EPIPHENOMENALISM:

A CONFLICT OF SCIENCE AND PHILOSOPHY

A thesis submitted in partial fulfilment of the requirement for the degree of Master of Arts in the Faculty of Graduate Studies and Research

by

John Lachs, B.A.

Department of Philosophy
McGill University

May 1957

Montreal, P.Q.

I N D E X

CHAPTER 1

A BRIEF SURVEY OF DESCARTES' THEORY OF MIND

For a long time now, I have been thinking about the problem of body and mind and about the strangely wonderful phenomenon of consciousness. The concept of consciousness has always been a relatively neglected topic. As for questions about body and mind, one's first reaction is inevitably the impatient exclamation " But there is no problem there at all; " We are not all mind; everyone knows we are not disembodied spirits. On the other hand, we are not all body either; we are different - and, we prefer to think, better - than the brutes. We are mentoid creatures; hybrids of mind and of body. The ingredients are blended into a unity -- arranged in a working order.

This is the primary thesis; this is what is revealed to us in our everyday experience (for every moment of experience is a moment of revelation) -- or at least this is the way we tend to interpret the given. And after all investigation and all reflection has been done our conviction may not have changed. For it is neither the task nor the intention of philosophy to discredit, much less to 'refute' everyday opinions of the immediately given. The common man's beliefs are usually based on a solid common sense which unfortunately often evaporates in the heat of education; and the immediately given is quite possibly the only thing we have. It is much more important to shed light on what is observed; to establish vantage points which can be used to house the mind, bases the view from

which shows the bare facts in an intelligible light. Many philosophers may have gone beyond the transcience of the experience of the senses, but few have taken the heroic course of denying all validity and reality to this world.

These are not original thoughts but they are valid thoughts. Descartes voices them in his famous letter of June 28, 1643 to Princess Elizabeth.[1] We have three primitive notions, Descartes holds, those of body, mind and the union of body and mind. The concept of mind is best understood by the rational faculty --- the pure intellect. The concept of body is best understood by the rational faculty with the aid of the imagination, as in the science of geometry. The union of body and soul, however, cannot be clearly and distinctly understood; it cannot even be clearly and distinctly imagined. It is known by means of the senses. For the senses themselves presuppose a union of body and soul (how else could a sense organ affect the knower?) and are not directed to body in general (as is the imagination), but to actually existing and sensibly presented bodies in particular. I am uneasy about these three concepts; I am not sure that I do in fact have them. Let me investigate.

What is my evidence for saying that I have a clear and distinct idea of body and a similarly clear and distinct idea of mind? First of all, what is a clear and distinct idea? For Descartes, an idea is an object of knowledge. When the idea is entertained, the theory of representative

---

1. Descartes: Philosophical Writings ed. by E.Anscombe and P.T. Geach (Nelson: Edinburgh,1954); pp. 279-82 (Hereafter referred to as 'Writings').

perception prompts him to say that it also constitutes knowledge of an object. Now a clear idea is one which we can recognize in the natural light of reason as self-identical and distinguishable from all other ideas. A distinct idea, on the other hand, is one which is not only clear and manifest in its presence, but the inner structure of which is also perfectly understood. Distinctness implies clearness; but the clearness of an idea does not necessarily imply its distinctness.[2]

The evidence Descartes adduces in favor of the view that we do in fact possess such clear and distinct ideas of body as separate from mind and of mind as separate from brain, is reasoned, or metaphysical. What I mean by this is as follows. Descartes nowhere refers to special experiences which could be taken as originating or symptomatically designating the presence of such ideas. He is not, for example, referring to the feeling-- which we may have after a long period of contemplation or of fasting---of sheer spirituality; to the experience of feeling as though we were disembodied minds.

