Borden, eds. *The Muses Flee Hitler. Cultural Transfer and Adaption 1930-1945.* Washington, D.C.: Smithsonian Institution Press, 1983.

[Juhos, 1965] B. Juhos. Gibt es in Osterreich eine wissenschaftliche Philosophie?, in: *Österreich - Geistige Provinz?,* pp. 232-244, 1965.

[Kallen, 1946] H. M. Kallen. Postscript: Otto Neurath, 1882-1945, in: *Philosophy and Phenomenological Research.* VI/4, 529-533, 1946.

[Kamlah, 1983] A. Kamlah. Die philosophiegeschichtliche Bedeutung des Exils (nichtmarxistischer) Philosophen zur Zeit des Dritten Reiches, in: *Dia1ektik* 7, 29-43, 1983.

[Katz, 1991] B. Katz. The Acculturation of Thought: Transformations of the Refugee Scholar in America, in: *Journal of Modern History* 63, 740-752, 1991.

[Kaufmann, 1936] F. Kaufmann. *Methoden1ehre der Sozialwissenschaften*. Wien: Springer, 1936.

[Kaufmann, 1944] F. Kaufmann. *Methodology of the Social Sciences*. New York-London: Oxford University Press, 1944.

[Kaufmann, 1949] F. Kaufmann. The Issue of Ethical Neutrality in Political Science, in: *Social Research* 16, 344-352, 1949.

[Kaufmann, 1950] F. Kaufmann. Basic Issues in Logical Positivism, in: [Farber, 1950, pp. 565-588].

[Kaufmann, 1978] F. Kaufmann. *The Infinite in Mathematics. Logicomathematical Writings.* Ed. by Brian McGuinness. With an Introduction by Ernest Nagel. Dordrecht-Boston-London: Reidel, 1978.

[Kazemier and Vuysje, 1962] B. H. Kazemier and D. Vuysje, eds. *Logic and Language. Studies Dedicated to Professor Carnap on the Occasion of his Seventeenth Birthday*. Dordrecht: Reidel, 1962.

[Kendler, 1989] H. Kend1er. The Iowa Tradition, in: *American Psychologist* 44/8, 1124-1132, 1989.

[Kitcher, 2001] P. Kitcher. *Science, Truth, and Democracy*. Oxford University Press, 2001.

[Koertge, 1998] N. Koertge, ed. *A House Built on Sand. Exposing Postmodernist Myths about Science*. Oxford University Press, 1998.

[Koppelberg, 1987] D. Koppelberg. *Die Aufhebung der analytischen Philosophie. Quine als Synthese von Carnap und Neurath.* Frankfurt/M.: Suhrkamp, 1987.

[Kraft, 1950] V. Kraft. *Der Wiener Kreis. Der Ursprung des Neopositivismus. Ein Kapitel der jüngsten Philosophiegeschichte.*Wien: Springer, 1950.

[Kroner, 1988] H.-P. Kroner. Über1egungen zur Wirkungsgeschichte der deutschsprachigen wissenschaft1ichen Emigration, in: [Stadler, 1988, pp. 82-92].

[Krohn, 1987] C.-D. Krohn. *Wissenschaft im Exil. Deutsche Sozial und Wirtschaftswissenschaftler in den USA und die New School for Social Research.* Frankfurt/M.: Campus, 1987.

[Langer, 1988] J. Langer, (Hrsg.). *Geschichte der österreichischen Soziologie. Konstituierung, Entwicklung und europäische Bezüge*. Wien: Verlag für Gese11schaftskritik, 1988.

[Lazarsfeld, 1993] P. F. Lazarsfeld. The Pre-History of the Vienna Institute for Advanced Studies (1973), in: [Fe1derer, 1993, pp. 9-50].

[Lehrer and Marek, 1997] K. Lehrer and J. C. Marek, eds. *Austrian Philosophy. Past and Present. Essays in Honor of Rudolf Haller*. Dordrecht-Boston-London: K1uwer, 1997.

