

EPISTEMOLOGY WITHOUT A KNOWING SUBJECT

K. R. POPPER

University of London, England

Allow me to start with a confession. Although I am a very happy philosopher I have, after a lifetime of lecturing, no illusions about what I can convey in a lecture. For this reason I shall make no attempt in this lecture to convince you. Instead I shall make an attempt to challenge you, and, if possible, to provoke you.

1. Three theses on epistemology and the third world

I might have challenged those who have heard of my adverse attitude towards Plato and Hegel by calling my lecture '*A theory of the Platonic world*', or '*A theory of the objective spirit*'.

The main topic of this lecture will be what I often call, for want of a better name, '*the third world*'. To explain this expression I will point out that, without taking the words 'world' or 'universe' too seriously, we may distinguish the following three worlds or universes: first, the world of physical objects or of physical states; secondly, the world of states of consciousness, or of mental states, or perhaps of behavioural dispositions to act; and thirdly, the world of *objective contents of thought*, especially of scientific and poetic thoughts and of works of art.

Thus what I call 'the third world' has admittedly much in common with Plato's theory of forms or ideas, and therefore also with Hegel's objective spirit, though my theory differs radically, in some decisive respects, from Plato's and Hegel's. It has more in common still with Bolzano's theory of a universe of propositions in themselves and of truths in themselves, though it differs from Bolzano also. My third world resembles most closely the universe of Frege's objective contents of thought.

It is not part of my view or of my argument that we might not enumerate our worlds in different ways, or not enumerate them at all. We might,

especially, distinguish more than three worlds. My term 'the third world' is merely a matter of convenience.

In upholding an objective third world I hope to provoke those whom I call '*belief philosophers*': those who, like Descartes, Locke, Berkeley, Hume, Kant, or Russell, are interested in our subjective beliefs, and their basis or origin. Against these belief philosophers I urge that our problem is to find better and bolder theories; and that *critical preference* counts, but *not belief*.

I wish to confess, however, at the very beginning, that I am a realist: I suggest, somewhat like a naive realist, that there is a physical world and a world of states of consciousness, and that these two interact. And I believe that there is a third world, in a sense which I shall explain more fully.

Among the inmates of my 'third world' are, more especially, *theoretical systems*; but just as important inmates are *problems* and *problem situations*. And I will argue that the most important inmates of this world are *critical arguments*, and what may be called – in analogy to a physical state or to a state of consciousness – *the state of a discussion* or the *state of a critical argument*; and, of course, the contents of journals, books and libraries.

Most opponents of the thesis of an objective third world will of course admit that there are problems, conjectures, theories, arguments, journals and books. But they usually say that all these entities are, essentially, symbolic or linguistic *expressions* of subjective mental states, or perhaps of behavioural dispositions to act; further, that these entities are means of *communication*, that is to say, symbolic or linguistic means to evoke in others similar mental states or behavioural disposition to act.

Against this, I have often argued that one cannot relegate all these entities and their content to the second world.

Let me repeat one of my standard arguments* for the (more or less) *independent existence of the third world*.

I consider two thought experiments:

Experiment (1). All our machines and tools are destroyed, also all our subjective learning, including our subjective knowledge of machines and tools, and how to use them. But *libraries and our capacity to learn from them* survive. Clearly, after much suffering, our world may get going again.

Experiment (2). As before, machines and tools are destroyed, and our subjective learning, including our subjective knowledge of machines and tools, and how to use them. But this time, *all libraries are destroyed also*, so that our capacity to learn from books becomes useless.

* The argument is adapted from POPPER [1962] Vol. II; cp. p. 108.

If you think about these two experiments, the reality, significance and the degree of autonomy of the third world (as well as its effects on the second and first worlds) may perhaps become a little clearer to you. For in the second case there will be no re-emergence of our civilization for many millennia.

I wish to defend in this lecture three main theses, all of which concern epistemology. Epistemology I take to be the theory of *scientific knowledge*.

My first thesis is this. Traditional epistemology has studied knowledge or thought in a subjective sense – in the sense of the ordinary usage of the words ‘I know’ or ‘I am thinking’. This, I assert, has led students of epistemology into irrelevancies: while intending to study scientific knowledge, they studied in fact something which is of no relevance to scientific knowledge. For *scientific knowledge* simply is not knowledge in the sense of the ordinary usage of the words ‘I know’. While knowledge in the sense of ‘I know’ belongs to what I call the ‘second world’, the world of *subjects*, scientific knowledge belongs to the third world, to the world of objective theories, objective problems, and objective arguments.

Thus my first thesis is that the traditional epistemology, of Locke, Berkeley, Hume, and even of Russell, is irrelevant, in a pretty strict sense of the word. It is a corollary of this thesis that a large part of contemporary epistemology is irrelevant also. This includes modern epistemic logic, *if we assume that it aims at a theory of scientific knowledge*. However, any epistemic logician can easily make himself completely immune from my criticism, simply by making clear that he does not aim at contributing to the *theory of scientific knowledge*.

My first thesis involves the existence of two different senses of knowledge or of thought: *knowledge or thought in the subjective sense*, consisting of a state of mind or of consciousness or a disposition to behave or to react, and *knowledge in an objective sense*, consisting of problems, theories, and arguments as such. Knowledge in this objective sense is totally independent of anybody’s claim to know; also it is independent of anybody’s belief, or disposition to assent; or to assert, or to act. Knowledge in the objective sense is *knowledge without a knower*: it is *knowledge without a knowing subject*.

Of thought in the objective sense Frege wrote: ‘I understand by a *thought* not the subjective act of thinking but its *objective content*...’.*

The two senses of thought and their interesting interrelations can be illustrated by the following highly convincing quotation from HEYTING [1962] p. 195 who says about Brouwer’s act of inventing his theory of the continuum:

* Cp. FREGE [1892] p. 32; italics mine.

'If recursive functions had been invented before, he [Brouwer] would perhaps not have formed the notion of a choice sequence which, I think, would have been unlucky.'

This quotation refers on the one hand to some *subjective thought processes* of Brouwer's and says that they might not have occurred (which would have been unfortunate) had the *objective problem situation* been different.

Thus Heyting mentions certain possible *influences* upon Brouwer's subjective thought processes, and he also expresses his opinion regarding the value of these subjective thought processes. Now it is interesting that influences, *qua* influences, must be subjective: only Brouwer's subjective acquaintance with recursive functions could have had that unfortunate effect of preventing him from inventing free choice sequences.

On the other hand, the quotation from Heyting points to a certain objective relationship between the *objective contents* of two thoughts or theories: Heyting does not refer to the subjective conditions or the electrochemistry of Brouwer's brain processes, but to an *objective problem situation in mathematics* and its possible influences on Brouwer's subjective acts of thought which were bent on solving these objective problems. I would describe this by saying that Heyting's remark is about the objective or third-world *situational logic* of Brouwer's invention, and that Heyting's remark implies that the third-world situation may affect the second world. Similarly, Heyting's suggestion that it would have been unfortunate if Brouwer had not invented choice sequences is a way of saying that the *objective content* of Brouwer's thought was valuable and interesting; valuable and interesting, that is, in the way it changed the objective problem situation in the third world.

To put the matter simply, if I say 'Brouwer's thought was influenced by Kant' or even 'Brouwer rejected Kant's theory of space' then I speak at least partly about acts of thought in the subjective sense: the word 'influence' indicates a context of thought processes or acts of thinking. If I say, however, 'Brouwer's thought differs vastly from Kant's', then it is pretty clear that I speak mainly about contents. And ultimately if I say 'Brouwer's thoughts are incompatible with Russell's', then, by using a *logical term* such as '*incompatible*', I make it unambiguously clear that I am using the word 'thought' only in Frege's objective sense, and that I am speaking only about the objective content, or the logical content, of theories.

Just as ordinary language unfortunately has no separate terms for 'thought' in the sense of the second world and in the sense of the third world, so it has no separate terms for the corresponding two senses of 'I know' and of 'knowledge'.

In order to show that both senses exist, I will first mention three subjective or second-world examples:

- (1) 'I *know* you are trying to provoke me, but I will not be provoked.'
- (2) 'I *know* that Fermat's last theorem has not been proved, but I believe it will be proved one day.'
- (3) From the entry 'Knowledge' in *The Oxford English dictionary*: *knowledge* is a 'state of being aware or informed'.

Next I will mention three objective or third-world examples:

- (1) From the entry 'Knowledge' in *The Oxford English dictionary*: *knowledge* is a 'branch of learning; a science; an art'.
- (2) 'Taking account of the present state of *metamathematical knowledge*, it seems possible that Fermat's last theorem may be undecidable.'
- (3) 'I certify that this thesis is an original and significant *contribution to knowledge*.'

These very trite examples have only the function of helping to clarify what I mean when I speak of 'knowledge in the objective sense'. My quoting *The Oxford English dictionary* should not be interpreted as either a concession to language analysis or as an attempt to appease its adherents. It is not quoted in an attempt to prove that 'ordinary usage' covers 'knowledge' in the objective sense of my third world. In fact, I was surprised to find in *The Oxford English dictionary* examples of objective usages of 'knowledge'. (I was even more surprised to find some at least *partly* objective usages of 'know': 'to distinguish, to be acquainted with (a thing, a place, a person); ... to understand'. That these usages may be partly objective will emerge from the sequel*.) At any rate, my examples are not intended as arguments. They are intended solely as illustrations.

My *first thesis*, so far not argued but only illustrated, was that traditional epistemology with its concentration on the second world, or on knowledge in the subjective sense, is irrelevant to the study of scientific knowledge.

My *second thesis* is that what is relevant for epistemology is the study of scientific problems and problem situations, of scientific conjectures (which I take as merely another word for scientific hypotheses or theories), of scientific discussions, of critical arguments, and of the role played by evidence in arguments; and therefore of scientific journals and books, and of experiments and their evaluation in scientific arguments; or, in brief: that the study of a *largely autonomous* third world of objective knowledge is of decisive importance for epistemology.

* See section 7.1, below.

An epistemological study as described in my second thesis shows that scientists very often do not claim that their conjectures are true, or that they 'know' them in the subjective sense of 'know', or that they believe in them. Although they do not, in general, claim to know, they do, in their research programmes, act on the basis of guesses about what is and what is not fruitful, and what line of research promises further results in the third world of objective knowledge. In other words, scientists act on the basis of a guess or, if you like, of a *subjective belief* (for we may so call the subjective basis of an action) concerning what is promising of impending *growth in the third world of objective knowledge*.

This, I suggest, furnishes an argument in favour of both my *first thesis* (of the irrelevance of a subjectivist epistemology) and of my *second thesis* (of the relevance of an objectivist epistemology).