Descartes holds that the search after truth begins when one grows dissatisfied with one's obscure and confused ideas. This intellectual crisis leads us to the method of systematic, hyperbolical doubt where we reject everything that can possibly be doubted. And what is it the veracity of which we cannot doubt? Our senses may deceive us and our imagination may deceive us. Our thoughts can be deceptive; we often find that we are

---

2. Descartes: Principles of Philosophy, I. #45-46 (Writings p.190). See A. Koyre's Introduction to "Writings", p.xxx. I will be referring to Principles of Philosophy hereafter as "Principles".

in error-- we may quite conceivably always be in error. But even though the content of our thoughts may be deceptive, the act of thought leads us to the first certainty. I cannot possibly doubt that I doubt - this is where Descartes reaches bedrock. I doubt therefore I exist. Dubito ergo sum. Doubting is a form of thinking --cogito ergo sum. I can be perfectly certain of this. My confused idea has become clear. I think therefore I am; this is the first building block of the Cartesian system.[3] In itself, as Gilson holds,[4] it is not the first principle, but it develops into the first principle and defines the orientation of the system. If it is the inevitable result of the Cartesian method, it is also in dangerous proximity to the source of a possible fallacy. And this is a "possible" fallacy in two senses. On the one hand it is possible to take a route from the " Cogito " that leads straightway into this fallacy. I shall try to show that Descartes has taken this route. On the other hand, since we are in possession of no certain knowledge concerning consciousness and the relation of mind and body, it is only possibly (at most, probably) the case that Descartes' view of the matter amounts to a fallacy. Once again, I shall argue that this is more probable than not.

Much has been written on the illegitimate introduction of the self - the " I " in the Cogito. His critics have insisted that on the basis of the evidence he presents, Descartes could say no more than that thoughts occur, or exist. We may even speculate (if we are at all linguistically

---

3. E. Gilson: Commentary on the Discours (Paris,1935) p.299
4. Versfeld shows this decisively. See Marthinus Versfeld: An Essay on the Metaphysics of Descartes., (Methuen & Co. London,1940), esp. Chapter V.

minded) as to whether the mistake did not originate in Latin grammar. In Latin no separate personal pronoun is used with verbs. The person-reference is already included in the verb form. But these speculations are totally invalid if we speak in Descartes' terms. For ideas are qualities or modes, and to nothing no affections or qualities belong. Everything there is must be either a substance or the modification of a substance. Now if thoughts (which are evidently not substances for I cannot clearly and distinctly conceive them as being in need of no other thing in order to exist) are thus modes, they must be the modes of some substance in which they inhere. This substance is the self.

And this is just the question. In what sense can we say and should we say that the self is a substance? If we apply the scholastic terminology of substance-essence (quality) - mode, we are bound to run into difficulties. Substances are reciprocally exclusive. From this it follows that the essential attributes of substances must be also toto coelo different. The modes of substances must be different as well. Between substances there is no point of contact. If we distinguish between mind the essential attribute of which is thinking (intellection, consciousness) with the ideas as its proper modes, on the one hand, and matter the essential attribute of which is extension with figures as its proper modes, on the other; how can we possibly conceive the cooperation or the interaction of the two? If there is no point of contact, there is no point where causal influence could be transferred. Where there is such a bifurcation, such a complete dichotomy of substances, we are inevitably doomed

to an inexplicable parallelism or an intellectually totally inadequate occasionalism, as the only logically consistent position. In the light of this we recognize Descartes' statement in his letter to Princess Elizabeth, as an ultimate defence against the violent objections to his philosophy of mind. Pure intellect knows, intuits the soul as seperate from the body, and intellect together with the faculty of imagination intuits the body as seperate from thinking substance: - the union of the two, however, we do not know, we do not and cannot even hope to conceive it clearly and distinctly in the natural light of reason, for it is an experiential fact revealed to us through the senses.

In fact, Descartes seems to be saying, we experience the unity of body and mind; in fact the two are united and they interact. Thus the seperation of body from mind, of extension from thought as different substances is the result of a reflective ontological analysis of the experience-situation. Such " real distinction " between two substances is, for Descartes, " discovered from the mere fact that we can clearly and distinctly conceive one without the other. "[5] Confusion is increased by Descartes' statement in the same paragraph, where he maintains that " even if we supposed that God had conjoined some corporeal substance to a conscious substance so closely that they could not be more closely joined, and had thus compounded a unity out of the two, yet even so they remain really distinct." What we are asked here is to conceive of the soul and of the body as a single thing (conjoined) and as two things (seperate), which is self-contradictory. Descartes attempts to resolve the

---

5. Principles 1.60, (Writings, p.193)

contradiction by the suggestion (mentioned above) that whereas we can clearly and distinctly conceive the latter, we can only feel or experience the former. I am uneasy about this solution, for it comes to this. What is closest to me, what I experience--what I am-- every day of my life is essentially unintelligible. The origin and the destination of the soul as an immortal substance is clearly and distinctly known, but common experience is inexplicable. It is highly improbable that this should be the case. For this amounts to little more than an emphasis on the rationalistic dualism of reason (thought) and experience, in order to explain or explain away the difficulties arising from the previous emphasis on the dualism of mind and body.