[Leinfellner, 1988] W. Leinfellner. Oskar Morgenstern, in: [Stadler, 1988, pp. 416-424].

[Leinfellner, 1993] W. Leinfellner. Der Wiener Kreis und sein Einf1uß auf die Sozia1wissenschaften, in: [Haller and Stadler, 1993, pp. 593-618].

[Leinfellner *et al.*, 1997] W. Leinfellner, *et al*, eds. *Game Theory, Experience, Rationality. In Honor of John C. Harsanyi.* Dordrecht-BostonLondon: K1uwer, 1997.

[Losee, 1980] J. Losee. *A Historical Introduction to the Philosophy of Science*. Oxford-New York-Toronto-Me1bourne: Oxford University Press, 1980.

[Mach, 1905] E. Mach. *Erkenntnis und Irrtum. Skizzen zur Psychologie der Forschung*. Leipzig: J.A.Barth, 1905. English: *Knowledge and Error. Sketches on the Psychology of Enquiry*. With an Introduction by Erwin N. Hiebert. Dordrecht-Boston 1976.

[Misak, 1995] C. J. Misak. *Verificationism. Its History and Prospects.* London-New York: Routledge, 1995.

[Marin, 1978] B. Marin. *Politische Organisation sozialwissenschaftlicher Forschungsarbeit. Fallstudie zum Institut für Höhere Studien*. Wien: Braumüller, 1978.

[McCulloch, 1956] W. S. McCulloch. Mysterium Iniquitatis - Of Sinful Man Aspiring into the Place of God, in: [Frank, 1956, pp. 159–170].

[Menger, 1934] K. Menger. *Moral, Wille und Weltgestaltung. Grundlegung zur Logik der Sitten.* Wien: Springer, 1934. Reprint: Hrsg. von Uwe Czaniera, Frankfurt/M: Suhrkamp 1997.

[Menger, 1994] K. Menger. *Reminiscences of the Vienna Circle and the Mathematical Colloquium.* Ed. by L. Golland/B. McGuinness/A. Sklar. Dordrecht-Boston-London: Kluwer, 1994.

[Misak, 1995] C. J. Misak. *Verificationism. Its History and Prospects.* London-New York: Routledge, 1995.

[Mises, 1951] R. von Mises. *Positivism. An Essay in Human Understanding.* New York: Dover. Ursprünglich 1939 erschienen unter dem Titel: *Kleines Lehrbuch des Positivismus. Einführung in die empiristische Wissenschaftsauffassung*. Den Haag: Van Stockum & Zoon, 1951. Reprint: 1990 Frankfurt/M: Suhrkamp, hrsg. von Friedrich Stadler.

[Molden, 1981] O. Molden. *Der andere Zauberberg. Das Phänomen Alpbach.* Wien-München-Zürich-New York: F. Molden, 1981.

[Monk, 1990] R. Monk. *Ludwig Wittgenstein. The Duty of Genius*, London: Jonathan Cape, 1990.

[Morgenstern, 1936] O. Morgenstern. Logistik und Sozialwissenschaften, in: *Zeitschrift far Nationalökonomie* 7, 1-24, 1936.

[Morgenstern and von Neumann, 1944] O. Morgenstern and J. von Neumann. *Theory of Games and Economic Behavior.* Princeton: Princeton University Press, 1944.

[Morris, 1935] C. Morris. Some Aspects of American Scientific Philosophy, in: *Erkenntnis* 5, 142-150, 1935.

[Morris, 1936] C. Morris. Opening Speech (For the American Delegates), "Semiotic and Scientific Empiricism", in: *Actes du Congres International de Philosophie Scientifique* I. Paris: Hermann & Cie., 22 bzw. 42-56, 1936.

[Morris, 1938] C. Morris. *Foundations of the Theory of Signs.* VIm. 1/2, 1938. Auch in Neurath/Carnap/Morris 1970, 77-138.

[Nagel, 1936] E. Nagel. Impressions and Appraisals of Analytic Philosophy in Europe, in: *Journal of Philosophy* 33, 1936. Auch in Nagel 1956, 191-246.