But I have a *third thesis*. It is this. An objectivist epistemology which studies the third world can help to throw an immense amount of light upon the second world of subjective consciousness, especially upon the subjective thought processes of scientists; but *the converse is not true*.

These are my three main theses.

In addition to my three main theses, I offer three supporting theses.

The first of these is that the third world is a natural product of the human animal, comparable to a spider's web.

The second and I think an almost crucial thesis is that the third world is largely *autonomous*, even though we constantly act upon it and are acted upon by it: it is autonomous in spite of the fact that it is our product and that it has a strong feed-back effect upon us; that is to say, upon us *qua* inmates of the second and even of the first world.

The third supporting thesis is that it is through this interaction between ourselves and the third world that objective knowledge grows, and that there is a close analogy between the growth of knowledge and biological growth; that is, the evolution of plants and animals.

2. A biological approach to the third world

In the present section of my talk I shall try to defend the existence of an autonomous world by a kind of biological or evolutionary argument.

A biologist may be interested in the behaviour of animals; but he may also be interested in some of the *non-living structures* which animals produce, such as spiders' webs, or nests built by wasps or ants, the burrows of badgers,

dams constructed by beavers, or paths made by animals in forests.

I will distinguish between two main categories of problems arising from the study of these structures. The first category consists of problems concerned with *the methods used* by the animals, or *the ways the animals behave* when constructing these structures. This first category thus consists of *problems concerned with the acts of production*, and with the relationships between the animal and the product. The second category of problems is concerned with the *structures themselves*. It is concerned with the chemistry of the materials used in the structure; with their geometrical and physical properties; with their dependence upon or their adjustment to special environmental conditions. *Very important* also is the *feed-back relation* from the properties of the structure to the behaviour of the animals. In dealing with this second category of problems – that is with the structures themselves – we shall also have to look upon the structures from the point of view of their biological *functions*. Thus some problems of the first category will admittedly arise when we discuss problems of the second category; for example ‘How was this nest built?’ and ‘What aspects of its structure are typical (and thus presumably traditional or inherited) and what aspects are variants adjusted to special conditions?’.

As my last example of a problem shows, problems of the first category – that is, problems concerned with the production of the structure – will sometimes be suggested by problems of the second category. This must be so, since both categories of problems are dependent upon *the fact that such objective structures exist*, a fact which itself belongs to the second category. Thus the existence of the *structures themselves* may be said to create both categories of problems. We may say that the second category of problems – problems connected with the structures themselves – is more fundamental: all that it presupposes from the first category is the bare fact that the structures are somehow *produced* by some animals.

Now these simple considerations may of course also be applied to products of *human* activity, such as houses, or tools, and also to works of art. Especially important for us, they apply to what we call ‘language’, and to what we call ‘science’.*

The connection between these biological considerations and the topic of my present lecture can be made clear by reformulating my three main theses. My first thesis can be put by saying that in the present problem situation in philosophy, few things are as important as the awareness of the distinction

* On these ‘artifacts’ cp. HAYEK [1967] p. 111.

between the two categories of problems – production problems on the one hand and problems connected with the produced structures themselves on the other. My second thesis is that we should realize that the second category of problems, those concerned with the products in themselves, is in almost every respect more important than the first category, the problems of production. My third thesis is that the problems of the second category are basic for understanding the production problems: contrary to first impressions, we can learn more about production behaviour by studying the products themselves than we can learn about the products by studying production behaviour. This third thesis may be described as an anti-behaviouristic and anti-psychologistic thesis.

In their application to what may be called ‘knowledge’ my three theses may be formulated as follows.

(1) We should constantly be aware of the distinction between problems connected with our personal contributions to the production of scientific knowledge on the one hand, and problems connected with the structure of the various products, such as scientific theories or scientific arguments, on the other.

(2) We should realize that the study of the products is vastly more important than the study of the production, even for an understanding of the production and its methods.

(3) We can learn more about the heuristics and the methodology and even about the psychology of research by studying theories, and the arguments offered for or against them, than by any direct behaviouristic or psychological or sociological approach. In general, we may learn a great deal about behaviour and psychology from the study of the products.

In what follows I will call the approach from the side of the products – the theories and the arguments – the ‘objective’ approach or the ‘third-world’ approach. And I will call the behaviourist, the psychological, and the sociological approach to scientific knowledge the ‘subjective’ approach or the ‘second-world’ approach.

The appeal of the subjective approach is largely due to the fact that it is *causal*. For I admit that the objective structures for which I claim priority are caused by human behaviour. Being causal, the subjective approach may seem to be more scientific than the objective approach which, as it were, starts from effects rather than causes.

Though I admit that the objective structures are products of behaviour, I hold that the argument is mistaken. In all sciences, the ordinary approach is from the effects to the causes. The effect raises the problem – the problem

to be explained, the explicandum – and the scientist tries to solve it by constructing an explanatory hypothesis.

My three main theses with their emphasis on the objective product are therefore neither teleological nor unscientific.

3. The objectivity and the autonomy of the third world

One of the main reasons for the mistaken subjective approach to knowledge is the feeling that a book is nothing without a reader: only if it is understood does it really become a book; otherwise it is just paper with black spots on it.

This view is mistaken in many ways. A wasp's nest is a wasp's nest even after it has been deserted; even though it is never again used by wasps as a nest. A bird's nest is a bird's nest even if it was never lived in. Similarly a book remains a book – a certain type of product – even if it is never read (as may easily happen nowadays).

Moreover, a book, or even a library, need not even have been written by anybody: a series of books of logarithms, for example, may be produced and printed by a computer. It may be the best and fullest series of books of logarithms – it may contain logarithms up to, say, 50 decimals. It may be sent out to libraries, but it may be found too cumbersome for use; at any rate, years may elapse before anybody uses it; and many figures in it (which represent mathematical theorems) may never be looked at as long as men live on earth. Yet each of these figures contains what I call 'objective knowledge'; and the question of whether or not I am entitled to call it by this name is of no interest.

The example of these books of logarithms may seem far-fetched. But it is not. I should say that almost every book is like this: it contains objective knowledge, true or false, useful or useless; and whether anybody ever reads it and really grasps its contents is almost accidental. A man who reads a book with understanding is a rare creature. But even if he were more frequent, there would always be plenty of misunderstandings and misinterpretations; and it is not the actual and somewhat accidental avoidance of such misunderstandings which turns black spots on white paper into a book, or an instance of knowledge in the objective sense. Rather, it is something more abstract. It is the possibility or potentiality of being understood, the dispositional character of being understood or interpreted, or misunderstood or misinterpreted, which makes of a thing a book. And this potentiality or disposition may exist without ever being actualized or realized.

To see this more clearly, we may imagine that after the human race has perished, some books or libraries may be found by some civilized successors of ours (no matter whether these are terrestrial animals which have become civilized, or some visitors from outer space). These books may be decyphered. They may be those logarithm tables never read before, for argument's sake. This makes it quite clear that neither its composition by thinking animals nor the fact that it has not actually been read or understood is essential for making a thing a book, and that it is sufficient that it can be decyphered.

Thus I do admit that in order to belong to the third world of objective knowledge, a book should – in principle, or virtually – be capable of being grasped (or decyphered, or understood, or 'known') by somebody. But I do not admit more.

We can thus say that there is a kind of Platonic (or Bolzanoesque) third world of books in themselves, theories in themselves, problems in themselves, problem situations in themselves, arguments in themselves, and so on. And I assert that even though this third world is a human product, there are many theories in themselves and arguments in themselves and problem situations in themselves which have never been produced or understood and may never be produced or understood by men.

The thesis of the existence of such a third world of problem situations will strike many as extremely metaphysical and dubious. But it can be defended by pointing out its biological analogue. For example, it has its full analogue in the realm of birds' nests. Some years ago I got a present for my garden – a nesting box for birds. It was a human product, of course, not a bird's product – just as our logarithm table was a computer's product rather than a human product. But in the context of the bird's world, it was part of an objective problem situation, and an objective opportunity. For some years the birds did not even seem to notice the nesting box. But after some years, it was carefully inspected by some blue tits who even started building in it, but gave up very soon. Obviously, here was a graspable opportunity, though not, it appears, a particularly valuable one. At any rate, here was a problem situation. And the problem may be solved in another year by other birds. If it is not, another box may prove more adequate. On the other hand, a most adequate box may be removed before it is ever used. The question of the adequacy of the box is clearly an objective one; and whether the box is ever used is partly accidental. So it is with all ecological niches. They are potentialities and may be studied as such in an objective way, up to a point independently of the question of whether these potentialities will ever be actualized by any living organism. A bacteriologist knows how to prepare

such an ecological niche for the culture of certain bacteria or moulds. It may be perfectly adequate for its purpose. Whether it will ever be used and inhabited is another question.

A large part of the objective third world of actual and potential theories and books and arguments arises as an unintended by-product of the actually produced books and arguments. We may also say that it is a by-product of human language. Language itself, like birds' nests, is an unintended by-product of actions which were directed at other aims.

How does an animal path in the jungle arise? Some animal may break through the underwood in order to get to a drinking place. Other animals find it easiest to use the same track. Thus it may be widened and improved by use. It is not planned – it is an unintended consequence of the need for easy or swift movement. This is how a path is originally made – perhaps even by men – and how language and any other institutions which are useful may arise, and how they may owe their existence and development to their usefulness. They are not planned or intended, and there was perhaps no need for them before they came into existence. But they may create a new need, or a new set of aims: the aim-structure of animals or men is not 'given', but it develops with the help of some kind of feed-back mechanism out of earlier aims, and out of results which were or were not aimed at*.

In this way, a whole new universe of possibilities or potentialities may arise: a world which is to a large extent *autonomous*.

A very obvious example is a garden. Even though it may have been planned with great care, it will as a rule turn out partly in unexpected ways. But even if it turns out as planned, some unexpected interrelationships between the planned objects may give rise to a whole universe of possibilities, of possible new aims, and of new *problems*.

The world of language, of expectations, theories and arguments, in brief, the universe of objective knowledge, is one of the most important of these man-created yet at the same time largely autonomous universes.

The idea of *autonomy* is central to my theory of the third world: although the third world is a human product, a human creation, it creates in its turn, as do other animal products, its own *domain of autonomy*.

There are countless examples. Perhaps the most striking ones, and at any

* See HAYEK [1967] Ch. 6, esp. pp. 96, 100, n. 12; DESCARTES [1637], cp. [1931] p. 89; POPPER [1960] p. 65; [1966] sect. XXIV.

rate those which should be kept in mind as our standard examples, may be found in the theory of natural numbers.