But now I can suddenly see that Descartes' sharp distinction between body and mind is not wholly fallacious at all. For don't we in fact know of this radical disjunction of things in the world? I think and I know and I am conscious and I have feelings. It is evident that stones are different from me in this sense, even if I can judge only from their behavior. The cumulative evidence of physiology goes to show (much more conclusively today than in Descartes' time) that the body is a machine, a precision mechanism that requires fuel, proper maintenance, in return for which it works, executes orders. If the body is not goal-directed of itself, as Descartes asserts it isn't, the final causes superimposed on its blind causality cannot originate in matter. Matter fills space and is divisible, perishable and compounded: it does not spontaneously generate

motion. The realm of matter is the realm of mechanical causality; a colorless, lifeless, insensitive world of extension. Bits of matter filling the universal void; all simply located and externally related to one another---atom propelling atom upon impact. In contrast with this bleak and senseless world, mechanical and mathematical, - the human mind knows the force of feelings and emotions. Contrast the warm life of the senses, the beauty of art, the intensity of pleasures and pains with the cold precision of matter. Random unplanned mechanical action could not sustain organic life for very long;- without final causes directing the actions of the body we would very soon die--of starvation or cold or heat or just from sheer inactivity. And the I (whatever this mysterious personal pronoun may stand for) does not fill space; if you reflect you cannot even answer the question " Where are you ?" You can say, " Here is my hand " and " Here is my neck " and " Here is my head "--- under special circumstances you can even say " Here is my brain ". But where is the "me" to which all these,- hands, neck, head and brain belong? Of which it makes no sense to say that it is mine for it is me. It is that to which all that belongs to me, belongs. The genitive of the personal pronoun is meaningless when it is applied to it.

The new physics taught Descartes that the material universe consisted of nothing but space-filling stuff in motion, assuming an ever-changing, infinite variety of forms. Color and sound, taste and smell did not

belong to this mechanico-mathematical continuum; they were our contributions. When the feather tickles me, the tickle is in me not in the feather: - this is the point where the world of science prepares the first clash with the world of our everyday experience. Hence the belief that the outer world of matter is quantitative; the inner world is qualitative. The outer world manifests constant diversity; the inner world is unified in the "I", the single and the simple, indivisible self-substance. Matter undergoes constant change; the self is unchanging and abiding. In this mood I ask myself boldly: " What is there that is in common between the mind and the body?" In this mood (and this is a legitimate mood) I am bound to end up by wondering about the famous question, " Why does the mind have a body?"

These, then, were good reasons for Descartes to decide in favor of the essential diversity of mind and body and consequently of mind and brain. They are also good reasons for us; reasons which should help us to understand Descartes and which we should remember after we have fallen into the opposite mood of emphasizing the unity of body and soul.

In the light of these considerations I think I can now appreciate why Descartes insisted on drawing a sharp line of division between body and mind. There were also other considerations involved in his refusal to allow the possession of a rational soul to the brutes; but we are not at present concerned with those. The question I would rather ask is this. If we distinguish between two things which seem to us to be quite different,

whose properties seem to be, as a matter of fact, even contradictory and mutually exclusive--is there any reason why we should state this distinction in the substance-attribute terminology? As a psychological argument in the case of Descartes, I would adduce the fact that he has not quite succeeded in breaking with the Scholastic tradition. He accepted the traditional subject-predicate form of proposition unquestioningly. He inherited the subtance-attribute mode of thought, and applied it as the one that was closest at hand. But there is more to it than this.

Let me try to state the point this way. If we discern radical differences between mind and matter, does this mean that we have to think of them as different substances--viz. toto coelo different? An unconditional, total dualism is intellectually unsatisfying. It introduces a radical bifurcation, a gap that cannot be bridged and forces us to resort to measures of artificial juggling and frequent reference to an omnipotent God, in order to hold the system together. Most important of all, it does violence to my every-day experience, the evidence of which I would be very reluctant to disregard. Such a Cartesian dualism will not do. But Descartes had gotten into this position by asserting that matter and mind were different substances. To understand him I must find out why he held this view.

