What Are the Aims of Science?

Aaron Sloman **Published 1976 in** *Radical Philosophy*
Written while author was at University of Sussex
Now at School of Computer Science University of Birmingham
http://www.cs.bham.ac.uk/~axs

I Introduction

If we are to understand the nature of science, we must see it as an activity
and achievement of the human mind alongside others, such as the achievements
of children in learning to talk and to cope with people and other objects in
their environment, and the achievements of non-scientists living in a rich and
complex world which constantly poses problems to be solved. Looking at
scientific knowledge as one form of human knowledge, scientific understanding
as one form of human understanding, scientific investigation as one form of
human problem-solving activity, we can begin to see more clearly what science
is, and also what kind of mechanism the human mind is.

I suggest that no *simple* slogan or definition, such as can be found in
textbooks of science or philosophy can capture its aims. Science is a complex
network of different interlocking activities with multiple practical and
theoretical aims and a great variety of methods. I shall try to describe some
of the aims and their relationships in this essay.

Oversimple characterisations, by both scientists and philosophers, have led to
unnecessary and crippling restrictions on the activities of some would-be
scientists, especially in the social and behavioural sciences, and to
harmfully rigid barriers between science and philosophy.

By undermining the slogan that science is the search for laws, and subsidiary slogans such as that quantification is essential, that scientific theories must be empirically refutable, and that the methods of philosophers cannot serve the aims of scientists, I shall try, in what follows, to liberate some scientists from the dogmas indoctrinated in universities and colleges. I shall also try to show philosophers how they can contribute to the scientific study of man, thereby escaping from the barrenness and triviality complained of so often by non-philosophers and philosophy students.

A side-effect which will be reported elsewhere, is to undermine some old philosophical distinctions and pour cold water on battles which rage around them -- like the distinction between subjectivity and objectivity, and the battles between empiricists and rationalists.

First crude subdivision of aims of science

Science has not just one aim but several. The aims of scientific investigation can be crudely subdivided as follows:

1 to extend man's knowledge and understanding of
the form and contents of the universe (*factual* aims),

2 to extend man's control over the universe, and to use this to improve the world (*technological* or practical aims),

3 to discover how things ought to be, what sorts of things are good or bad and how best to further the purposes of nature or God (*normative* aims).

Whether the third aim make s sense (and many scientists and philosophers would dispute this) depends on whether it is possible to derive values and norms from facts. I shall not discuss it as it is not relevant to the main purposes of this enquiry.

The second kind of aim will not be given much attention either, except when relevant to discussions of the first kind of aim.

The first kind of aim, like the others, is of course much wider than science. We all, including infants and children, aim to extend our knowledge and understanding: science is unique only in the rigour, system, and amount of co-operation between individuals involved in its methods. For the present, however, I shall not explore the peculiarities of science, since what it has in common with other forms of acquisition of knowledge has been too long neglected, and it is the common features I want to describe.

In particular, notice that one cannot have the aim of *extending* one's knowledge unless one presupposes that one's knowledge is incomplete, or perhaps even includes mistakes. This means that pursuing this aim requires systematic self-criticism in order to find the gaps and errors. This distinguishes both science and perhaps the curiosity of young children from some other belief systems, such as dogmatic theological systems and political ideologies. But it does not distinguish science from philosophy.

A further subdivision: form and content

The factual aim, extending knowledge and understanding, can be further
subdivided as follows:

1.a Extending knowledge of what sorts of things are possible and impossible in
the world, and how or why they are (the aim of interpreting the world, or
learning about its *form*)

1.b Extending knowledge of what particular objects, events, processes, or
states of affairs exist or existed in particular places at particular times
(the aim of acquiring 'historical' knowledge, or learning about the
contents of the world).

A similar distinction pervades the writings of Karl Popper, though he would
disagree with some of the things I say below about (l.a). Different branches
of science tend to stress one or other of these aims, though both aims are
usually present to some extent. For instance, physics is more concerned with
aim (1.a), whereas astronomy is perhaps more concerned with (1.b). Geology,
geography, biology, anthropology, human history, sociology, and some kinds of
linguistics tend to be more concerned with (1.b), i.e. with learning about the
particular contents of particular parts of the universe. Chemistry, some
branches of biology, economics and psychology attempt to investigate truths
not so restricted in scope. In the jargon of philosophers, (1.a) is concerned
with universals, (1.b.) with particulars.

However, the two scientific aims are very closely linked. One cannot discover
what sorts of things are possible, nor test explanatory theories, except by
discovering particular facts about what actually exists or occurs. Conversely,
one cannot really understand *particular* objects, events, processes etc,
except insofar as one classifies and explains them in the light of more
general knowledge about what *kinds* of things there can be and how or
why.

These two aims are closely linked in all forms of learning about the world,
not only in science.

Notice that I have characterised these aims in a dynamic form: the aim is to
extend knowledge, to go on learning. Some might say that the aim is to arrive
at some terminal state when everything is known about the form and content of
the world, or at least the form. There are serious problems about whether this
suggestion makes sense: for example, how could one tell that this goal had
been reached? But I do not wish to pursue the matter. For the present, it is
sufficient to note that it makes sense to talk of extending knowledge, that is
removing errors and filling gaps, whether or not any final state of complete
knowledge is possible. Some of the criteria for deciding what is an extension
or improvement will be mentioned later.

Many philosophers of science have found it hard to explain the sense in which
science makes progress, or is cumulative. (e.g. Kuhn, 1962, last chapter).
This is because they tend to think of science as being mainly concerned with
laws; and supposed laws are constantly being refuted or replaced by others.

Very little seems to survive.

But if we see science as being also concerned with knowledge of what is possible, then it is obviously cumulative. For a single instance demonstrates a new possibility and, unlike a law, this cannot be refuted by new discoveries, even if the possibility is re-described from time to time as the language of scientists evolves. Hypotheses about the limits of these possibilities (laws) lack this security, for they are constantly subject to revision as the boundaries are pushed further out, by newly discovered (or created) possibilities. Explanations of possibilities and their limits constantly have to be refined or replaced, for the same reason. But this is all a necessary part of the process of learning and understanding more about what is possible in the world. It is an organic, principled growth, even if people do sometimes disagree about what is and what is not progress, for reasons to be described later.

Let us now look more closely at aim (l.a), the aim of extending knowledge of the *form* of the world.

The interpretative aims of science

The aims listed below together constitute the aim (l.a) of interpreting the world, or learning about its form. They are all so closely related that to treat them as separate aims would be artificial. Similarly, to call some of them 'scientific' and others 'metaphysical' or 'philosophical', as empiricists and Popperians tend to do, is to ignore their interdependence. Rather, they are all aspects of one aim. For convenience I shall talk of them as separate aims, but this will be qualified by describing their connections. They are:

(a) Development of *new concepts and symbolisms* making it possible to conceive of, represent, think about and ask questions about new kinds or ranges of possibilities (e.g. new kinds of physical substances, events, processes, animals, mental states, human behaviour, languages, social systems, etc.). This aim includes the construction of taxonomies, typologies, scales of measurement and notations for structural descriptions. This extension of our conceptual and symbolic powers is one of the major functions of mathematics in science.