Pace Kronecker, I agree with Brouwer that the sequence of natural numbers is a human construction. But although we create this sequence, it creates, in its turn, its own autonomous problems. The distinction between odd and even numbers is not created by us: it is an unintended and unavoidable consequence of our creation. Prime numbers, of course, are similar unintended autonomous and objective facts; and in their case it is obvious that there are many facts here for us to *discover*: there are conjectures like Goldbach's. And these conjectures, though they refer indirectly to objects of our creation, refer directly to problems and facts which have somehow emerged from our creation and which we cannot control or influence: they are hard facts, and the truth about them is often hard to discover.

This exemplifies what I mean when I say that the third world is largely autonomous, though created by us.

But the autonomy is only a partial one: the new problems lead to new creations or constructions – such as recursive functions, or Brouwer's free choice sequences – and may thus add new objects to the third world. And every such step will create *new unintended facts*; *new unexpected problems*; and often also *new refutations**.

There is also a most important feed-back effect from our creations upon ourselves; from the third world upon the second world. For the new emergent problems stimulate us to new creations.

The process can be described by the following somewhat oversimplified schema (see my [1966] especially p. 24):

$$P_1 \rightarrow TT \rightarrow EE \rightarrow P_2.$$

That is, we start from some problem P_1 , proceed to a tentative solution or tentative theory TT , which may be (partly or wholly) mistaken; in any case it will be subject to error elimination EE which may consist of critical discussion or experimental tests; at any rate, new problems P_2 arise from our own creative activity; and these new problems are not in general intentionally created by us, they emerge autonomously from the field of new relationships which we cannot help bringing into existence with every action, however little we intend to do so.

The autonomy of the third world, and the feed-back of the third world

* An example of the latter is Lakatos's 'concept-stretching refutation'; see LAKATOS [1963–64].

upon the second and even the first, are among the most important facts of the growth of knowledge.

Following up our biological considerations, it is easy to see that they are of general importance for the theory of Darwinian evolution: they explain how we can lift ourselves by our own bootstraps. Or in more highbrow terminology, they help to explain 'emergence'.

4. Language, criticism, and the third world

The most important of human creations, with the most important feedback effects upon ourselves and especially upon our brains, are the higher functions of human language; more especially, the *descriptive function and the argumentative function*.

Human languages share with animal languages the two lower functions of language: (1) self-expression and (2) signalling. The self-expressive function or symptomatic function of language is obvious: all animal language is symptomatic of the state of some organism. The signalling or release function is likewise obvious: we do not call any symptom linguistic unless we assume that it can release a response in another organism.

All animal languages and all linguistic phenomena share these two lower functions. But human language has many other functions*. Strangely enough, the most important of the higher functions have been overlooked by almost all philosophers. The explanation of this strange fact is that the two lower functions are always present when the higher ones are present, so that it is always possible to 'explain' every linguistic phenomenon, in terms of the lower functions, as an '*expression*' or a '*communication*'.

The two most important higher functions of human languages are (3) the *descriptive function* and (4) the *argumentative function***.

With the descriptive function of human language, the regulative idea of *truth* emerges, that is, of a description which fits the facts***.

* For example, advisory, hortative, fictional, etc.

** See POPPER [1963] especially chapters 4 and 12, and the references on pp. 134, 293 and 295 to BÜHLER [1934]. Bühler was the first to discuss the decisive difference between the lower functions and the descriptive function. I found later, as a consequence of my theory of criticism, the decisive distinction between the descriptive and the argumentative functions. See also POPPER [1966] section XIV and note 47.

*** One of the great discoveries of modern logic was Alfred Tarski's re-establishment of the (objective) correspondence theory of truth (truth = correspondence to the facts). The present essay owes everything to this theory; but I do not of course wish to implicate Tarski in any of the crimes here committed.

Further regulative or evaluative ideas are content, truth content, and verisimilitude*.

The argumentative function of human language presupposes the descriptive function: arguments are, fundamentally, about descriptions: they criticize descriptions from the point of view of the regulative ideas of truth; content; and verisimilitude.

Now two points are all-important here:

(1) Without the development of an exosomatic descriptive language – a language which, like a tool, develops outside the body – there can be *no object* for our critical discussion. But with the development of a descriptive language (and further, of a written language), a linguistic third world can emerge; and it is only in this way, and only in this third world, that the problems and standards of rational criticism can develop.

(2) It is to this development of the higher functions of language that we owe our humanity, our reason. For our powers of reasoning are nothing but powers of critical argument.

This second point shows the futility of all theories of human language that focus on *expression and communication*. As we shall see, the human organism which, it is often said, is to express itself, depends in its structure very largely upon the emergence of the two higher functions of language.

With the evolution of the argumentative function of language, criticism becomes the main instrument of further growth. (Logic may be regarded as *the organon of criticism*; see my [1963] p. 64.) The autonomous world of the higher functions of language becomes the world of science. And the schema, originally valid for the animal world as well as for primitive man,

$$P_1 \rightarrow TT \rightarrow EE \rightarrow P_2$$

becomes the schema of the growth of knowledge through error elimination by way of systematic *rational criticism*. It becomes the schema of the search for truth and content by means of rational discussion. It describes the way in which we lift ourselves by our bootstraps. It gives a rational description of evolutionary emergence, and of our *self-transcendence by means of selection and rational criticism*.

To sum up, although the meaning of 'knowledge', like of all words, is

* See the previous note and POPPER [1962a] especially p. 292; and POPPER [1963] chapter 10 and Addenda.

unimportant, it is important to distinguish between different senses of the word.

(1) Subjective knowledge which consists of certain inborn dispositions to act, and of their acquired modifications.

(2) Objective knowledge, for example, scientific knowledge which consists of conjectural theories, open problems, problem situations, and arguments.

All work in science is work directed towards the growth of objective knowledge. We are workers who are adding to the growth of objective knowledge as masons work on a cathedral.

Our work is fallible, like all human work. We constantly make mistakes, and there are objective standards of which we may fall short – standards of truth, content, validity, and others.

Language, the formulation of problems, the emergence of new problem situations, competing theories, mutual criticism by way of argument, all these are the indispensable means of scientific growth. The most important functions or dimensions of the human language (which animal languages do not possess) are the descriptive and the argumentative functions. The growth of these functions is, of course, of our making, though they are unintended consequences of our actions. It is only within a language thus enriched that critical argument and knowledge in the objective sense become possible.

The repercussions, or the feed-back effects, of the evolution of the third world upon ourselves – our brains, our traditions (if anybody were to start where Adam started, he would not get further than Adam did) our dispositions to act (that is, our beliefs*) and our actions, can hardly be overrated.

As opposed to all this, *traditional epistemology* is interested in the second world: in knowledge as a certain kind of belief – justifiable belief, such as belief based upon perception. As a consequence, this kind of belief philosophy cannot explain (and does not even try to explain) the decisive phenomenon that scientists criticize their theories and so kill them. *Scientists try to eliminate their false theories, they try to let them die in their stead. The believer – whether animal or man – perishes with his false beliefs.*

5. Historical remarks

5.1. *Plato and Neo-Platonism*

For all we know, Plato was the discoverer of the third world. As Whitehead remarked, all Western philosophy consists of footnotes to Plato.

* The theory that beliefs may be gauged by readiness to bet was regarded as well known in 1771; see KANT [1778] p. 852.

I will make only three brief remarks on Plato, two of them critical.

(1) Plato discovered not only the third world, but part of the influence or feed-back of the third world upon ourselves: he realized that we try to grasp the ideas of his third world; also that we use them as explanations.

(2) Plato's third world was divine; it was unchanging and, of course, true. Thus there is a big gap between his and my third world: my third world is man-made and changing. It contains not only true theories but also false ones, and especially open problems, conjectures, and refutations.

And while Plato, the great master of dialectical argument, saw in it merely a way leading to the third world, I regard arguments as among the most important inmates of the third world; not to speak of open problems.

(3) Plato believed that the third world of forms or ideas would provide us with ultimate explanations (that is, explanation by essences; see my [1963] chapter 3). Thus he writes for example: 'I think that if anything else apart from the idea of absolute beauty is beautiful, then it is beautiful *for the sole reason* that it has some share in the idea of absolute beauty. *And this kind of explanation applies to everything.*' (PLATO, *Phaedo*, 100 C.)

This is a theory of *ultimate explanation*; that is to say, of an explanation whose explicans is neither capable nor in need of further explanation. And it is a theory of *explanation by essences*, that is, by hypostasized words.

As a result, Plato envisaged the objects of the third world as something like non-material things or, perhaps, like stars or constellations – to be gazed at, and intuited, though not liable to be touched by our minds. This is why the inmates of the third world – the forms or ideas – became concepts of things, or essences or natures of things, rather than theories or arguments or problems.

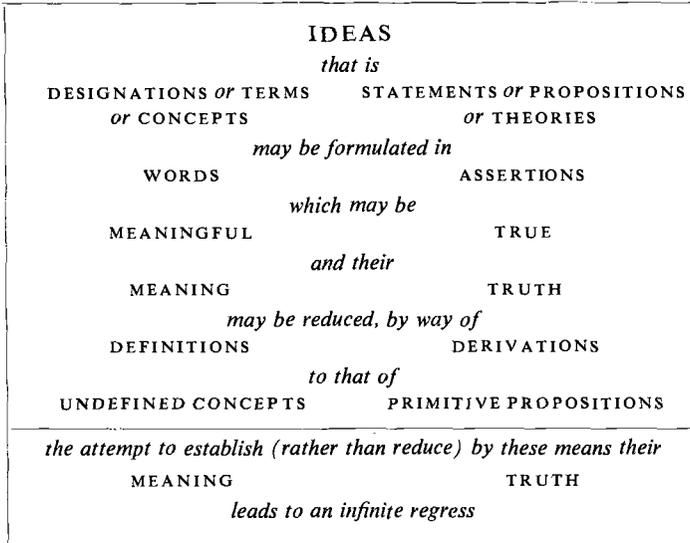
This had the most far-reaching consequences for the history of philosophy. From Plato until today, most philosophers have either been nominalists* or else what I have called essentialists. They are more interested in the (essential) meaning of words than in the truth and falsity of theories.

I often present the problem in the form of a table (see the next page).

My thesis is that *the left side of this table is unimportant*, as compared to the right side: what should interest us are theories; truth; argument. If so many philosophers and scientists still think that concepts and conceptual systems (and problems of their meaning, or the meaning of words) are comparable in importance to theories and theoretical systems (and problems of their truth, or the truth of statements), then they are still suffering from

* Cf. WATKINS [1965] Ch. VIII, esp. pp. 145f., and POPPER [1959] pp. 420–2; [1963] pp. 18ff., 262, 297f.

Plato's main error*. For concepts are partly means of formulating theories, partly means of summing up theories. In any case their significance is mainly instrumental; and they may always be replaced by other concepts.