[Nagel, 1936] E. Nagel. Impressions and Appraisals of Analytic Philosophy, Journal of Philosophy XXXIII. Reprint: Nagel, E. 1956. *Logic without Metaphysics and other Essays in the Philosophy of Science*, Glencoe, Ill.: The Free Press, 191-246, 1936.

[Nagel, 1956] E. Nagel. *Logic without Metaphysics and other Essays in the Philosophy of Science.* Glencoe: The Free Press, 1956.

[Neurath, 1935] O. Neurath. Pseudorationalismus der Falsifikation *Erkenntnis* V, 353-365, 1935. Reprint: Neurath 1983, 121-131.

[Neurath, 1936] O. Neurath. Erster Internationaler Kongress für Einheit der Wissenschaft in Paris 1935, *Erkenntnis* 5, 377-428, 1936.

[Neurath, 1941] O. Neurath. Universal Jargon and Terminology, in: *Proceedings of the Aristotelian Society* 41, 127-148, 1941. In deutscher Übersetzung in Neurath 1981, 901-924.

[Neurath, 1942] O. Neurath. International Planning for Freedom, *The New Commonwealth Quaterly*, April/July 1942, 23-28, 1942. Reprint: Neurath 1973, 422-440. (Neurath was member of the Editorial Board of the NCQ and its successor, *The London Quaterly of World Affairs*).

[Neurath, 1945] O. Neurath. The Road to Serfdom, *The London Quaterly of World Affairs*, Jan. 1945, 121f.

[Neurath, 1973] O. Neurath. *Empiricism and Sociology*. Ed. by M. Neurath and R.S. Cohen, Dordrecht-Boston: Reidel, 1973.

[Neurath, 1981] O. Neurath. *Gesammelte philosophische und methodologische Schriften*. 2 Bande. Hrsg. von Rudolf Haller und Heiner Rutte. Wien: Hölder-Pichler-Tempsky, 1981.

[Neurath, 1983] O. Neurath. *Philosophical Papers 1913-1946*. Ed. by R.S. Cohen and M. Neurath. Dordrecht-Boston-Lancaster: Reidel, 1983.

[Neurath, 1987] O. Neurath. The New Encyclopedia, in: *Unified Science*, p. 136f. Ed. by Brian McGuiness. Dordrecht: Reidel, 1987.

[Neurath *et al.*, 1938] O. Neurath, E. Brunswik, C. L. Hull, G. Mannoury, and J. H. Woodger. *Zur Enzyklopadie der Einheitswissenschaft. Vorträge.* Den Haag: VanStockum & Zoon, 1938. Wiederabgedruckt in: Einheitswissenschaft 1992.

[Neurath *et al.*, 1970] O. Neurath, R. Carnap,and C. Morris, eds. *Foundations of the Unity of Science. Toward an International Encyclopedia of Unified Science*. 2 Vlms. Chicago-London: The University of Chicago Press, 1970.

[Österreich, 1965] *Österreich- - Geistige Provinz?* 1965. Wien-Hannover-Bern: Forum Verlag.

[Österreicher, 1995] *Österreicher im Exil. USA 1938-1945*. DÖW (Hrsg.) 1995. Eine Dokumentation. 2 Bande. Ein1eitung, Auswahl und Bearbeitung: Peter Eppe1. Wien: Osterreichischer Bundesverlag.

[Pap, 1955] A. Pap. *Analytische Erkenntnistheorie. Kritische Übersicht aber die neueste Entwick1ung in den USA und England*. Wien: Springer, 1955.

[Popper, 1934] K. R. Popper. *Logik der Forschung. Zur Erkenntnistheorie der modernen Naturwissenschaft*, Wien: Verlag Julius Springer, 1934. Published: 1935 (= Schriften zur wissenschaftlichen Weltauffassung, ed. by Ph. Frank and M. Schlick, Vol. 9). First English edition: *The Logic of Scientific Discovery*, London: Hutchinson & Co. New York: Basic Books 1959.Sixth (revised) impression 1972.