(b) Extending knowledge of what kinds of things (including events and processes) are *possible in the world*; i.e. what kinds of things are not merely conceivable or representable but really can exist or occur. Finding out what actually exists, and trying to make new things exist, are often means to this end. We can distinguish knowledge of absolute possibility concerning a phenomenon X (X can exist) from knowledge of relative possibility (X can exist in conditions C). Extending knowledge of relative possibilities for X is an important way of extending knowledge of what is possible. All this should be distinguished from the aim (e) below, of finding out what kinds of things are most likely, common or frequent, either absolutely or in specified conditions. The latter is a concern with *probabilities* not *possibilities*.

Aim (b) clearly presupposes aim (a), for one can only acknowledge possibilities that one can conceive of, describe or represent.

(c) Constructing theories to *explain known possibilities:* i.e. theories about the underlying structures, mechanisms, and processes capable of generating such possibilities. For instance, a theory of the constituents of atoms may explain the possibility of chemical elements with different properties. 'How is this possible?' is the typical form of a request for this kind of explanatory theory, and should be contrasted with the question 'Why is this so?' or 'Why is this impossible?', discussed in (f) below.

(d) Finding *limitations on combinations* of known possibilities. These are often called laws of nature: for instance to say that it is a law of nature that all Xs are Ys is to say that it is *impossible* for something to be both an X and not a Y. It is these laws, limitations or impossibilities which make the world relatively stable and predictable. This aim, like (c), presupposes aim (b), since one can discover limitations of possibilities only if one already knows about those possibilities. (This aim of science is the one most commonly stressed in the writings of scientists and philosophers. It subsumes the aim of discovering causal connections, since X causes Y if the occurrence of X makes the non-occurrence of Y impossible.)

(e) Finding *regular or statistical correlations* between different possibilities, for instance correlations of the form 'In conditions C, 90% of all Xs are Ys'. This is a search for probabilities. It presupposes aim (b) for the same reason as (d) does. Except in quantum physics, the search for such statistical correlations is really only a stopgap or means towards acquiring a deeper understanding of the sort described in (d), above.

Alternatively, it may be an aim of a historical science: facts about relative frequencies and proportions of various kinds of objects, events or processes are often important facts about the contents of a particular part of the world. For instance, most of the correlations unearthed by social scientists are culture-relative. Such information may have practical value despite its theoretical poverty.

(f) Constructing *theories to explain known impossibilities, laws and correlations*. Such theories answer 'Why?' questions, and are generally refinements of the theories described in (c). That is, explaining limits of possibilities (i.e. explaining laws) presupposes or refines an explanation of the possibilities limited. The theory of molecules composed of atoms which can recombine explains the *possibility* of chemical change. Further refinements concerning weights and valencies of atoms explain the observed *limitations*: the laws of constant and multiple proportions.

[Picture of Newcomen's beam engine added in original.]

(g) Finding and eliminating inadequate concepts, symbolisms, and laws, about what is and is not possible, and inadequate explanations of possibilities and laws. That this is an aim of science is, as already remarked, implied by saying that an aim of science is to extend knowledge. As many philosophers of science have pointed out, it is not generally possible to *prove*

explanatory theories in science: at most they can only be refuted or shown to be inadequate in some way. Moreover, when several candidates survive refutation, the most that can be done is to compare their relative merits and faults, without necessarily establishing the absolute superiority of one over the other.

It is often assumed that the only kinds of proper tests are empirical (i.e. observations of new facts, in experiments or in nature). However, we shall see that many important tests are not empirical. We shall also see that just as negative instances count against laws, so do positive instances provide support for theories about possibilities.

If forced to summarise all this in a single slogan, one could say:

A major aim of science is to find out what sorts of things are and are not possible in the world, and to explain how and why.

Though too short to be clear, this may be a useful antidote to more common slogans stressing the discovery and explanation of laws and regularities. Such slogans lead to an excessive concern with prediction, control and testing, topics mainly relevant to aims (d) to (g), while insufficient attention is paid to the more fundamental aims (a) to (c), especially in psychology and social science. The result is often sloppy research, theorising and teaching.

More about interpretative and historical aims of science

Unlike the historical scientist, the interpretative scientist is interested in actual objects, events or situations only insofar as they are *specimens of what is possible*. The research chemist is not interested in the fact that this particular sample of water was, on a certain day, decomposed into hydrogen and oxygen in *that* laboratory, except insofar as this illustrates something universal, such as the *possibility* of decomposing water. This possibility refutes the theory that water is a chemical element and corroborates the alternative hypothesis that all water is composed of hydrogen and oxygen, and also more general theories about possible kinds of transformations of matter. Similarly, although an 'historical' biologist may be interested in recording, for a fascinated public, the flora and fauna of a foreign isle, or the antics of a particularly intelligent chimpanzee, the 'interpretative' biologist is interested only insofar as they illustrate something, such as what *kinds* of plants and animals can exist (or can exist in certain conditions) , or what *kinds* of behaviour are possible for a chimpanzee or for some other class containing the animal in question.

In short, the interpretative scientist studies the *form* of the world, using the contents only as evidence, whereas the historical scientist simply studies the *contents*.

There is, of course, no reason. why any one science, or scientist, should be classified entirely as interpretative, or entirely as historical. Different elements may intermingle in one branch of science. For instance, a linguist studying a particular dialect is an interpretative scientist insofar as he is not concerned merely to record the actual set of sentences uttered by certain speakers of that dialect, but to characterise the full range of sentences that would or could be intelligible to an ordinary speaker of that dialect, namely, a range of possibilities. However, insofar as he is interested merely in finding out exactly what dialect is intelligible to a certain spatio-temporally restricted group of persons, he is an historical linguist, as contrasted with a linguist who is interested in this dialect primarily as a sample of the kinds of language which human societies can develop: the attempt to characterise this set of possible languages is often called the search for linguistic universals.

Thus a richer philosophical terminology would be required for a precise description of hybrid historical and interpretative aims. This is not relevant to our present concerns and will not be pursued further. Instead, in II-IV below, I will concentrate mainly upon an analysis of the first three components of the interpretative aim, outlined above. These three are very tightly interconnected.

It is very hard to describe the distinctions between them accurately, and I am sure I do not yet understand these matters aright. Moreover, each of them could be further subdivided. Detailed historical analysis is required here, so that similarities and differences between cases can be described accurately and a more satisfactory typology developed; a contribution to the scientific study of science. Alas, this will require the help of persons more scholarly than I.