Contents and objects of thought seem to have played an important part in Stoicism and in Neo-Platonism: Plotinus preserved Plato's separation between the empirical world and Plato's world of Forms or Ideas. Yet like Aristotle**, Plotinus destroyed the transcendence of Plato's world by placing it into the consciousness of God.

* The error, which is traditional, is known as 'the problem of universals'. This should be replaced by 'the problem of theories', or 'the problem of the theoretical content of all human language'. See POPPER [1959] sections 4 (with the new footnote *1) and 25.

Incidentally, it is clear that of the famous three positions – *universale ante rem*, *in re*, and *post rem* – the last, in its usual meaning, is anti-third-world and tries to explain language as expression, while the first (Platonic) is pro-third-world. Interestingly enough, the (Aristotelian) middle position (*in re*) may be said either to be anti-third-world or to ignore the problem of the third world. It thus testifies to the confusing influence of conceptualism.

** Cp. ARISTOTLE, *Metaphysics* XII (A), 7: 1072b21f.; and 9: 1074b15 to 1075a4. This passage (which Ross sums up: 'the divine thought must be concerned with the most divine object, which is itself') contains an implicit criticism of Plato. Its affinity with Platonic ideas is especially clear in lines 25 f.: 'it thinks of that which is most divine and precious, and it does not change; for change would be change for the worse...'. (See also ARISTOTLE, *De Anima* 429b27ff., esp. 430a4.)

Plotinus criticized Aristotle for failing to distinguish between the First Hypostasis (Oneness) and the Second Hypostasis (the divine intellect). Yet he followed Aristotle in identifying God's acts of thought with their own contents or objects; and he elaborated this view by taking the Forms or Ideas of Plato's intelligible world to be the immanent states of consciousness of the divine intellect.*

5.2. *Hegel*

Hegel was a Platonist (or rather a Neo-Platonist) of sorts and, like Plato, a Heraclitean of sorts. He was a Platonist whose world of Ideas was changing, evolving. Plato's 'Forms' or 'Ideas' were objective, and had nothing to do with conscious ideas in a subjective mind; they inhabited a divine, an unchanging, heavenly world (super-lunar in Aristotle's sense). By contrast Hegel's Ideas, like those of Plotinus, were conscious phenomena: thoughts thinking themselves and inhabiting some kind of consciousness, some kind of mind or 'Spirit'; and together with this 'Spirit' they were changing or evolving. The fact that Hegel's 'Objective Spirit' and 'Absolute Spirit' are subject to change is the only point in which his Spirits are more similar to my 'third world' than Plato's world of Ideas (or Bolzano's world of 'statements in themselves').

The most important differences between Hegel's 'Objective Spirit' and 'Absolute Spirit' and my 'third world' are these:

(1) According to Hegel, though the Objective Spirit (comprising artistic creation) and Absolute Spirit (comprising philosophy) both consist of human productions, man is not creative. It is the hypostasized Objective Spirit, it is the divine self-consciousness of the Universe, that moves man: 'individuals ... are instruments', instruments of the Spirit of the Epoch, and their work, their 'substantial business', is 'prepared and appointed independently of them'. (Cp. HEGEL [1830] paragraph 551.) Thus what I have called the autonomy of the third world, and its feed-back effect, becomes with Hegel omnipotent: it is only one of the aspects of his system in which his theological background manifests itself. As against this I assert that the individual creative element, the relation of give-and-take between a man and his work, is of the greatest importance. In Hegel this degenerates into the doctrine that the great man is something like a medium in which the Spirit of the Epoch expresses itself.

(2) In spite of a certain superficial similarity between Hegel's dialectic and

* Cp. PLOTINUS, *Enneades* II, 4, 4 ([1883] p. 153, 3); III, 8, 11 ([1883] p. 346, 6); V, 3, 2-5; V, 9, 5-8; VI, 5, 2; VI, 6, 6-7.

my evolutionary schema

$$P_1 \rightarrow TT \rightarrow EE \rightarrow P_2$$

there is a fundamental difference. My schema works through error elimination, and on the scientific level through conscious criticism under the regulative idea of the search for truth.

Criticism, of course, consists in the search for contradictions and in their elimination: the difficulty created by the demand for their elimination constitutes the new problem (P_2). Thus the elimination of error leads to the objective growth of our knowledge – of knowledge in the objective sense. It leads to the growth of objective verisimilitude: it makes possible the approximation to (absolute) truth.

Hegel, on the other hand, is a relativist*. He does not see our task as the search for contradictions, with the aim of eliminating them, for he thinks that contradictions are as good as (or better than) non-contradictory theoretical systems: they provide the mechanism by which the Spirit propels itself. Thus rational criticism plays no part in the Hegelian automatism, no more than does human creativity**.

(3) While Plato lets his hypostasized Ideas inhabit some divine heaven, Hegel personalizes his Spirit into some divine consciousness: the Ideas inhabit it as human ideas inhabit some human consciousness. His doctrine is throughout that the Spirit is not only conscious but a self. As against this, my third world has no similarity whatever to human consciousness; and though its first inmates are products of human consciousness, they are totally different from conscious ideas or from thoughts in the subjective sense.

5.3. Bolzano and Frege

Bolzano's statements in themselves and truths in themselves are, clearly, inhabitants of my third world. But he was far from clear about their relationship to the rest of the world***.

It is, in a way, Bolzano's central difficulty which I have tried to solve by comparing the status and autonomy of the third world to those of animal products, and by pointing out how it originates in the higher functions of the human language.

* See POPPER [1963] chapter 15; POPPER [1962] Addendum to vol. ii: 'Facts, Standards and Truth: A Further Criticism of Relativism'.

** See LAKATOS [1963] p. 234, footnote 1 (Offprint p. 59).

*** BOLZANO [1837] Vol. I, §19, p. 78, says that statements (and truths) in themselves have no being ('*Dasein*'), existence, or reality. Yet he also says that a statement in itself is *not merely* 'something stated, thus presupposing a person who stated it'.

As to Frege, there can be no doubt about his clear distinction between the subjective acts of thinking, or thought in the subjective sense, and objective thought or thought content*.

Admittedly, his interest in subordinate clauses of a sentence, and in indirect speech, made him the father of modern epistemic logic**. But I think that he is in no way affected by the criticism of epistemic logic which I am going to offer (see section 7 below): as far as I can see, he was not thinking in these contexts of epistemology in the sense of a theory of scientific knowledge.

5.4. *Empiricism*

Empiricism – say, of Locke, Berkeley and Hume – has to be understood in its historical setting: its main problem was, simply, religion *versus* irreligion; or more precisely, the rational justification, or justifiability, of Christianity, as compared to scientific knowledge.

This explains why knowledge is throughout regarded as a kind of belief – belief justified by evidence, especially by perceptual evidence, by the evidence of our senses.

Though their positions with respect to the relation of science and religion differ widely, Locke, Berkeley*** and Hume agree essentially in the demand (which Hume sometimes feels is an unattainable ideal) that we should reject all propositions – and especially propositions with existential import – for which the evidence is insufficient, and accept only those propositions for which we have sufficient evidence: which can be proved, or verified, by the evidence of our senses.

This position can be analysed in various ways. A somewhat sweeping analysis would be the following chain of equations or equivalences most of which can be supported by passages from the British empiricists and even from Bertrand Russell.****

* See the quotation in section 1 above from FREGE [1892] p. 32, and FREGE [1894].

** The way leads from FREGE to RUSSELL [1922] p. 19 and WITTGENSTEIN [1922] 5,542.

*** For Berkeley's position compare section 1 of ch. 3 and ch. 6 of POPPER [1963].

**** Cp. RUSSELL [1906–7] p. 45: 'Truth is a quality of beliefs'; RUSSELL [1910]: 'I shall use the words "belief" and "judgment" as synonyms.' (p. 172, footnote); or: '... judgment is... a multiple relation of the mind to the various other terms with which the judgment is concerned.' (p. 180). He also holds that 'perception is always true (even in dreams and hallucinations)' (p. 181). Or cp. RUSSELL [1959] p. 183: '... but from the point of view of the theory of knowledge and of the definition of truth it is sentences expressing belief that are important'. See also RUSSELL [1922] pp. 19f., and Ducasse's '*epistemic attitudes*' in DUCASSE [1940], pp. 701–711. It is clear that both Russell and Ducasse belong to those traditional epistemologists who study knowledge in its subjective or second-world sense. The tradition far transcends empiricism.

p is verified or demonstrated by sense experience = there is sufficient reason or justification for us to believe p = we believe or judge or assert or assent or know that p is true = p is true = p .

One remarkable thing about this position which *conflates the evidence, or proof, and the assertion to be proved*, is that anybody who holds it ought to reject the law of the excluded middle. For it is obvious that the situation may arise (in fact, it would be practically the normal situation) that neither p nor not- p can be fully supported, or demonstrated, by the evidence available. Yet it seems that this was not noticed by anybody before Brouwer.

This failure to reject the law of the excluded middle is particularly striking in Berkeley; for if

$$esse = percipi$$

then the truth of any statement about reality can be established only by perception statements. Yet Berkeley, very much like Descartes, suggests in his Dialogues* that we should reject p if there is 'no reason to believe in it'. The absence of such reasons may be compatible, however, with the absence of reasons to believe in non- p .

6. Appreciation and criticism of Brouwer's epistemology

In the present section I wish to pay homage to L. E. J. Brouwer.

It would be presumptuous of me to try to praise and even more presumptuous to try to criticize Brouwer as a mathematician. But it may be permissible for me to try to criticize his epistemology and his philosophy of intuitionist mathematics. If I dare to do so, it is in the hope of making a contribution, however slight, to the clarification and further development of Brouwer's ideas.

In his Inaugural Lecture BROUWER [1912] starts from Kant. He says that Kant's intuitionist philosophy of geometry – his doctrine of the pure intuition of space – has to be abandoned in the light of non-Euclidean geometry. But, Brouwer says, we do not need it, since we can arithmetize geometry: we can take our stand squarely on Kant's theory of arithmetic, and on his doctrine that arithmetic is based upon the pure intuition of time.

* See the second dialogue between Hylas and Philonous (BERKELEY [1949] p. 218, lines 15f.): 'It is to me a sufficient reason not to believe the existence of any thing, if I see no reason for believing it.' Compare DESCARTES [1637] Pt. IV (first paragraph): 'Any opinion should be rejected as manifestly false [*aperte falsa* in the Latin version] if the slightest reason for doubt can be found in it.'

I feel that this position of Brouwer's can no longer be sustained; for if we say that Kant's theory of space is destroyed by non-Euclidean geometry, then we are bound to say that his theory of time is destroyed by special relativity. For Kant says explicitly that there is only *one* time, and that the intuitive idea of (absolute) simultaneity is decisive for it*.