[Popper, 1974] K. R. Popper. Intellectual Autobiography, in *The Philosophy of Karl Popper*. Ed. by P.A. Schilpp, La Salle, Ill.: OpenCourt, 3-181, 1974.

[Pabisch, 1989] P. Pabisch, ed. *From Wilson to Waldheim*. Riverside: Ariadne Press, 1989.

[Platt and Hoch, 1996] J. Platt and P. K. Hoch. The Vienna Circle in the United States and Empirical Research Methods in Sociology, in: [Ash and Sollner, 1996, pp. 224-245].

[Quine, 1951] W. V. O. Quine. Two Dogmas of Empiricism, in: *Philosophical Review* 60, 20-43, 1951. Auch in Quine 1953.

[Quine, 1953] W. V. O. Quine. *From a Logical Point of View. 9 Logico-Philosophical Essays.* Cambridge, Mass.: Harvard University Press, 1953.

[Ramsey, 1923] F. P. Ramsey. Correspondence with Moritz Schlick. Schlick Papers, Vienna Circle Archives, Haarlem, NL, 1923.

[Raven and Krohn, 1996/2000] D. Raven and W. Krohn. Edgar Zilsel. His Life and Work, in: [Raven *et al.*, 2000, pp. xix-lxi].

[Raven *et al.*, 2000] D. Raven, W. Krohn, and R. S. Cohen, eds. Edgar Zilsel, *The Social Origins of Modern Sciences.* Dordrecht-Boston-London: Kluwer, 2000.

[Regis, 1989] E. Regis. *Einstein, Godel & Co. Genialitat und Exzentrik Die Princeton-Geschichte*. Basel-Boston-Berlin: Birkhauser, 1989

[Reichenbach, 1938] H. Reichenbach. *Experience and Prediction. An Analysis of the Foundations and the Structure of Knowledge*. The University of Chicago Press 1938. Deutsch: *Erfahrung und Prognose. Eine Analyse der Grundlagen und der Struktur der Erkenntnis.* Mit Erläuterungen von Alberto Coffa. Braunschweig-Wiesbaden: Vieweg 1983.

[Reichenbach, 1951] H. Reichenbach. *The Rise of Scientific Philosophy.* University of California Press 1951. Deutsch: *Der Aufstieg der wissenschaftlichen Philosophie.* Berlin:Grunewald: Herbig 1951

[Reisch, 1995] G. Reisch. *A History of the International Encyclopedia of Unified Science*. Ph.D. Thesis. Chicago 1995.