II
The Role of Concepts and Symbolisms

Individuals (and cultural groups) can differ not only in the things they know or believe, but also in the possibilities they can grasp, the concepts they have, the generative power of the languages they use, the questions they can ask.

As new concepts and symbolisms are developed, and the language extended, new questions become askable. For instance, people who grasp the concepts 'hotter' and 'longer' can understand the question whether metal rods get longer when they are made hotter. And they may even be able to grasp crude distinctions between metals according to which grows longer faster when heated. But in order to learn to think about whether the change in length is *proportional* to the change in temperature, so that they can then use the constant of proportionality (divided by the length of the rod) to define a numerical 'coefficient of expansion' for each metal, they need to grasp numerical representations of differences in temperature and length ('hotter by how much?', 'longer by how much?').

Similarly, although people may have a crude grasp of distinctions between velocity and acceleration, and may be able to detect gross changes in either, on the basis of their own experiences of moving things, being moved, and perceiving moving objects, nevertheless, until they have learnt how to relate concepts of distance and time to numerical interval scales, they cannot easily make precise distinctions between different velocities, or between acceleration and rate of change of acceleration, nor think of precise relations between these concepts.

These familiar examples show the power of extending scientific language by introducing numerical concepts and notations corresponding to old non-numerical concepts. This sort of thing has been so important in physics that many biological and social scientists have been deluded into thinking it part of the definition of a scientist that he uses numbers!

The replacement of Roman numerals with the Arabic system is an example of a powerful notational advance. Another was the Cartesian method of using arithmetic to represent geometry and vice versa.

However, *non-numerical* conceptual and notational devices have also been important, such as concepts used in describing structures of plants and animals, concepts used for describing structures of mechanical systems and electrical circuits (geometrical and topological concepts), taxonomies or typologies, and grammatical concepts (see N. Chomsky, 1957).

All sorts of notations besides numerical and algebraic ones have played an important role in extending the abilities of scientists to express what they know and want to find out, for instance pictures, diagrams, maps, models, graphs, flow charts, and non-numerical computer programs. Examples include the diagrams used in the study of levers, pulleys, bending beams, and other mechanical systems, the 'pictures' of molecules used by chemists, for instance in the following representation of the formation of water from hydrogen and oxygen

$$(\text{H-H, H-H, O-O}) \Rightarrow (\text{H-O-H, H-O-H})$$
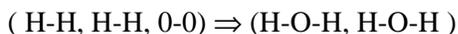
circuit diagrams used in electronics, optical drawings showing the paths of light rays, plates showing tracks of subatomic particles, and the 'trees' used to represent deep-structures of sentences.

Concepts may also be used without being represented explicitly by any external symbol. There are philosophers who dispute that these are cases of the use of concepts, but in the face of well known facts I can only regard this as verbal quibbling. We know that young children and other animals can discriminate, recognise and react intelligently to things which they cannot name or describe. The consistency and appropriateness of their behaviour shows that they act on the basis of reasons, even if they cannot articulate them or are unaware of them.

The same is true of an adult who cannot describe the features of musical compositions which enable him to recognise styles of composers and appreciate their music, or the cues which enable him to judge another's mood. No doubt this is true also of many scientists, especially when they are in the early phases of some kind of conceptual development. They may then, like children and chimpanzees, be unable to articulate fully the reasons they have for some of the decisions they take about interpreting evidence and assessing hypotheses.

Our minds may use symbolisms which we can only translate into actions, not spoken or written language -- until we have extended our language. Non logicians can often distinguish valid from invalid arguments without being able to say how. They have not learnt the overt language of logicians.

Even after going a stage further and learning how to articulate their reasons, scientists may not yet have learned how to teach their new concepts to colleagues and rival theorists. So attempts at rational persuasion break down. This has misled some philosophers and historians of science into thinking that there are no reasons, so that the decisions of scientists are irrational or non-rational. This is as silly as assuming that a mathematician is irrational simply because he cannot explain a theorem to a four year old child. The child may have much to learn before he can understand the problem, let alone the reasoning, and the mathematician may be a poor teacher.

Concepts are not simple things which you either grasp or don't grasp, or which can be completely conveyed by an explicit definition or axiomatic characterisation. For instance, as work of Piaget has shown so clearly, and Wittgenstein less clearly, very many of our familiar concepts, like 'number', 'more', 'cause', 'moral', and 'language', are very complex structures of which different fragments may be grasped at different times. A child aged four or five may be able to count flawlessly well beyond 20, and use counting to get correct answers to simple addition problems, yet be quite unable to count backwards, or to answer questions like 'What comes before 6?', 'Is 9 after (bigger than) 5?'. (Yet he succeeds with the numbers on a clock in front of him, so he understands the questions.) Does he or does he not understand the concept of number, or of 6? His failure does not prove that he is irrational: he still has some procedures to invent for himself -- if his parents and teachers don't destroy his creative abilities.

The more of one's concepts and associated procedures one is able to represent explicitly in symbols of some sort, the greater one's power to explore possibilities systematically by manipulating those symbols. For instance, by explicitly characterising aspects of our intuitive grasp of spatial structures in the form of axioms and definitions, one becomes able to experiment with alterations in the axioms and definitions, and thereby invent concepts of non-euclidean or other new sorts of geometries. In this way one can learn to think about new sorts of possibilities without waiting to be confronted with them. (This kind of thing may also happen below the level of consciousness, in children and scientists, as part of the process of learning and discovery.) Of course, one may also extrapolate too far, and construct representations of things which are not really possible in the world, so empirical investigation

of some sort is required to discover whether things which are conceivable or representable can also exist. For instance, merely analysing the concept of an element with atomic number 325 will not decide whether such a thing can occur. This is the reason for distinguishing the first aim of interpretative science, namely extending concepts and symbolisms, from the second aim, namely extending knowledge of what is really possible. (I believe many of these ideas are to be found in Kant (1781)).

Two phases in knowledge-acquisition:
understanding and knowing
(Grasping concepts and knowing which propositions are true.)

It is not always noted in epistemological discussions that there are two important phases or steps in the acquisition of knowledge. Discovering that P is true first of all requires the ability to understand the possibility that P might be true and might be false, which requires grasping the concepts used in the proposition P. The second phase is finding out that p is true, for instance by empirical observation, use of testimony, inference from what is already known, or some combination of these. In the first phase one is able to ask a question, in the second one has an answer. (There may be primitive kinds of knowledge-acquisition in which questions are never understood, only information acquired and used. But science is not like this).