It might be argued – on lines somewhat parallel to a remark of Heyting's** – that Brouwer might not have developed his epistemological and philosophical ideas about intuitionist mathematics had he known at the time of the analogy between Einstein's relativization of time and non-Euclidean geometry. To paraphrase Heyting, this would have been unfortunate.

However, it is unlikely that Brouwer would have been overmuch impressed by special relativity. He might have given up citing Kant as a precursor of his intuitionism. But he could have retained his own theory of a *personal* time – of a time of our own intimate and immediate experience. (See BROUWER [1949]). And this was in no way affected by relativity, even though Kant's theory was affected.

Thus we need not treat Brouwer as a Kantian. Yet we cannot sever him from Kant too easily. For Brouwer's idea of intuition, and his use of the term 'intuition', cannot be fully understood without analysing its Kantian background.

For Kant, *intuition is a source of knowledge*; and 'pure' intuition ('the pure intuition of space and time') is an unfailing source of knowledge: from it springs *absolute certainty*. This is most important for the understanding of Brouwer who clearly adopts this epistemological doctrine from Kant.

It is a doctrine with a history. Kant took it from Plotinus, St. Thomas, Descartes, and others. Originally, intuition meant, of course, perception: it is what we see, or perceive, if we look at, or if we direct our gaze on to, some object. But at least from Plotinus on, there developed a contrast between *intuition* on the one hand, and *discursive* thinking on the other. Intuition is God's way of knowing everything at a glance, in a flash, timelessly. Discursive thought is the human way: as in a discourse, we argue step by step, which takes time.

Now Kant upheld the doctrine (against Descartes) that we do not possess

* In the *Transcendental Aesthetic* (KANT [1778] pp. 46f; Kemp-Smith's translation, pp. 74f.), Kant stresses under point 1) the *a priori* character of simultaneity; under points 3) and 4) that there can be only *one* time; and under point 4) that time is *not a discursive concept*, but 'a pure form of ... intuition' (or more precisely, *the pure form of sensual intuition*). In the last paragraph before the Conclusion on p. 72 (Kemp-Smith, p. 90) he says explicitly that the intuition of space and time is not an intellectual intuition.

** See the quotation from Heyting in section 1 above.

a faculty of intellectual intuition, and that, for this reason, our intellect – our concepts – remain empty or analytic, unless indeed they are applied to material which is either given to us by our senses (sense intuition), or unless they are ‘*concepts constructed in our pure intuition of space and time*’*. Only in this way can we obtain synthetic knowledge *a priori*: our intellect is essentially discursive; it is bound to proceed by logic, which is empty – ‘analytic’.

According to Kant, sense intuition presupposes pure intuition: our senses cannot do their work without ordering their perceptions into the framework of space and time. Thus space and time are prior to all sense-intuition; and the theories of space and time – geometry and arithmetic – are *a priori* valid. The source of their *a priori* validity is the human faculty of *pure intuition*, which is strictly limited to this field, and which is strictly distinct from the intellectual or discursive way of thinking.

Kant maintained the doctrine that the *axioms of mathematics* were based on pure intuition (KANT [1778] p. 760 f.): they could be ‘seen’ or ‘perceived’ to be true, in a non-sensual manner of ‘seeing’ or ‘perceiving’. In addition, pure intuition was involved in *every step of every proof in geometry* (and in mathematics generally)**: to follow a proof we need to look at a (drawn) figure. This ‘looking’ is not sense-intuition but pure intuition, as shown by the fact that the figure might often be convincing even though drawn in a very rough manner, and by the fact that the drawing of a triangle might represent for us, in *one* drawing, an infinity of possible variants – triangles of all shapes and sizes.

Analogous considerations hold for arithmetic which, according to Kant, is based on counting; a process which in its turn is essentially based on the pure intuition of time.

Now this theory of the sources of mathematical knowledge suffers in its Kantian form from a severe difficulty. Even if we admit everything that Kant says, we are left puzzled. For Euclid’s geometry, whether or not it uses pure

* See KANT [1778] p. 741: ‘To construct a concept means to exhibit this *a priori* intuition [the ‘pure intuition’] which corresponds to the concept.’ See also p. 747: ‘We have endeavoured to make it clear how great the difference is between the discursive use of reason through concepts and the intuitive use through the construction of concepts.’ On p. 751, the ‘*construction of concepts*’ is further explained: ‘we can determine our concepts in our *a priori* intuition of space and time in as much as we create the *objects themselves* by way of a uniform synthesis’. (The italics are partly mine.)

** Cp. KANT [1778] pp. 741–764. See, for example, the end of p. 762 where he says about proofs in mathematics (‘even in algebra’): ‘all inferences are made safe ... by placing them plainly before our eyes’. Cp., for example, also the top of p. 745 where Kant speaks of a ‘chain of inferences’, and ‘always guided by intuition’. (In the same passage (p. 748) ‘to construct’ is explained as ‘to represent in intuition’.)

intuition, certainly makes use of intellectual argument, of logical deduction. *It is impossible to deny that mathematics uses discursive thought.* Euclid's discourse moves through propositions and whole books step by step: it was not conceived in one single intuitive flash. Even if we admit, for the sake of the argument, the need for pure intuition *in every single step without exception* (and this admission is difficult for us moderns to make), the step-wise, discursive and logical procedure of Euclid's derivations is so unmistakable, and it was so generally known and imitated (Spinoza, Newton) that it is difficult to believe that Kant can have ignored it. In fact Kant knew all this probably as well as anybody. But his position was forced upon him, by (1) the structure of the *Critique* in which the 'Transcendental Aesthetic' precedes the 'Transcendental Logic', and (2) by his sharp distinction (I should suggest untenably sharp distinction) between intuitive and discursive thought. As it stands, one is almost inclined to say that there is not merely a lacuna here in Kant's exclusion of discursive arguments from geometry and arithmetic, but a contradiction.

That this is not so was shown by Brouwer who filled the lacuna. I am alluding to Brouwer's theory of *the relation between mathematics on the one hand and language and logic on the other.*

Brouwer solved the problem by making a sharp distinction between *mathematics as such* and *its linguistic expression and communication.* Mathematics itself he saw as an extra-linguistic activity, essentially an activity of mental construction on the basis of our pure intuition of time. By way of this construction we create in our intuition, in our mind, the objects of mathematics which afterwards – after their creation – we can try to describe, and to convey to others. Thus the linguistic description, and the discursive argument with its logic, comes after the essentially mathematical activity: it always comes after an object of mathematics – such as a proof – has been constructed.

This solves the problem which we uncovered in Kant's *Critique.* What at first sight appears to be a contradiction in Kant is removed, in a most ingenious way, by the doctrine that we must sharply distinguish between two levels, one level intuitive and mental and essential for mathematical thought, the other discursive and linguistic and essential for communication only.

Like every great theory, this theory of Brouwer's shows its worth by its fertility. It solved three great sets of problems in the philosophy of mathematics with one stroke:

- (1) *Epistemological problems* concerning the source of mathematical cer-

tainty; the nature of mathematical evidence; and the nature of mathematical proof. These problems were solved, respectively, by the doctrine of intuition as a source of knowledge; by the doctrine that we can intuitively see the mathematical objects we have constructed; and by the doctrine that a mathematical proof is a sequential construction, or a construction of constructions.

(2) *Ontological problems* concerning the nature of mathematical objects and the nature of their mode of existence. These problems were solved by a doctrine which had two sides: on the one side there was *constructivism*, and on the other there was a *mentalism* which located all mathematical objects in what I call the 'second world'. Mathematical objects were constructions of the human mind, and they existed solely as constructions in the human mind. Their objectivity – their character as objects, and the objectivity of their existence – rested entirely in the possibility of repeating their construction at will.

Thus Brouwer in his inaugural lecture could imply that, for the intuitionist, mathematical objects existed in the human mind; while for the formalist, they existed 'on paper'*

(3) *Methodological problems* concerning mathematical proofs.

We may quite naively distinguish two main ways of being interested in mathematics. One mathematician may be interested mainly in theorems – in the truth or falsity of mathematical propositions. Another mathematician may be interested mainly in proofs: in questions of the existence of proofs of some theorem or other, and in the character of the proofs. If the first interest is preponderant (which seems to be the case for example with Polya), then it is usually linked with an interest in the discovery of mathematical 'facts' and thus with a Platonizing mathematical heuristic. If the second kind of interest is preponderant, then proofs are not merely means of making sure of theorems about mathematical objects, but they are mathematical objects themselves. This, it seems to me, was the case with Brouwer: those constructions which were proofs were not only creating and establishing mathematical objects, they were at the same time themselves mathematical objects – perhaps even the most important ones. Thus to assert a theorem was to assert the existence of a proof for it, and to deny it was to assert the existence

* Cp. the end of the third paragraph of BROUWER [1912]. Brouwer speaks there about the existence not of mathematics but of 'mathematical exactness', and *as it stands*, the passage therefore applies to the problems (1) and (3) even more closely than to the ontological problem (2). But there can be no doubt that it was meant to apply to (2) also. The passage reads in Dresden's translation: 'The question where mathematical exactness does exist is answered differently The intuitionist says: in the human intellect. The formalist says: on paper.'

of a refutation; that is, a proof of its absurdity. This leads immediately to Brouwer's rejection of the law of the excluded middle, to his rejection of indirect proofs, and to the demand that existence can be proved only by the actual construction – the making visible as it were – of the mathematical object in question.

It also leads to Brouwer's rejection of 'Platonism' by which we may understand the doctrine that mathematical objects have what I call an 'autonomous' mode of existence: that they may exist without having been constructed by us, and thus without having been proved to exist.

So far I have tried to understand Brouwer's epistemology, mainly by conjecturing that it springs from an attempt to solve a difficulty in Kant's philosophy of mathematics. I now proceed to what I announced in the title of this section – to an appreciation and criticism of Brouwer's epistemology.

From the point of view of the present paper, it is one of Brouwer's great achievements that he saw that mathematics – and perhaps I may add, the third world – is created by man.

This idea is so radically anti-Platonic that it is understandable that Brouwer did not see that it can be combined with a kind of Platonism. I mean the doctrine of the (partial) *autonomy* of mathematics, and of the third world, as sketched in section 3 above.

Brouwer's other great achievement, from a philosophical point of view, was his anti-formalism: his recognition that mathematical objects must exist before we can talk about them.

But let me turn to a criticism of Brouwer's solution of the three main sets of problems of the philosophy of mathematics discussed earlier in the present section.

(1') *Epistemological problems*: Intuition in general, and the theory of time in particular.

I do not propose to change the name 'Intuitionism'. Since the name will no doubt be retained, it is the more important to give up the mistaken philosophy of intuition as an infallible source of knowledge.

There are no authoritative sources of knowledge, and no 'source' is particularly reliable*. Everything is welcome as a source of inspiration, including 'intuition'; especially if it suggests new problems to us. But nothing is secure, and we are all fallible.