[Reisch, 2005] G. Reisch. *How the Cold War Transformed Philosopy of Science. To the Icy Slopes of Logic.* Cambridge University Press 2005.
[Richardson and Uebel, 2006] A. Richardson and T. Uebel, eds. *The Cambridge Companion of Logical Empiricism.* Cambridge University Press, 2006.
[Ringer, 1997] F. Ringer. *Max Weber's Methodology. The Unification of the Cultural and Social Sciences.* Harvard University Press, 1997.
[Runes, 1944] D. Runes, ed. *The Dictionary of Philosophy.* London: Routledge, 1944.
[Russell, 1905] B. Russell. On Denoting, in *Mind* 14, 479-473, 1905.
[Russell, 1914] B. Russell. *Our Knowledge of the External World as a Field for Scientific Method* in Philosophy, London: Open Court, 1914.
[Russell, 1936] B. Russell. The Congress of Scientific Philosophy, in *Actes du Congrés de Philosophie scientifique*, Sorbonne 1935, Paris: Hermann & Cie, 10-12, 1936.
[Russell, 1938] B. Russell. On the Importance of Logical Form, 1938. Reprint: Carnap, Morris, Neurath 1971, 39ff.
[Russell, 1940] B. Russell. *An Inquiry into Meaning and Truth*, London: Allen & Unwin, 1940.
[Rutkoff and Scott, 1986] P. M. Rutkoff and W. B. Scott. *New School. A History of The New School for Social Research.* New York-London: The Free Press-Collier Macmillan Pub, 1986.
[Schlick, 1930/31] M. Schlick. Ü ber wissenschaftliche Weltauffassung in den Vereinigten Saaten von Amerika, in: *Erkennntis* 1, 75f, 1930/31.
[Schlick, 1930] M. Schlick. Die Wende der Philosophie, *Erkenntnis* I, 4-11, 1930.
[Schlick, 1931] M. Schlick. The Future of Philosophy, *Proceedings of the Seventh International Congress of Philosophy*, held at Oxford, 1930, London, 112-116, 1931.
[Schurz and Dorn, 1993] G. Schurz and G. J. W. Dorn. Report: After Twenty Years. Die Entwicklung der Wissenschaftstheorie in Osterreich 1971-1990, in: *Journal for General Philosophy of Science* 24/2, 315-347, 1993.
[Skorupski, 1993] J. Skorupski. *English-Language Philosophy 1750 to 1945.* Oxford-New York: Oxford University Press, 1993.
[Sluga, 1999] H. Sluga. What Has History to Do with Me? Wittgenstein and Analytic Philosophy, *Inquiry,* 41, 99-121, 1999.
[Smith, 1986] L. D. Smith. *Behaviorism and Logical Positivism. A Reassessment of the Alliance.* Stanford: Stanford University Press, 1986.
[Somerville, 1936] J. Somerville. Logical Empiricism and the Problem of Causality in Social, in: *Erkenntnis* 6, 405-411, 1936.
[Sorell, 1991] T. Sorell. *Scientism. Philosophy and the Infatuation with Science*, London-New York: Routledge, 1991.
[Snow, 1959] C. P. Snow. *The Two Cultures: and a Second Look.* An Expanded Version of the two Cultures and the Scientific Revolution. Cambridge University Press 1959/1964.
[Sokal and Bricmont, 1998] A. Sokal and J. Bricmont. *Fashionable Nonsense. Postmodern Intellectuals' Abuse of Science.* New York: Picador, 1998. French edition: *Impostures Intellectuelles.* Paris: Editions Odile Jacob. German edition: *Eleganter Unsinn. Wie Denker der Postmoderne die Wissenschaften mißbrauchen.* München: Beck 1999.
[Spohn, 1991] W. Spohn, ed. *Erkenntnis Orientated: A Centennial Volume for Rudolf Carnap and Hans Reichenbach.* Dordrecht-Boston-London: K1uwer, 1991.
[Srubar, 1988] I. Srubar, (Hrsg.). *Exi1, Wissenschaft, Identität. Die Emigration deutscher Sozia1wissenschaft1er 1933-1945.* Frankfurt/M.: Suhrkamp, 1988.
[Stadler, 1987/88] F. Stadler (Hrsg.). *Vertriebene Vernunft. Emigration und Exil österreichischer Wissenschaft*, 1987/88. 2 Bande. Wien-München: Jugend und Vo1k. 2. Auflage Münster: LIT Verlag 2004.
[Stadler, 1988] F. Stadler (Hrsg.). *Kontinuität und Bruch 1938-1945/1955. Beitrage zur österreichischen Ku1tur- und Wissenschaftsgeschichte*, 1988. Wien-München: Jugend und Vo1k. 2. Auflage Münster: LIT Verlag 2004.
[Stadler, 1990] F. Stadler, (Hrsg.). Richard von Mises. *K1eines Lehrbuch des Positivismus. Einführung in die empiristische Wissenschaftsauffassung.* Frankfurt/M.: Suhrkamp, 1990.
[Stadler, 1993] F. Stadler, ed. *Scientific Philosophy: Origins and Developments.* Dordrecht-Boston-London: Kluwer, 1993.
[Stadler, 1997] F. Stadler. *Studien zum Wiener Kreis. Ursprung, Entwick1ung und Wirkung des Logischen Empirismus im Kontext.* Frankfurt/M.: Suhrkamp, 1997.
[Stadler, 1997a] F. Stadler (Hrsg.). *Phänomeno1ogie und Logischer Empirismus. Zentenarium Felix Kaufmann.* Wien-New York: Springer, 1997.