Usually philosophers plunge into discussions of such questions as whether we can know anything about the future, or rationally believe anything about the future, without first asking how a rational being can even *think* about the future or *think* about alternative possible future states of affairs. They are therefore attempting to assess the rationality of certain decisions on the basis of a drastically incomplete account of the resources that might enter into the decision-making process. The reason why this has been shirked is partly because it is so hard to do, partly because of an unwarranted restriction of rationality to relations between evidence and belief-contents, and partly because many philosophers (unlike Wittgenstein, 1956), think that the investigation of conceptual mechanisms is a task for psychologists not philosophers.

However, most psychologists never even think of the important questions, and those who do usually lack the techniques of conceptual analysis required for teaching them: so the job does not get done.

(Piaget seems to be an exception. But many of his followers seem capable only of repeating his experiments, and not of extending his conceptual analyses.)

There is a need for a tremendous amount of research into what it is to understand various sorts of concepts, and what makes it possible. There is also a need for some kind of taxonomy of types of conceptual change, whether in individuals or in cultures.

The efforts made so far by psychologists to produce such taxonomies capture only a tiny fragment of the range of possible developments. Here are some examples of possibilities of conceptual change which still require adequate explanations. Going from grasping a relation like 'hotter' or 'longer' to grasping that it can be used to define equivalence classes of objects of the same temperature or length. Going from this to grasping the possibility of comparing *differences* in temperature or length (i.e. understanding an interval scale). Going from grasping some general concept defined in terms of a structure, or a function, or some combination of structure and function, to grasping systematic principles for subdividing that concept into different categories. Learning to separate the structural and functional aspects of a hybrid concept, like 'knife', or 'experiment'. Changing a concept by changing the theories in which it is embedded, in the way that the concept of mass was changed by going from Newtonian mechanics to Einstein's mechanics.

Developing a new more powerful symbolism for an old set of concepts: e.g. inventing differential calculus notation for representing changes, or using the concept of a mathematical function to generalise earlier concepts of regularity or correlation. Coming to see something in common between things one has never previously classified together, like mass and energy, particles and waves, straight lines and geodesics on a sphere.

Going from knowing a set of formulae and how to manipulate them to being able to see their relevance to a variety of new concrete problems, e.g. going from understanding algebra to being able to apply it in real life. (There are many other cases not so closely linked to science, e.g. the growth of social, moral and political consciousness Learning to feel shame, embarrassment, or guilt, as contrasted with annoyance or regret, requires complex cognitive development, and the same is true of many other human emotions. Some concepts e.g. 'impertinent' are only intelligible in certain cultures.)

Until these conceptual changes are better understood, discussion of 'incommensurability' of scientific theories and of the role of rationality in science is premature. Meanwhile education will continue to be largely a hit and miss affair, with teachers not knowing what they are doing or how it works.

To sum up so far. A system of concepts and symbols constitutes a language. (This statement is grossly inaccurate, but will do for present purposes.) A language which is used to formulate one theory will usually also contain resources for formulating alternatives, including the negation of the theory and versions of the theory in which some predicate, relational expression or numerical constant is replaced by another. So concepts and symbols are tools for generating possibilities or questions for investigation. They have greater generative power than theories. The scientist who usefully extends the language of science, unlike one who simply proposes a new theory using existing concepts and symbols, extends the hypothesis-forming powers of the scientists who understand him. In this sense conceptual advances are more profound.

The important differences between modern scientists and those of the distant past therefore concern not merely the statements and theories thought to be true or false, but also which statements and theories could be thought of at all. Not only are more answers known now, but more questions are intelligible.

Criticising conceptual systems

Sometimes old questions become unaskable as a result of conceptual change, like questions about phlogiston or absolute velocity, or perhaps 'medical' questions like 'What did he do to deserve this affliction?' Modern medical science contains no means of generating possibilities constituting answers to this question, though both laymen and some medical men (on Sundays?) may still formulate them. (Incompatible systems of concepts and theories may coexist in one mind -- but that's another story.) So science is served not only by extending and differentiating existing concepts: rejection of a concept or typology or mode of representation may also serve the aims of science by reducing the variety of dead-end questions and theories.

Concepts, typologies, taxonomies, and symbolisms can, like theories, be rationally criticised, and rejected or modified.

There are several ways in which a typology and associated notation can be criticised. For instance one may be able to show:

```
(a) that there are some possibilities it doesn't allow for,
(b) that it represents as possible some cases which
    are not really possible,
(c) that some of the subdivisions it makes are of no
    theoretical importance,
(d) that some category within it should be sub-divided into
    two or more categories, because their instances have
    different relations to the other categories,
(e) that the principle of subdivision is too vague to decide
    all known cases,
(f) that the classification procedure generates inconsistent
    classifications for some instances,
(g) that the notation used does not adequately reflect the
    structural properties of the typology, or of the instances,
(h) that the concepts used generate questions which apparently
    cannot be answered by scientific investigation (like the
    question 'How fast is the Earth moving through the ether?')
(i) that more powerful explanatory theories can be developed
    using other tools for representing possibilities.
    (It may be that some of these criteria are used,
    unconsciously of course, not only by scientists, but also
    by young children developing their conceptual systems.)
```

Several of these criteria will remain rather obscure until later. In particular, the first two can only be understood on the basis of a distinction between what is conceivable or representable and what is really possible in the world. We now examine this, in order to explain the difference between the first two interpretative aims of science.

III
Conceivable or Representable versus Really Possible

The second interpretative aim of science is to find out what kinds of things really are possible in the world and not merely conceivable. This includes such aims as finding out what sorts of physical substances, what kinds of transformations of energy, what kinds of chemical reactions, what kinds of astronomical objects and processes, what kinds of plants and animals, what kinds of animal behaviour, what kinds of mental development, what kinds of mental abnormality, what kinds of language and what kinds of social changes can exist or occur.

This aim is indefinitely extensible: having found out that Xs can exist or occur, one can then try to find out whether Xs can exist or occur in specified conditions C1, C2, C3,... Similarly, having found that objects can have one range of properties which can change (e.g. length) and can also have another range of properties which can change (e.g. temperature) one can then try to find out whether these properties can change independently of each other in the same object, such as a bar of metal, or a particular object in specified circumstances, such as a bar of metal under constant pressure or tension. Such further exploration of the *limits of combinations of known possibilities* merges into the search for laws and regularities, as explained previously.

We can conceive of, or describe, a lump of wood turning spontaneously into gold, or a human living unclothed in a vacuum, but it does not follow that these things really can exist. What is the difference?

First we look at what it is for something to be conceivable, representable, or describable.