* I have dealt with this problem at length in my lecture 'On the sources of knowledge and of ignorance' which now forms the Introduction to POPPER [1963].

Besides, Kant's sharp distinction between intuition and discursive thought cannot be upheld. 'Intuition', whatever it may be, is largely the product of our cultural development, and of our efforts in discursive thinking. Kant's idea of one standard type of pure intuition shared by us all (perhaps not by animals in spite of a similar perceptual outfit) can hardly be accepted. For after having trained ourselves in discursive thought, our intuitive grasp becomes utterly different from what it was before.

All this applies to our intuition of time. I personally find Benjamin Lee Whorf's report on the Hopi Indians* and their utterly different intuition of time convincing. But even if this report should be incorrect (which I think unlikely), it shows possibilities which neither Kant nor Brouwer ever considered. Should Whorf be right, then our intuitive grasp of time – the way in which we 'see' temporal relations – would partly depend on our language and the theories and myths incorporated in it: *our own European intuition of time would owe much to the Greek origins of our civilization, with its emphasis on discursive thought.*

At any rate, our intuition of time may change with our changing theories. The intuitions of Newton, Kant and Laplace differ from Einstein's; and the role of time in particle physics differs from that in the physics of continua, especially optics. While particle physics suggests a razor-like unextended instant, a '*punctum temporis*' which divides the past from the future, and thus a time coordinate consisting of (a continuum of) unextended instants, and a world whose 'state' may be given for any such unextended instant, the situation in optics is very different. Just as there are spatially extended grids in optics whose parts co-operate over a considerable distance of space, so there are temporally extended events (waves possessing frequencies) whose parts co-operate over a considerable distance of time. Thus owing to optics, *there cannot be in physics a state of the world at an instant of time.* This argument should, and does, make a great difference to our intuition: what has been called the specious present of psychology is neither specious nor confined to psychology, but is genuine and occurs already in physics**.

Thus not only is the general doctrine of intuition as an infallible source of knowledge a myth, but our intuition of time, more especially, is just as subject to criticism and correction as is, according to Brouwer's own admission, our intuition of space.

* Cp. 'An American Indian model of the universe' in WHORF [1956].

** Cp. GOMBRICH [1964] especially p. 297: 'If we want to pursue this thought to its logical conclusion the *punctum temporis* could not even show as a meaningless dot, for light has a frequency.' (The argument can be supported by considering boundary conditions.)

The main point here I owe to Lakatos's philosophy of mathematics. It is that mathematics grows through criticism of guesses and bold informal proofs. This presupposes their linguistic formulation, and their status in the third world. Language, at first merely a means of communicating descriptions of prelinguistic objects, thus becomes an *essential part* of the scientific enterprise, even in mathematics, which in its turn becomes part of the third world. And there are layers, or levels, in language (whether or not they are formalized in a hierarchy of metalanguages).

Were the intuitionist epistemology correct, mathematical competence would be no problem. (Were Kant's theory correct, it would not be understandable why we – or more precisely Plato and his school – had to wait so long for Euclid*.) Yet it is a problem, since even highly competent intuitionist mathematicians can disagree on some difficult points**. It is not necessary for us to enquire which side in the disagreement is in the right. It is sufficient to point out that, once an intuitionist construction can be criticized, the problem raised can only be solved *by using argumentative language in an essential way*. Of course, the essential critical use of language does not commit us to the use of arguments banned by intuitionist mathematics (though there is a problem here, as will be shown). My point at the moment is merely this: once the admissibility of a proposed intuitionist mathematical construction can be questioned – and of course it can be questioned – language becomes more than a mere means of communication which could in principle be dispensed with: it becomes, rather, the indispensable medium of critical discussion. Accordingly it is no longer only the intuitionist construction 'which is objective in the sense that it is irrelevant which subject makes the construction'***; rather, the objectivity, even of intuitionist mathematics, rests, as does that of all science, upon the criticizability of its arguments. But this means that language becomes indispensable as the medium of argument, of critical discussion.****

It is for this reason that I regard Brouwer's subjectivist epistemology, and the philosophical justification of his intuitionist mathematics, as mistaken. There is a give and take between construction, criticism, 'intuition', and even tradition, which he fails to consider.

* Cp. the corresponding remark on Kant's aprioristic view of Newton's physics in POPPER [1963] chapter 2, the paragraph to which the footnote 63 is attached.

** Cp. S. C. Kleene's comments in KLEENE and VESLEY [1965] pp. 176–83, on BROUWER [1951] pp. 357–8, which Kleene criticizes in the light of Brouwer's note on page 1248 of BROUWER [1949].

*** Heyting in LAKATOS [1967] p. 173.

**** Cp. LAKATOS [1963–4], especially pp. 229–35.

I am, however, prepared to admit that even in his erroneous view of the status of language Brouwer was partly right. Although the objectivity of all science, including mathematics, is inseparably linked with its criticizability, and therefore with its linguistic formulation, Brouwer was right in reacting strongly against the thesis that mathematics is *nothing but* a formal language game or, in other words, that there are no such things as extra-linguistic mathematical objects; that is to say, thoughts (or in my view, more precisely, thought contents). As he insisted, mathematical talk is *about* these objects; and in this sense, mathematical language is secondary to these objects. But this does not mean that we could construct mathematics without language: there can be no construction without constant critical control, and no critical control without giving our constructs linguistic form and treating them as objects of the third world. Although the third world is not identical with the world of linguistic forms, it arises together with argumentative language: it is a by-product of language. This explains why, once our constructions become problematic, systematized, and axiomatized, language may become problematic too, and why formalization may become a branch of mathematical construction. This, I think, is what Professor Myhill means when he says that '*our formalizations correct our intuitions while our intuitions shape our formalizations*'*. What makes this remark particularly worth quoting is that, having been made in connection with Brouwerian intuitionist proof, it seems indeed to provide a correction of Brouwerian epistemology.

(2') *Ontological problems*: That the objects of mathematics owe their existence partly to language was sometimes seen by Brouwer himself. Thus he wrote in 1924: 'Mathematics is based upon ["*Der Mathematik liegt zugrunde*"] an unlimited sequence of signs or symbols ["*Zeichen*"] or of finite sequences of symbols ...'**. This need not be read as an admission of the priority of language: no doubt the crucial term is 'sequence', and the idea of a sequence is based upon the intuition of time, and upon construction based upon this intuition. Yet it shows that Brouwer was aware that signs or symbols were needed to carry out the construction. My own view is that discursive thought (that is, sequences of linguistic arguments) has the strongest influence upon our awareness of time, and upon the development of our intuition of sequential order. This in no way clashes with Brouwer's constructivism; but it does clash with his subjectivism and mentalism. For the objects of mathematics can now become citizens of an objective third world:

* J. MYHILL [1967] p. 175 (my italics). Also cp. LAKATOS [963-4].

** BROUWER [1924] p. 244.

though originally constructed by us – the third world originates as our product – the thought contents carry with them their own unintended consequences. The series of natural numbers which we construct creates prime numbers – which we *discover* – and these in turn create problems of which we never dreamt. *This is how mathematical discovery becomes possible.* Moreover the most important mathematical objects we discover – the most fertile citizens of the third world – are *problems*, and new kinds of *critical arguments*. Thus a new kind of mathematical existence emerges: the existence of problems; and a new kind of intuition: the intuition which makes us see problems, and which makes us understand problems prior to solving them. (Think of Brouwer's own central problem of the continuum.)

The way in which language and discursive thought interact with more immediate intuitive constructions (an interaction which, incidentally, destroys that ideal of absolute evidential certainty which intuitive construction was supposed to realize) has been described in a most enlightening way by Heyting. I may perhaps quote the beginning of a passage of his from which I have derived not only stimulation but also encouragement: 'It has proved not to be intuitively clear what is intuitively clear in mathematics. It is even possible to construct a descending scale of grades of evidence. The highest grade is that of such assertions as $2+2=4$. $1002+2=1004$ belongs to a lower grade; we show this not by actual counting, but by reasoning which shows that in general $(n+2)+2=n+4$ [Statements like this] have already the character of an implication: "If a natural number n is constructed, then we can effect the construction, expressed by $(n+2)+2=n+4$ ".'* In our present context, Heyting's 'grades of evidence' are of secondary interest. What is primarily important is his beautifully simple and clear analysis of the unavoidable interplay between intuitive construction and linguistic formulation which necessarily involves us in discursive – and therefore logical – reasoning. The point is stressed by Heyting when he continues: 'This level is formalized in the free-variable calculus.'

A last word may be said on Brouwer and mathematical Platonism. The autonomy of the third world is undeniable, and with it, Brouwer's equation '*esse = construi*' must be given up; *at least* for problems. This may lead us to look anew at the problem of the logic of intuitionism: *without giving up the intuitionist standards of proof*, it may be important for critical rational discussion to distinguish sharply between a thesis and the evidence for it. But this distinction is destroyed by intuitionist logic which

* Cp. HEYTING [1962] p. 195.

results from the *conflation of evidence, or proof, and the assertion to be proved* *.

(3') *Methodological problems*: The original motive of Brouwer's intuitionist mathematics was security: the search for safer methods of proof; in fact, for infallible methods. Now if you want more secure proofs, you must be more severe concerning the admissibility of demonstrative argument: you must use weaker means, weaker assumptions. Brouwer confined himself to the use of logical means which were weaker than those of classical logic **. To prove a theorem by weaker means is (and has always been) an intensely interesting task, and one of the great sources of mathematical problems. Hence the interest of intuitionist methodology.

But I suggest that this holds for proofs only. For criticism, for refutation, we do not want a poor logic. While an organon of demonstration should be kept weak, an organon of criticism should be strong. In criticism we do not wish to be confined to demonstrate impossibilities: we do not claim infallibility for our criticism, and we are often content if we can show that some theory has counter-intuitive consequences. In an organon of criticism, weakness and parsimony is no virtue, since it is a virtue in a theory that it can stand up to strong criticism. (It seems therefore plausible that in the critical debate – the metadepbate – of the validity of an intuitionist construction, the use of full classical logic may be admissible.)

7. Subjectivism in logic, probability theory and science

In view of what has been said in section 5, especially on empiricism, it is not surprising that neglect of the third world – and consequently a subjectivist epistemology – should be still widespread in contemporary thought. Even where there is no connection with Brouwerian mathematics there are often subjectivist tendencies to be found within the various specialisms. I will here refer to some such tendencies in logic, probability theory, and physical science.

7.1. *Epistemic logic*

Epistemic logic deals with such formulae as '*a* knows *p*' or '*a* knows that *p*' and '*a* believes *p*', or '*a* believes that *p*'. It usually symbolizes these by

'*Kap*' or '*Bap*'

* Cp. section 5.4 above.