[Stadler, 1997b] F. Stadler, (Hrsg.). *Wissenschaft a1s Ku1tur. Osterreichs Beitrag zur Moderne.* Wien-New York: Springer, 1997.

[Stadler, 1997c] F. Stadler. Die andere Kulturgeschichte. Am Beispiel von Emigration und Exi1 der österreichischen Inte1lektue1len 19301945", in: Steininger/Gehler (Hrsg.) II, 499-558, 1997.

[Stadler, 2001] F. Stadler. *Studien zum Wiener Kreis. Ursprung, Entwicklung und Wirkung des Logischen Empirismus im Kontext*, Frankfurt/M.: Suhrkamp. (Sonderausgabe), 2001.

[Stadler, 2001b] F. Stadler. *The Vienna Circle. Studies in the Origins, Development, and Influence of Logical Empiricism,* Wien-New York: Springer, 2001.

[Stadler, 2001c] F. Stadler. *The Vienna Circle. Studies in the Origins, Development, and Influence of Logical Empiricism.* Wien-New York: Springer, 2001.

[Stadler, 2004] F. Stadler. Transfer and Transformation of Logical Empiricism: Quantitative and Qualitative Aspects, in [Hardcastle and Richardson, 2004].

[Stadler and Weibel, 1995] F. Stadler and P. Weibel, eds. *The Cultural Exodus from Austria*, Wien-New York: Springer, 1995.

[Stebbing, 1933] S. Stebbing. *Logical Positivism and Analysis, Annual Philosophical Lecture, British Academy, London:* Oxford University Press, 1933.

[Stebbing, 1935] S. Stebbing. Notes on an Informal Conference on Logical Positivism, held at Belsize Park, London, 5-6th January, 1935 in Vienna Circle Archives, Haarlem, NL, Neurath papers, 1935.

[Stebbing, 1939-40] S. Stebbing. Language and Misleading Questions, *Erkennntis/The Journal of Unified Science*, VIII, 1-6, 1939-40.

[Stebbing, 1944] S. Stebbing. *Ideals and Illusions*, London: Watts and Co, 1944.

[Stebbing, 1944a] S. Stebbing. *Men and Moral Principles.* L.T. Hobhouse memorial Trust Lectures, No. 13, London: Oxford University Press, 1944.

[Stegmüller, 1979] W. Stegmüller. *Rationale Rekonstruktion von Wissenschaft und ihrem Wandel. Mit einer autobiographischen Einleitung.* Stuttgart: Reclam, 1979.

[Steininger and Gehler, 1997] R. Steininger and M. Gehler, eds. *Osterreich im 20. Jahrhundert. Ein Studienbuch in zwei Banden.* Wien-Köln-Weimar: Böhlau, 1997.

[Strauss, 1991] H. A. Strauss. Wissenschaftsemigration als Forschungsproblem, in: Strauss u.a. (Hrsg.), 9-23, 1991.

[Strauss *et al.*, 1991] H. A. Strauss, K. Fischer, Ch. Hoffmann, and A. Söllner, eds. *Die Emigration der Wissenschaften nach 1933. Diszip1ingeschicht1iche Studien.* München-London-New York-Paris: K.G. Saur, 1991.

[Suppe, 1977] F. Suppe, ed. *TheSstructure of Scientific Theories.* Urbana-Chicago: University of Illinois Press, 1977.

[Thiel, 1984] C. Thiel. Folgen der Emigration deutscher und österreichischer Wissenschaftstheoretiker und Logiker zwischen 1933 und 1945, in: *Berichte zur Wissenschaftsgeschichte* 7, 227–256, 1984.

[Topitsch, 1960] E. Topitsch (Hrsg.). *Prob1eme der Wissenschaftstheorie. Festschrift für Viktor Kraft.* Wien: Springer, 1960.