As philosophers well know, the subjective feeling of intelligibility, the feeling of having understood or imagined something, is no guarantee that anything consistent was understood, imagined or conceived of. If someone claims to be able to conceive of the set of all sets which do not contain themselves, then provided he is using words in the normal way, we can show, by Russell's well known argument, using steps that he will accept if he is reasonable, that he was wrong, or that his 'conceiving' amounted to little more than repeating the phrase, or some equivalent, to himself. A sentence, phrase, picture, diagram, or other complex symbol will, if intelligible, be part of a language which includes syntactic and semantic rules in accordance with which the symbol is to be interpreted. The mere fact that the symbol is syntactically well-formed does not guarantee that it can be interpreted, though it may mislead us into thinking it can. More precisely, it may have a *sense* but necessarily fail to have any *denotation*.

Thus the question 'Does the table exist more slowly than the chair?' is syntactically perfect but we can show that so long as the words are used according to normal semantic rules there can be no answer to the question. For, 'more slowly' when qualifying a verb requires that verb to denote a process or sequence involving changes other than the change of time, so that the rate of change or succession can be measured against time. Existence is

not such a process, so rates of existence cannot be compared. (For more on this see Sloman, 1971b. For non-verbal examples see Clowes, 1971)

So we can use the notion of what is coherently describable or representable in some well defined language or representational system, as an objective semantic notion. What is conceivable to a person, will be what is coherently representable in some symbolic system which he uses, not necessarily fully consciously. It may be very hard, even for him, to articulate the system he uses, but that does not disprove its existence. These notions are as objective as the notion of logical consistency, which is a special case. However the mere fact that something is, in this sense, representable or conceivable does not mean that it really can exist. Conversely, what can exist need not be representable or conceivable using the symbolic resources available to scientists (or others) at any particular time: their language may need to be extended. Scientists (like children) may be confronted with an instance of some possibility, like inertial motion, diffraction, or curvature of space-time, without seeing it as such because they lack the concepts. (Kuhn, 1962, chapter X, has over-dramatised this by saying they inhabit a different world.)

The word 'possible' as I have used it, and as others use it, tends to slide between the two cases (a) used as a synonym for 'consistently representable or describable using some language', as in 'logically possible', and (b) used to refer to what can occur or exist in the world. But what is the difference between (a) and (b)?

This is not an easy question to answer completely. The main difference is that conceivability or representability can be established simply by analysing the sentence or other symbol used and checking that the syntactic and semantic rules of the language in question do not rule out a consistent interpretation (which is not always easy), whereas checking whether something really is or is not possible requires empirical investigation of some sort.

If an actual example is found, that conclusively establishes the possibility. The corresponding kind of impossibility is very much harder to establish, and perhaps can never be conclusively established, though one can often be fairly sure that something is not possible in the world either because of extensive and varied attempts to realise it, or on the basis of inference from some well established theory. (For instance, I am convinced by physical and biological arguments that it is impossible for a human being to live without clothing in a vacuum.)

However, possibility is not the same as actual existence. To say that it is possible for ten drugged alligators to be painted with red and yellow stripes and then piled on top of one another, is not to say that this ever has happened or will happen. Similarly, to say that several courses of action are possible for me, is not to say that I shall actually follow all of them. So, in saying that one of the aims of interpretative science is to find out which kinds of things are possible in the world, I do not mean that the aim is to find out which kinds actually exist, as in historical science. The latter is just a means, since existence establishes possibility. (See part one for more

on this.)

What other means are there of deciding that something is really possible, besides finding an instance? Alas, the only answer I can give to this is that we can reasonably, though always tentatively perhaps, infer that something is possible if we have an explanation of its possibility. What this amounts to is roughly the following: (a) we can consistently represent it using symbolic resources which have already been shown to be useful in representing what is actual, and (b) it is not ruled out by any well established law or theory specifying limitations on possibilities. Perhaps an extra condition is required: (c) it should differ from already realised possibilities only/in ways which are in some sense well understood. However, it is not clear that (c) adds anything to (a) and (b).

It is clear that these conditions do not conclusively prove something to be possible, for they rest on current theories of the limitations of what is possible and such theories, being empirical, are bound to include errors and omissions, at any stage in the advance of science. Further, these conditions do not yield clear decisions in all cases. For instance, is it reasonable to believe that it is possible for a normal human being to be trained (perhaps starting from birth) to run a mile in three minutes? It may not be clear whether we already know enough to settle such a question.

The above conditions for proving unrealised possibilities need to be further defined and illustrated. For the present, however, my aim is simply to indicate roughly how something can be shown to be possible without producing an instance. So I have demonstrated that possibility is a different concept from conceivability (or coherent representability), and also different from existence.

But I still have not given anything approximating to a complete analysis: this would require very much more than describing the criteria for deciding whether something is possible or not. It would also require analysis of the role of the concept of possibility in our thinking, problem-solving, deliberating, regretting, blaming, praising, etc, and its relations to a whole family of modal words, such as 'may', 'can', 'might', 'could', 'would', etc. A mammoth task. (For some useful beginnings see Gibbs, 1970).

At any rate, we cannot analyse 'Things of type X are possible' as *synonymous* with 'Either things of type X already exist, or else they are consistently representable in our symbolic system without being ruled out by known laws', since this would define real possibility in terms of the *current* system of concepts and beliefs. We could try a formula like 'Things of type X are possible if and only if they either exist or are consistently representable in some useful representational system and are not ruled out by any true laws'.

But this has the disadvantage of presupposing that there exists some complete set of true laws formulated in some unspecified language which correctly defines all the limitations on what is possible in the world. It is by no means clear that such a presupposition is intelligible. Moreover as a

definition it introduces a circularity, since it is notoriously hard to define the concept of a law without presupposing the concept of possibility or some related concept.

Despite the remaining obscurities, I hope I have done enough to indicate both that the first two aims of interpretative science are different, and also that they are very closely related. Now for a closer look at the third aim - the aim of explaining possibilities. I feel least satisfied with what I have to say about this.

IV
Explanations of Possibilities

A request for an explanation of a possibility or range of possibilities is characteristically expressed in the form 'How is X possible?' Unfortunately, the study of the role of such explanations in our thought is made more difficult by the fact that not everyone who requires, seeks or finds such an explanation, or who learns one from other people, asks this sort of question explicitly, or fully articulates the explanation when he has understood it. This partially explains why the role of possibilities and their explanations in science has not been widely acknowledged.
Roughly, an explanation of a possibility or range of possibilities can be defined to be some theory or system of representation which *generates* the possibility or set of possibilities, or rather representations or descriptions thereof. Even more briefly, an explanation of a range of possibilities is a grammar for those possibilities. There is much to be clarified in these formulations, but first some examples from the history of science.

Examples of theories purporting to explain possibilities

The examples which follow are not all *correct* explanations. Some have already been superseded and others probably will be. The ancient theory of epicycles used geometry to explain how it was possible for the apparent paths of planets to exhibit irregularities while the actual paths were constructed out of regular circular motions. Known forms of motion were compounded in a representation of new ones.