** These remarks hold only for the *logic* of intuitionism which is part of classical logic, while intuitionist mathematics is not part of classical mathematics. See especially Kleene's remarks on 'Brouwer's principle' in KLEENE and VESLEY [1965] p. 70.

where 'K' or 'B' respectively stand for the relationships of knowing or believing, and *a* is the knowing or believing subject and *p* the known or believed proposition or state of affairs.

My first thesis in section 1 implies that this has nothing to do with scientific knowledge: that the scientist, I will call him 'S', does neither know nor believe. What does he do? I will give a very brief list:

- 'S tries to understand *p*.'
- 'S tries to think of alternatives to *p*.'
- 'S tries to think of criticisms of *p*.'
- 'S proposes an experimental test for *p*.'
- 'S tries to axiomatize *p*.'
- 'S tries to derive *p* from *q*.'
- 'S tries to show that *p* is not derivable from *q*.'
- 'S proposes a new problem *x* arising out of *p*.'
- 'S proposes a new solution of the problem *x* arising out of *p*.'
- 'S criticizes his latest solution of the problem *x*.'

The list could be extended at some length. It is miles removed in character from 'S knows *p*' or 'S believes *p*' or even from 'S mistakenly believes *p*' or 'S doubts *p*'. In fact, it is quite an important point that we may doubt without criticizing, and criticize without doubting. (That we may do so was seen by Poincaré in *Science and hypothesis*, which may be in this point contrasted with Russell's *Our knowledge of the external world*.)

7.2. Probability theory

Nowhere has the subjectivist epistemology a stronger hold than in the field of the calculus of probability. This calculus is a generalization of Boolean algebra (and thus of the logic of propositions). It is still widely interpreted in a subjective sense, as a *calculus of ignorance, or of uncertain subjective knowledge*; but this amounts to interpreting Boolean algebra, including the calculus of propositions, as a *calculus of certain knowledge* – of certain knowledge *in the subjective sense*. This is a consequence which few Bayesians (as the adherents of the subjective interpretation of the probability calculus now call themselves) will cherish.

This subjective interpretation of the probability calculus I have combatted for 33 years. Fundamentally, it springs from the same epistemic philosophy which attributes to the statement 'I know that snow is white' a greater epistemic dignity than to the statement 'snow is white'.

I do not see any reason why we should not attribute still greater epistemic dignity to the statement 'In the light of all the evidence available to me I

believe that it is rational to believe that snow is white'. The same could be done, of course, with probability statements.

7.3. *Physical science*

The subjective approach has made much headway in science since about 1926. First it took over quantum mechanics. Here it became so powerful that its opponents were regarded as nitwits who should rightfully be silenced. Then it took over statistical mechanics. Here Szilard proposed in 1929 the by now almost universally accepted view that we have to pay for subjective information by physical entropy increase; which was interpreted as a proof that physical entropy is lack of knowledge and thus a subjective concept, and that knowledge or information is equivalent to physical negentropy. This development was neatly matched by a parallel development in information theory which started as a perfectly objective theory of channels of communication but was later linked with Szilard's subjectivist information concept.

Thus the subjective theory of knowledge has entered science on a broad front. The original point of entry was the subjective theory of probability. But the evil has spread into statistical mechanics, the theory of entropy, into quantum mechanics, and into information theory.

It is of course not possible to refute in this lecture all these subjectivist theories. I cannot do more than mention that I have combatted them for years (most recently in my [1967]). But I do not harbour any illusions. There will be many more years before the tide will turn – if it ever does.

There are only two final points I wish to make.

First, I shall try to indicate what epistemology or the logic of discovery looks like from an objectivist point of view, and how it may be able to throw some light on the biology of discovery.

Secondly, I shall try to indicate, in the last section of this lecture, what the *psychology of discovery looks like, from the same objectivist point of view.*

8. The logic and the biology of discovery

Epistemology becomes, from an objectivist point of view, the theory of the growth of knowledge. It becomes the theory of problem-solving, or, in other words, of the construction, critical discussion, evaluation, and critical testing, of competing conjectural theories.

I now think that with respect to competing theories it is perhaps better

to speak of their 'evaluation' or 'appraisal', or of the 'preference' for one of them, rather than of its 'acceptance'. Not that words matter. The use of 'acceptance' causes no harm as long as it is kept in mind that all acceptance is tentative and, like belief, of passing and personal rather than objective and impersonal significance*.

The evaluation or appraisal of competing theories is partly prior to testing (*a priori*, if you like, though not in the Kantian sense of the terms which means '*a priori* valid') and partly posterior to testing (*a posteriori*, again in a sense which does not imply validity). Also prior to testing is the (empirical) content of a theory, which is closely related to its (virtual) explanatory power, that is to say, its power to solve pre-existing problems – those problems which give rise to the theory, and with respect to which the theories are *competing theories*.

Only with respect to some pre-existing set of problems can theories be (*a priori*) evaluated, and their values compared. Their so-called simplicity too can be compared only with respect to the problems in whose solution they compete.

Content and virtual explanatory power are the most important regulative ideas for the *a priori* appraisal of theories. They are closely related to their degree of testability.

The most important idea for their *a posteriori* appraisal is truth or, since we need a more accessible comparative concept, what I have termed 'nearness to truth', or 'verisimilitude'**. It is important that while a theory without content can be true (such as a tautology), verisimilitude is based upon the regulative idea of truth content; that is to say, on the idea of the amount of interesting and important true consequences of a theory. Thus a tautology, though true, has zero *truth content* and zero verisimilitude. It has of course the probability *one*. Generally speaking, content and testability and verisimilitude*** can be measured by *improbability*.

The *a posteriori* evaluation of a theory depends entirely upon the way it has stood up to severe and ingenious tests. But severe tests, in their turn, presuppose a high degree of *a priori* testability or content. Thus the *a posteriori* evaluation of a theory depends largely upon its *a priori* value: theories which are *a priori* uninteresting – of little content – need not be tested be-

* For instance, I have no objection whatever to Lakatos's use of the terms 'acceptance₁' and 'acceptance₂' in his 'Changes in the problem of inductive logic', § 3 (LAKATOS [1968]).

** Cp. POPPER [1963] especially chapter 10, section 3, and addendum 6; also POPPER [1962a] especially p. 292.

*** Cp. POPPER, 'A theorem on truth content', in FEYERABEND and MAXWELL [1966].

cause their low degree of testability excludes *a priori* the possibility that they may be subjected to really significant and interesting tests.

On the other hand, highly testable theories are interesting and important even if they fail to pass their test; we can learn immensely from their failure. Their failure may be fruitful, for it may actually suggest how to construct a better theory.

Yet all this stress upon the fundamental importance of *a priori* evaluation could perhaps be interpreted as ultimately due to our interest in high *a posteriori* values – in obtaining theories which have a high truth content and verisimilitude, though they remain of course always conjectural or hypothetical or tentative. What we are aiming at are theories which are not only intellectually interesting and highly testable, but which have actually passed severe tests better than their competitors; which thus solve their problems better; and which, should their conjectural character become manifest by their refutation, give rise to new, unexpected, and fruitful, problems.

Thus we can say that science begins with problems and proceeds from there to competing theories which it evaluates *critically*. Especially significant is the evaluation of their verisimilitude. This demands severe critical tests, and therefore presupposes high degrees of testability, which are dependent upon the content of the theory, and therefore can be evaluated *a priori*.

In most cases, and in the most interesting cases, the theory will ultimately break down and thus raise new problems. And the advance achieved can be assessed by the intellectual gap between the original problem and the new problem which results from the breakdown of the theory.

This cycle can again be described by our repeatedly used diagram:

$$P_1 \rightarrow TT \rightarrow EE \rightarrow P_2;$$

that is: problem P_1 – tentative theory – evaluative elimination – problem P_2 .

The evaluation is always *critical*, and its aim is the discovery and *elimination of error*. The growth of knowledge – and thus the learning process – is not a repetitive or a cumulative process but one of error elimination. It is Darwinian selection, rather than Lamarckian instruction.

This is a brief description of epistemology from an objective point of view: the method, or logic, of aiming at the growth of objective knowledge. But although it describes the growth of the third world, it can be interpreted as a description of biological evolution. Animals, and even plants, are problem-solvers. And they solve their problems by the method of competitive tentative solution and error elimination.

The tentative solutions which animals and plants incorporate into their anatomy and their behaviour are biological analogues of theories; and *vice versa*: theories correspond, as do many exosomatic products such as honeycombs, and especially exosomatic tools, such as spider webs, to endosomatic organs and their ways of functioning. Just like theories, organs and their functions are tentative adaptations to the world we live in. And just like theories, or like tools, new organs and their functions, and also new kinds of behaviour, exert their influence on the first world which they may help to change. (A new tentative solution – a theory, an organ, a new kind of behaviour – may discover a new virtual ecological niche and thus may turn a virtual niche into an actual one.) Behaviour or organs may also lead to the emergence of new problems. And in this way they may influence the further course of evolution, including the emergence of new biological values.

All this holds, especially, for sense organs. They incorporate, more especially, theory-like expectations. Sense organs, such as the eye, are prepared to react to certain selected environmental events – to those events which they ‘expect’, and *only* to those events. Like theories (and prejudices) they will in general be blind to others: to those which they do not understand, which they cannot interpret (because they do not correspond to any specific problem which the organism is trying to solve).*

Classical epistemology which takes our sense perceptions as ‘given’, as the ‘data’ from which our theories have to be constructed by some process of induction, can only be described as pre-Darwinian. It fails to take account of the fact that the alleged data are in fact adaptive reactions, and therefore interpretations which incorporate theories and prejudices and which, like theories, are impregnated with conjectural expectations; that there can be no pure perception, no pure datum; exactly as there can be no pure observational language, since all languages are impregnated with theories and myths. Just as our eyes are blind to the unforeseen or unexpected, so our languages are unable to describe it (though our languages can grow – as can our sense organs, endosomatically as well as exosomatically).

This consideration of the fact that theories or expectations are built into our very sense organs shows that the epistemology of induction breaks down even before having taken its first step. It cannot start from sense data or perceptions and build our theories upon them, since there are no such things as sense data or perceptions which are not built upon theories (or expectations, that is, the biological predecessors of linguistically formulated theo-

* Cp. my remarks in LAKATOS and MUSGRAVE [1968] p. 163.

ries). Thus the 'data' are no basis of, no guarantee for, the theories: they are not more secure than any of our theories or 'prejudices' but, if anything, less so (assuming for argument's sake that sense data exist and are not philosophers' inventions). Sense organs incorporate the equivalent of primitive and uncritically accepted theories, which are less widely tested than scientific theories. Moreover, there is no theory-free language to describe the data, because myths (that is, primitive theories) arise together with language. There are no living things, neither animals nor plants, without problems and their tentative solutions which are equivalent to theories; though there may well be, or so it seems, life without sense-data (at least in plants).