[Uebel, 1991] T. E. Uebel, ed. *Rediscovering the Forgotten Vienna Circle. Austrian Studies on Otto Neurath and the Vienna Circle.* Dordrecht-Boston-London: Kluwer, 1991.

[Uebel, 1992] T. E. Uebel. *Overcoming Logical Positivism from Within. The Emergence of Neurath's Naturalism in the Vienna Circle's Protocol Sentence Debate.* Amsterdam-Atlanta, 1992.

[Uebel, 2000] T. E. Uebel. *Vernunftkritik und Wissenschaft. Otto Neurath und der Erste Wiener Kreis.* Wien-New York: Springer, 2000.

[Veröffentlichungen, 1991] *Veröffentlichungen* 1991ff. Veröffentlichungen des Instituts Wiener Kreis. Hrsg. von Friedrich Stadler. Wien-New York: Springer, 1991.

[Vienna Circle, 1993] *Vienna Circle Institute Yearbooks.* Ed. by Friedrich Stadler. Dordrecht-Boston-London: Kluwer, 1933.

[Waismann, 1965] F. Waismann. *The Principles of Linguistic Philosophy*, ed. by R. Harré, London-Melbourne-Toronto: Macmillan, 1965.

[Waismann, 1976] F. Waismann. *Logik, Sprache, Phi1osophie.* Hrsg. von G.P. Baker/B. McGuinness/J. Schulte. Stuttgart: Reclam, 1976.

[Warnock, 1971] G. J. Warnock. *Englische Philosophie im 20. Jahrhundert.* Stuttgart: Reclam, 1971.

[Wartofsky, 1963] M. W. Wartofsky, ed. *Boston Studies in the Philosophy of Science. Proceedings of the Boston Colloquium for the Philosophy of Science 1961/1962*. Dordrecht: Reidel, 1963.

[Wissenschaftliche, 1919] *Wissenschaftliche Weltauffassung: Der Wiener Kreis* 1929. Translation: The Scientific Conception of the World: The Vienna Circle, Dordrecht-Boston-London: Reidel 1973. Reprinted in Neurath, O. 1973, 299-318.

[Zecha, 1970] G. Zecha. Die gegenwärtige Situation der Wissenschaftstheorie in Osterreich, in: *Zeitschrift far allgemeine Wissenschaftstheorie* 1/2, 284-321, 1970.

[Zilian, 1990] H. G. Zilian. *K1arheit und Methode. Felix Kaufmanns Wissenschaftstheorie*. Amsterdam-Atlanta 1990.

[Zilian, 1997] H. G. Zilian. Felix Kaufmann - Leben und Werk, in: [Stadler, 1997a, pp. 9-22].

[Zilsel, 1932/33] E. Zilsel. Bemerkungen zur Wissenschaftslogik, in: *Erkenntnis* 3, 143-161, 1932/33.

[Zilsel, 1939] E. Zilsel. Preprint: The Social Roots of Science, Harvard, 1939.

[Zilsel, 1942] E. Zilsel. Problems of Empiricism, in: [Neurath *et al.*, 1970, pp. 803-844].

[Zilsel, 1972] E. Zilsel. *Die Entstehung des Geniebegriffs. Ein Beitrag zur Ideengeschichte des Frühkapitalismus.* Tübingen: Mohr 1926. Reprint: Hildesheim-New York: Georg Olms, 1972.

[Zilsel, 1976] E. Zilsel. *Die sozialen Ursprünge der neuzeitlichen Wissenschaft*. Hrsg von Wolfgang Krohn. Frankfurt/M.: Suhrkamp, 1976.

[Zilsel, 1990] E. Zilsel. *Die Geniereligion. Ein kritischer Versuch über das moderne Persönlichkeitsideal, mit einer historischen Begründung*. Hrsg und eingeleitet von Johann Dvorak. Frankfurt/M.: Suhrkamp., 1990

INDEX