The atomic theory after Dalton explained how various kinds of chemical transformations were possible without any change in basic substances. (It also explained why the range of possibilities was restricted according to the laws of constant and multiple proportions, so that it was vastly superior to previous atomic theories.)

The theory of natural selection explained how it was possible for undirected (' random') mutations to lead to apparently purposive or goal-directed changes in biological species. The theory of genes explained how it was possible for offspring to inherit some but not all of the characteristics of each parent, and for different siblings to inherit different combinations. The theory that atoms were composed of protons, neutrons and electrons explained many of the possibilities summarised in the periodic table of the elements, and explained

how it was possible for one element to be transformed into another.

Einstein's theories of relativity explained how it is possible for mass and energy to be interconvertible, and for light rays to be curved even in a vacuum. Other possibilities explained before specimens were produced include lasers and super-conductivity.

The examples given so far are theories which not only explained possibilities, but also contained enough detail to make prediction and, in some cases, control, possible. In the case of the human sciences this is rare. Marx's social theories explained how it was possible for large numbers of people to collaborate peacefully in social and economic practices against their own interest. He also explained how it was possible for such systems to generate forces tending to their own overthrow.

Popper has tried to explain how it is possible for the growth of scientific knowledge to be based on rational comparisons and assessment of theories, even though no theory can ever be proved to be right or even probable.

Chomsky's theory that human minds contain representations of generative grammars explains how it is possible for sentences never before heard or uttered nevertheless to be part of a person's language. The theory (see T. Winograd, 1973) that human minds contain certain sorts of procedures or programs explains how it is possible for new sentences to be produced or understood.

Freud's theories explained how it is possible for apparently meaningless slips and aberrations of behaviour to be significant actions. Piaget's theories about the structure of many familiar concepts explain how it is possible for a child to show in some behaviour that he has grasped the concept and in others that he has not.

Known possibilities for which explanations are still lacking include the following. The possibility of the growth of an oak from an acorn or a chicken from an egg. Fragments of the mechanisms are of course understood already, but there is as yet no explanation of how such an apparently simple structure as a seed or fertilised ovum can *control* its own development in such a way as to produce such an apparently complex structure as a plant or animal.

Another unexplained possibility is the evolution of animals with specific intelligent abilities (like the ability to learn to use tools, or to learn to use language) from species lacking these abilities, and in particular the evolution of human beings. In the case of human psychology, there are very many possibilities taken for granted as part of common sense, yet still without even fragmentary explanations, for instance the possibility of a newborn infant learning whatever human language happens to be spoken around it, the possibility of relating one's actions to tastes, preferences, principles, hopes, fears, knowledge, abilities, and social commitments, and the possibility of changing one's moral attitudes through personal experience.

Formal requirements for explanations of possibilities

The explanations listed earlier may not be *correct* explanations, but they at least meet formal conditions for explaining certain possibilities, or perhaps would do if precisely formulated. These conditions will be described below. They are generalisations and elaborations of the basic idea, familiar from writings of philosophers like Popper, Hempel and Nagel, that to explain something by means of a theory is to deduce it from the theory, perhaps with some additional premisses. As normally formulated, this assumes that both the theory and what it explains are expressed in the form of sentences, using natural language supplemented by the technical language of the science concerned.

It is also normally assumed that the deduction is *logical*, that is the inference from the theory to what it explains can be shown to be valid according to the rules of inference codified by logicians.

(This is sometimes generalised to permit cases where the inference is only probabilistic.)

This concept of deduction and the related notion of explanation needs to be generalised in two ways.

First of all, other means of representation besides sentences may be used, such as maps, diagrams, three-dimensional models or computer programs.

Secondly, the *forms of inference* include not only the logical forms (like 'All A's are B's, All B's are C's. Therefore All A's are C's'), but also the manipulation of other representations (see Sloman, 1971a). An example is the manipulation of diagrams representing molecular structures, in order to explain the possibility of chemical reactions, like the production of water from hydrogen and oxygen.

On this view the use of models and so-called 'analogies' in science is simply a change of language: one configuration is used to represent another.

All the usual talk about isomorphism of models in this context is as misconceived as the theory that sentences in natural language must be isomorphic with things they describe: there are many more kinds of non-verbal representations than isomorphic models. (See Goodman, 1968; Clowes, 1971; and Toulmin, 1953. I was helped to see all this by an unpublished paper by Max Clowes, called 'Paradigms and syntactic models'.)

Further requirements for explanations of possibilities

We now have a minimal requirement for a theory T formulated in sentences or other symbolic apparatus to be an explanation of some range of possibilities, namely:

(1) Statements or other representations of the range of possibilities should be validly derivable from T, according to whatever criteria for validity are associated with the 'language' of T.

Perhaps one of the most important illustrations of this is the use of the theory of bonds between atoms (the theory of valencies) to explain the possibility of a very large number of chemical compounds and transformations. Knowing the kinds of bonds into which the various atoms can enter, one can generate representations of large numbers of chemical compounds, and chemical reactions, using diagrams of the sorts mentioned previously. Here one range of (relatively primitive) possibilities is used to explain another range. The possibility of many kinds of atoms with different chemical bonding potentials was itself explained later on by a more economical theory which assumed atoms could be made up of a nucleus containing positively charged protons, neutrons with no charge, and electrons with negative charge. Thus, postulating a small number of primitive subatomic components and principles according to which they could be combined into atoms, physicists could generate representations of a wide range of possible atoms, and therefore of possible molecules and reactions.

These theories eventually had to be revised and refined of course, but that does not affect the point that at least part of the scientific function of those theories while they survived was to explain a range of possibilities according to criterion (1). While they worked, they provided 'generative grammars' for known ranges of possibilities.

However, there are additional requirements if T is to be a *good* explanation of the possibilities in question. Rival theories are assessed according to how well they meet these additional requirements, namely:

(2) The theory should be as *definite* as possible: i.e. there should be a clear demarcation between what it does and what it does not explain.

(3) T should be *general*, i.e. it should explain many different possibilities, preferably including some possibilities not known about before the theory was invented. This criterion should be used with caution. Insofar as a theory generates some possibilities not yet established by actual instances, efforts should be made to find or create instances, and they should not be types of things thought to be impossible because of some more general theory. If repeated efforts to find actual instances fail, this does not disprove the theory, but it does reduce its credit. So a theory should not explain too many things.

(4) T should account for *fine structure*: i.e. the descriptions or representations of possibilities generated by T should be rich and detailed.

(5) T should be non-circular: i.e. the possibilities assumed in T should not be of essentially the same character as the possibilities T purports to explain.

(6) The derivations from T should be *rigorous*: i.e. within the range of possibilities explained by T, the procedures by which those possibilities are deduced or derived should be explicitly specified so that they can be publicly assessed, and not left to the intuitions of individuals.