Thus life proceeds, like scientific discovery, from old problems to the discovery of new and undreamt of problems. And this process – that of invention and selection – contains in itself a rational theory of emergence. The steps of emergence which lead to a new level are in the first instance the new problems (P_2) which are created by the error elimination (EE) of a tentative theoretical solution (TT) of an old problem (P_1).

9. Discovery, humanism and self-transcendence

For a humanist our approach is important for it suggests a new way of looking at the relation between ourselves – the subjects – and the object of our endeavours: the growing objective knowledge, the growing third world.

The old subjective approach of interpreting knowledge as a relation between the subjective mind and the known object – a relation called by Russell 'belief' or 'judgment' – took those things which I regard as objective knowledge merely as *utterances or expressions* of mental states (or as the corresponding behaviour). This approach may be described as an *epistemological expressionism* because it is closely parallel to the expressionist theory of art. A man's work is regarded as the expression of his inner state: the emphasis is entirely upon the causal relation, and on the admitted but overrated fact that the world of objective knowledge, like the world of painting or music, is created by men.

This view is to be replaced by a very different one. It is to be admitted that the third world, the world of objective knowledge (or more generally of the objective spirit) is man-made. But it is to be stressed that this world exists to a large extent autonomously; that it generates its own problems, especially those connected with methods of growth; and that its impact on any one of us, even on the most original of creative thinkers, vastly exceeds the impact which any of us can make upon it.

But it would be a mistake to leave things at that. What I regard as the most important point is not the sheer autonomy and anonymity of the third world, and the admittedly very important point that we always owe almost everything to our predecessors and to the tradition which they created: that we thus owe to the third world especially our rationality – that is, our subjective mind, the practice of critical and self-critical ways of thinking. More important than all this is, I suggest, the relation between ourselves and our work, and what can be gained for us from this relation.

The expressionist believes that all he can do is to let his talent, his gifts, express themselves in his work. The result is good or bad, according to the mental or physiological state of the worker.

As against this I suggest that everything depends upon the give and take between ourselves and our work; upon the product which we contribute to the third world, and upon that constant feed-back that can be amplified by self-criticism. The incredible thing about life, evolution, and mental growth is just this method of give and take, this interaction between our actions and their results by which we constantly transcend ourselves, our talents, our gifts.

This self-transcendence is the most striking and important fact of all life and all evolution, and especially of human evolution.

In its pre-human stages it is of course less obvious, and so it may indeed be mistaken for something like self-expression. But on the human level the self-transcendence can be overlooked only by a real effort. As it happens with our children, so it does with our theories: they tend to become largely independent of their parents. And as it may happen with our children, so with our theories: we may gain from them a greater amount of knowledge than we originally imparted to them.

The process of learning, of the growth of subjective knowledge, is always fundamentally the same. It is *imaginative criticism*. This is how we transcend our local and temporal environment by trying to think of circumstances *beyond* our experience: by criticizing the universality, or the structural necessity, of what may, to us, appear (or what philosophers may describe) as the 'given' or 'habit'; by trying to find, construct, invent, new situations – that is, *test* situations, *critical* situations; and by trying to locate, detect and challenge our prejudices and habitual assumptions.

This is how we lift ourselves by our bootstraps out of the morass of our ignorance; how we throw a rope into the air and then swarm up it – if it gets any purchase, however precarious, on any little twig.

What makes our efforts differ from those of the amoeba is only that our

rope may get a hold in a third world of critical discussion: a world of language, of objective knowledge. This makes it possible for us to discard some of our competing theories. So if we are lucky, we may succeed in surviving some of our mistaken theories (and most of them are mistaken), while the amoeba will perish with its theory, its belief, and its habits.

Seen in this light, life is discovery – the discovery of new facts, of new possibilities, by way of trying out possibilities conceived in our imagination. On the human level, this trying out is done almost entirely in the third world, by attempts to represent, in the theories of this third world, our first world, and perhaps our second world, more and more successfully; by trying to get nearer to the truth – to a fuller, a more complete, a more interesting, powerful and relevant truth – relevant to our problems.

What may be called the second world – the world of the mind – becomes, on the human level, more and more the link between the first and the third world: all our actions in the first world are influenced by our second-world grasp of the third world. This is why it is impossible to understand the human mind without understanding the third world, the objective mind or ‘spirit’; and why it is impossible either to interpret the third world as a mere expression of the second, or the second as the mere reflection of the third.

There are three senses of the verb ‘to learn’ which have been insufficiently distinguished by learning theorists: ‘to discover’; ‘to imitate’; ‘to make habitual’. All three may be regarded as forms of discovery, and all three operate with trial and error methods which contain a (not too important and usually much overrated) element of chance. ‘To make habitual’ contains a minimum of discovery – but it clears the decks for further discovery; and its apparently repetitive character is misleading.

In all these different ways of learning or of acquiring or producing knowledge the method is Darwinian rather than Lamarckian. It is selection rather than instruction. But selection is a two-edged sword: it is not only the environment that selects and changes us – it is also we who select and change the environment. On the human level, we do this by co-operation with a whole new objective world – the third world, the world of objective tentative knowledge which includes objective new tentative aims and values. We do not mould or ‘instruct’ this world by expressing in it the state of our mind; nor does it instruct us: both, we ourselves and the third world grow through mutual struggle and selection. This, it seems, holds at the level of the enzyme and the gene – the genetic code may be conjectured to operate by selection or rejection rather than by instruction or command – and through all levels, up to the articulate and critical language of our theories.

References

- ARISTOTLE, *Metaphysics*.
 ARISTOTLE, *De Anima*.
 BERKELEY, Three dialogues between Hylas and Philonous, in: *Works*, eds. Luce and Jessop, Vol. II (1949).
 BOLZANO, B., *Wissenschaftslehre* (1837).
 BROUWER, L.E.J., Inaugural lecture, 14 October 1912; transl. A. Dresden, *Bull. Am. Math. Soc.* **20** (1914) 81–96.
 BROUWER, L.E.J., *Math. Ann.* **93** (1924).
 BROUWER, L.E.J., Consciousness, philosophy, and mathematics, in: *Proc. 10th Intern. Congress of Philosophy* (1949) Vol. I, fascicule II.
 BROUWER, L.E.J., On order in the continuum, and the relation of truth to non-contradictority, *Kon. Ned. Acad. Recht. Wet. Proc. Sect. Sci.* **54** (1951).
 BÜHLER, K., *Sprachtheorie* (1934).
 DESCARTES, R., *Discourse de la methode* (1637); transl. E.S. Haldane and G.R.T. Ross, Vol. I (1931).
 DUCASSE, C.J., *J. Phil.* **37** (1940).
 FEYERABEND and MAXWELL, editors, *Mind, matter and method, essays in philosophy and science in honor of Herbert Feigl* (1966).
 FREGE, G., Ueber Sinn und Bedeutung, *Z. Phil. und phil. Kritik* **100** (1892) 25–50.
 FREGE, G., Review of HUSSERL [1891], *Z. Phil. und phil. Kritik* **103** (1894) 313–332.
 GOMBRICH, E.H., Moment and movement in art, *J. Warburg and Court. Inst.* **27** (1964).
 GOMPERZ, H., *Weltanschauungslehre*, Vol. II/1 (1908).
 GOMPERZ, H., *Über Sinn und Sinngebilde. Verstehen und Erkennen* (1929).
 HAYEK, F. A., *The constitution of liberty* (1960).
 HAYEK, F. A., *Studies in philosophy, politics, and economics* (1967).
 HEGEL, G.W.F., *Enzyklopädie der philosophischen Wissenschaften* (third edition, 1830).
 HEINEMANN, F., *Plotin* (1921).
 HENRY, P., Plotinus' place in the history of thought, in: *Plotinus, the enneads*, transl. S. MacKenna (2nd ed., 1956).
 HEYTING, A., After thirty years, in: *Logic, methodology and philosophy of science*, eds. E. Nagel, P. Suppes and A. Tarski (1962) pp. 194 ff.
 HEYTING, A., *Intuitionism* (1966).
 HEYTING, A., Informal rigour and intuitionism in: LAKATOS [1967].
 HUSSERL, E., *Philosophie der Arithmetik* (1891).
 HUSSERL, E., *Logische Untersuchungen*, Vol. I (2nd ed., 1913).
 KANT, I., *Kritik der reinen Vernunft* (first edition 1770, second edition 1778).
 KLEENE, S. C. and R. VESLEY, *The foundations of intuitionistic mathematics* (Amsterdam, North-Holland Publ. Co., 1965).
 LAKATOS, I., *Proofs and refutations*, *Brit. J. Phil. of Sci.* **14** (1963–64).
 LAKATOS, I., editor, *Problems of the philosophy of mathematics* (Amsterdam, North-Holland Publ. Co., 1967).
 LAKATOS, I., editor, *The problem of inductive logic* (Amsterdam, North-Holland Publ. Co., 1968).
 LAKATOS, I. and A. MUSGRAVE, eds., *Problems in the philosophy of science* (Amsterdam, North-Holland Publ. Co., 1968).
 MYHILL, J., Remarks on continuity and the thinking subject, in: LAKATOS [1967].
 PLATO, *Phaedo*.
 PLOTINUS, *Enneades*, ed. R. Volkman (1883, 1884).
 POPPER, K.R., *Logik der Forschung* (1934), *The logic of scientific discovery* (1959 and later editions).

- POPPER, K. R., *The poverty of historicism* (2nd ed., 1960).
- POPPER, K. R., *The open society and its enemies* (fourth ed. 1962 and later editions).
- POPPER, K. R., *Truth and the growth of knowledge*, in: *Logic, methodology and philosophy of science*, eds. E. Nagel, P. Suppes and A. Tarski (1962a).
- POPPER, K. R., *Conjectures and refutations* (1963 and later editions).
- POPPER, K. R., *Of clouds and clocks* (1966).
- POPPER, K. R., *Quantum mechanics without 'the observer'*, in: *Quantum theory and reality*, ed. Mario Bunge (1967).
- RUSSELL, B., *On the nature of truth*, in: *Aristotelian Soc. Proc.* 7 (1906–7), pp. 28–49.
- RUSSELL, B., *Philosophical essays* (1910).
- RUSSELL, B., *Introduction to WITTGENSTEIN'S Tractatus* (1922).
- RUSSELL, B., *My philosophical development* (1959).
- WATKINS, J. W. N., *Hobbes's system of ideas* (1965).
- WHORF, B. L., *Language, thought and reality* (1956).
- WITTGENSTEIN, L., *Tractatus logico-philosophicus* (1922).