(7) The theory T should be *plausible*: that is, insofar as it makes any assertions or has any presuppositions about what is the case or what is possible, these should not contradict any known facts. However, sometimes the development of a new theory may lead to the refutation of previously widely held beliefs, so this criterion has to be used with great discretion.

(8) The theory should be *economical*: i.e. it should not include assumptions or concepts which are not required to explain the possibilities it is used to explain. Sometimes economy is taken to mean the use of relatively few concepts or assumptions, from which others can be derived as necessary. The latter is not always a good thing to stress, since great economy in primitive concepts can go along with uneconomical derivations and great difficulty of doing anything with the theory, that is, with heuristic poverty.

(9) The theory should be *rich in heuristic power*: i.e. the concepts, assumptions, symbolisms, and transformation procedures of the theory should be such as to make the detection of gaps and errors, the structuring of problem-solving strategies, the recognition of relevant evidence, and so on, easily manageable. (See McCarthy and Hayes, 1969 and my 1971a for more on this)

These criteria therefore indicate ways in which theories explaining possibilities may be criticised rationally. For instance, one may be able to show (by a logical or mathematical argument) that the theory does not in fact generate the range of possibilities it is said to explain.

(Nearly all psychological theories put forward to explain known human possibilities, such as perception, fail on this point: the theories generate the required range of possibilities only in the mind of a sympathetic audience supplying a large and unspecified set of additional assumptions.)

A theory may be criticised by showing that it explains too much, including things which so far appear to be impossible. The theory may not explain enough of the known fine structure of the possibilities (like theories of speech understanding which don't explain how hearers can cope with complex syntactic ambiguities, or developmental theories in biology which don't explain how a chicken's egg can grow into something like its mother or father in so many detailed ways). The explanation may be circular, like theories which attempt to explain human mental functioning by assuming the existence of a spirit or soul with essentially all the abilities it is intended to explain. The theory may be so indefinite that it is not clear what it does and what it does not explain.

A theory may also be criticised less directly by criticising the specification of the range of possibilities which it is meant to explain (e.g. criticising the typology on which it is based).

For instance the specification may describe a set of structures in ways which are not related to their functions, like describing sentences in terms of transition probabilities between successive words. Or the set of possibilities explained may be shown to be only a sub-range of some wider set of

possibilities which the theory cannot cope with.

For instance, a theory which explains how statements are constructed and understood can be criticised if it cannot be extended to account for questions, commands, threats, requests, promises, bets, contracts, and other types of verbal communication which are clearly functionally related to statements in that they use related syntactic structures and almost the same vocabulary. If it turns out that a physical theory of the interactions of atoms and their components can only explain the possibility of chemical reactions involving relatively simple molecules, then that will show an inadequacy in the theory. Similarly, if an economic theory can explain only the possibility of economic processes occurring when there is a very restricted amount of information flow in a community, then that theory is not good enough. Finally, if a theory of the function of moral language accounts only for abusive and exhortative uses of that kind of language, then it is clearly inadequate since moral language can be used in a much wider range of ways.

In some cases, whether a theory explaining some specified range of possibilities satisfies these criteria or not, or whether it satisfies them better than a rival theory, is not an empirical question. It is a question to be settled by logical and mathematical investigations of the structure of the theory and of what can be derived from it.

Sometimes the theory is too complex for its properties to be exhaustively surveyed. If so, one can only try out various derivations or manipulations in test cases. This is partly analogous to an empirical investigation in that the results are always partial and cannot be worked out in advance by normal human reasoning. Similarly testing a complex computer program may feel like conducting some kind of experiment. Nevertheless, as already remarked, the connections so discovered are not empirical, but logical or mathematical in nature.

(The criteria listed here can be justified by showing how using them is necessary for furthering the interpretative and practical aims of science listed in part one.)

Prediction and testing

A theory may meet the conditions above without being of any use in predicting or explaining particular events or in enabling events or processes to be controlled. This is why I have stressed the explanation of *possibilities*. Although it explains how certain sorts of phenomena are possible, the underlying mechanism or structure postulated may, at the time the theory is proposed, be unobservable, so that observation of its state cannot be used to predict actual occurrences of those phenomena.

Similarly, no techniques may be available for manipulating the mechanisms, so that the theory provides no basis for controlling the phenomena. For instance, the theory of evolution explains the possibility of a wide range of biological developments without providing a basis for predicting them. Similarly, a

theory explaining the possibility of my uttering sentences of particular forms need not provide any basis for predicting when I will utter anyone sentence, or for making me utter it, or even for explaining exactly why I uttered the particular sentence I did utter at a particular time. This is because the theory may simply postulate a certain kind of sentence-generating mechanism, available in my mind as a resource to be used along with other resources.

How any particular resource is used on any particular occasion, may be the result of myriad complex interactions between such factors as my purposes, preferences, hopes, fears and moral principles, what I believe to be the case at the time, what I know about the likely effects of various actions, how much I am distracted and so on. The theory which explains the possibility of generating and under standing sentences need not specify all the interactions between the postulated mechanism and other aspects of the mind. So it need not provide a basis for prediction and control.

This is true of any explanation of an ability, skill, talent or power, in terms of a mechanism making it possible. The explanation need not specify the rest of the system of which that resource is a part, nor specify the conditions under which the resource is used. And even if it does, the specification need not refer to either observable conditions or manipulable conditions. So such explanations of possibilities, though they contribute to scientific understanding, need not contribute to predictions of actual events.

It is not possible to refute such a theory, if it merely explains possibilities, and entails or explains no impossibilities. For it is a fact about the logic of possibility that 'X is possible' does not entail 'X will occur at some time or other'.

Similarly 'X never occurs' does not entail 'X is impossible'. Newtonian mechanics entails that it is possible for some very large body passing near the earth to deflect the earth from its orbit, and it explains this possibility: but the fact that this never occurs casts no doubt on the theory. Similarly, a grammatical theory may explain the possibility of the utterance of a certain rather complex English sentence, and even though nobody ever utters that sentence naturally, this casts no doubt on the theory. A psychological theory may imply that it is possible for a human being to count backwards from 99 to 0 to the tune of 'Silent night, holy night', without being refuted merely by the fact that nobody ever does this. Only a much more complex theory, taking into account a rich set of motives and beliefs, could ever be used to predict such a performance, and perhaps be refuted by its non-occurrence.

Lack of predictive power, practical utility, or refutability need not prevent the scientific merits of an explanation of a range of possibilities from being discussed rationally, and compared with the merits of rival explanations, in accordance with the criteria listed above. Nor does it prevent such a theory from giving deep insight, of a kind which provides a firm basis for building more elaborate theories which do permit predictions and explanations of particular events, and which are empirically refutable.

I therefore see no reason for calling such theories nonsensical, as some of the logical positivists would, nor for banishing them from the realm of science into metaphysics, as Popper does (though he admits that metaphysical theories may be rationally discussable and may be a useful stimulus to the development of what he calls scientific theories).

I am not here arguing over questions of meaning: define 'science' as you will, my point remains that among the major merits of the generally agreed most profound scientific theories is the fact that they satisfy the above criteria for being good explanations of possibilities, and therefore give us good insights into the nature of the kinds of objects, events or processes that can exist or occur in the universe. If unrefutable theories are to be dubbed 'metaphysical', then what I am saying is that even important scientific theories have a metaphysical component, and that the precision, generality, fine structure, non-circularity, rigour, plausibility, economy and heuristic power of the metaphysical component are among the objective criteria by which scientific theories are in fact assessed (and should be assessed).

The development of such 'metaphysical' theories is so intimately bound up with the development of science that to insist on a demarcation is to make a trivial semantic point, of no theoretical interest. Moreover, it has bad effects on the training of scientists.

Empirical support for explanations of possibilities

Further, even though a theory explaining only certain possibilities is not refutable empirically, that does not mean that empirical evidence is wholly irrelevant to it. For instance, if a kind of possibility explained by the theory is observed for the first time after the theory was constructed, then this is empirical corroboration for the theory, even though the theory did not specify that the phenomenon ever would occur, or that it would occur in the particular conditions in which it did. Observing an actual instance of a possibility explained by some theory provides support for that theory at least to the extent of showing that there is something for it to explain: it shows that the theory performs a scientific function.

However, the support *adds* to previous knowledge only if it is a new kind of possibility. Mere repetition of observations or experiments does not increase support for a theory: it merely checks that no errors were made in previous instances.

In those contexts all the normal stress on repeatability of scientific experiments is unnecessary and has misled many psychologists and social scientists into making impossible demands of empirical studies of man and society. Repetition may be a useful check on whether the phenomenon really is possible (since it permits more independent witnesses to observe it), and it provides opportunities for more detailed examination of exactly what occurred, but is not logically necessary. If a phenomenon occurs only once, then it is possible, and its possibility needs explaining.

Any explanation of that possibility is therefore not gratuitous, and the only question that should then arise is not whether the explanation is science or pseudo-science, or metaphysics, but whether a *better* explanation can be found for the same possibility, that is, an explanation meeting more of the criteria (2) to (9) above; or perhaps serving additional scientific aims besides explaining possibilities.

The frantic pursuit of repeatability and statistically significant correlations is based on a belief that science is a search for laws. This has blinded many scientists to the need for careful description and analysis of what *can* occur, and for the explanation of its possibility.

Instead they try to find what *always* occurs -- a much harder task -- and usually fail. Even if something is actually done by very few persons, or only by one, that still shows that it is possible for a human being, and this possibility needs explanation, as much as any other established fact. This justifies elaborate and detailed investigation and analysis of particular cases: a task usually shirked because of the search for statistically significant correlations. Social scientists have much to learn from historians and students of literature -- despite all the faults of the latter.

I have gone on at such great length about describing and explaining possibilities because the matter is not generally discussed in books on philosophy of science. But I do not wish to deny the importance of trying to construct theories which can be used to explain and predict what actually occurs, or which explain impossibilities and observed regularities. Of two theories explaining the same range of possibilities, one which also explains more impossibilities and permits a wider variety of predictions and explanations of actual events to be made on the basis of observation, is to be preferred, since it serves to a greater degree the aims of science listed previously. I think it is quite premature to seek such predictive explanations in psychology and social science: these sciences still seem too far from having good explicit descriptions and explanations of possibilities we all know about, as linguistics was until recently.

This discussion is still very sketchy and unsatisfactory. Much finer description and classification of different sorts of explanations is required. But enough for now!

V
Some thoughts on Popper

For reasons which I do not fully understand, Popper is apparently strongly opposed to all this talk of concepts and possibilities (see, for instance, pp123-4 of his (1972) where he describes it as an error to think that concepts and conceptual systems or problems about meaning are comparable in importance to theories and theoretical systems, or to problems of truth.) As far as I can tell, his argument rests on the curious assumption that concepts or meanings are purely subjective things, and that the only objective criteria by which they can be assessed or criticised are ones which concern the truth of statements and theories containing them. I hope I have said enough to refute

these claims. Roughly, our disagreement seems to hinge on Popper's view that the only place for rationality in science is in the selection from among hypotheses expressible in a given language, whereas I have tried to show that there are rational ways of deciding how to extend a language, and therefore how to extend the set of expressible hypotheses.

I admit that there are still serious gaps in my discussion: a theory of concept-formation is still lacking.

Finally, even if it is agreed that science uses rational *means* to pursue the aims described here, the question arises: are these *aims* rational? Is it rational to pursue them? I believe there is no answer to this. If someone genuinely prefers the life of a mystic or hermit or 'primitive' tribesman to the pursuit of knowledge and understanding of the universe, then that preference must be respected. However, I believe that the aims and criteria described here are part of the mental mechanism with which every human child is born -- but for which it would not be possible to learn all that human children do learn. So one can reject science only after one has used it, however unconsciously, for some years.

THANKS

REFERENCES

Chomsky, N. (1957) Syntactic structures, Mouton, 'The Hague

Clowes, M. (1971) 'On Seeing Things', Artificial Intelligence, 2, 1971

Gibbs, B,.R. (1970) 'Real Possibility', American Philosophical Quarterly, October 1970

Goodman, N. (1968) Languages of Art, Oxford University Press

Kant, I. (1781) Critique of Pure Reason, translated by N.K. Smith, Macmillan, 1929

Kuhn, T.S. (1962) The Structure of Scientific Revolutions, Phoenix Books

McCarthy, J. and Hayes, P. (1969) 'Some philosophical problems from the standpoint of artificial intelligence', in Machine Intelligence 4, ed. D. Michie and B. Meltzer, Edinburgh University Press

Popper, K.R. (1972) Objective Knowledge, Oxford

University Press

Sloman, A. (1971a). 'Interactions between philosophy
and Artificial Intelligence: the role of intuition
and non-logical reasoning in intelligence', in
Proc. 2nd International Joint Conference on Artificial
Intelligence (The British Computer Society,
1971), and in Artificial Intelligence 2 pp209-25, 1971

Sloman, A. (1971b) 'Tarski,Frege and the Liar Paradox',
in Philosophy XLVI, April 1971

Toulmin, S. (1953) The Philosophy of Science,
Hutchinson

Watson, J. (1968) The Double Helix, Penguin Books

Winograd, T. (1973) Understanding Natural Language
Edinburgh University Press

Wittgenstein, L. (1956) Remarks on the Foundations
of Mathematics, Basil Blackwell, Oxford