In Principles of Philosophy 1.60 we find the following statement.

" Real distinction between two or more substances... is discovered from the mere fact that we can clearly and distinctly conceive the one without the other."[6] It seems, then, that we do have a clear and distinct idea of the mind as seperate from everything else. This idea is in fact the first clear and distinct idea we have; it is the Cogito. " Cogito ergo sum"...this is the first indubitably certain proposition in Descartes' system. It is the first truth; it relieves me of the hyperbolical doubt. Sometimes I think I even experience the detachment of a disembodied spirit; as though I were nothing but consciousness and conseciousness were self-sufficient and autonomous--as though nothing but thoughts were necessary for the existence of thoughts. In this mood Descartes confidently exclaims; " Assuredly the conception of this "I" precisely as such (as a thinking being) does not depend on things of whose existence I am not yet aware."[7] I have a clear and distinct idea of mind, of the act of thinking, which I cannot doubt. In this idea antecedent conditions are not thought, whence Descartes assumes that no antecedent conditions are needed for its generation. " A substance is in need of no other thing in order to exist"[8], if I can conceive and understand something clearly and distinctly apart from all other things, then it is truly distinct and different from them-- for God can effect the seperation. Created substances, according to Descartes' definition, are things that need only the co-operation of God in order to exist.

---

6. Writings, p.193
7. Meditations on First Philosophy VI, (Writings, pp.114-115)
8. Principles I, 51 (Writings p.192)

I tend to challenge the view that we can clearly and distinctly understand mind as separate from matter. We must recall that the Cogito is the first term, the first unconditional certainty only in the order of proof. Indubitably, on its discovery depends the whole future course of the argument. But the order of demonstration is not necessarily identical with the order of generation; the logical order of proof is not necessarily and always the same as the order of being. It may very well be the case that though in thought I can conceive of nothing more ultimate than the Cogito; in fact, viz. in its actuality, this thinking being whose existence has been affirmed is dependent on certain antecedent conditions which are not thought of when it is contemplated, but without which it could not possibly exist. I wonder if Descartes' argument amounts to more than the disputable (fallacious?) statement that the idea of an object can be analyzed into the idea of the conditions of that object. It is certainly true that this is not always the case. Let me give you an example.

I can think of a house without at the same time thinking of all the bricks, wood, cement, electrical wiring, plumbing and working hours that have gone into the building and maintenance of that house. In fact I can think of the house even though I have no idea what these building materials were and how a house is usually built. I can always think of the finished product without at the same time also thinking of the process of production. Does it not seem probable that this is just where Descartes was wrong; he held that if we can think of the mind as seperate from

everything else, then it is in fact independent and in need of nothing beyond itself in order to exist?

It is evidently the case that the mind when contemplating its own nature tends to disregard the necessary conditions of its existence. It is, to use a suggestive phrase for a feature I can't describe any better, internal-to-itself. I order to consider itself it does not have to go outside of or beyond itself. I would venture to say that Descartes was probably familiar with this characteristic and consequent tendency of the mind to disregard the conditions of its conscious functioning; however, he was misled--through his order-mistake--into believing that from the fact that consciousness is the first term in the logical order of exposition it follows that it is also a first term, and is thus without necessary antecedents, in the ontological order of being. However, this does not follow. The inference can only be legitimized if we additionally assume the rationalistic premiss of " a nosse ad esse est vera consequentia". If we grant this we have bridged the gap between thought and being and our problems are solved. However, I would be very reluctant to grant this premiss, since it assumes what we have set out to prove. More correctly, it resolves the problem by granting likely or desirable premisses instead of instigating further investigations, even in the face of the possibility that the outcome of these investigations might be intellectually disturbing or undesirable.

The resort, then, to a rationalistic principle for the justification

of Descartes' position is of little use to us. It is of interest, however, because it reveals that without it Descartes can be shown to have made an order-mistake. It is this mistake alone that enabled him to conclude that the pure act of thought could exist without antecedent conditions (brain processes); that consciousness was in need of nothing beyond itself in order to exist. In accordance with this he concluded that mind was a substance completely independent of and different from matter, in general, and from the body and the brain to which it was attached, in particular.

I must repeat, I am not at present arguing that it is inconceivable that this should be the case. I am only saying that a radical dualism will not quite do justice to my experience, the evidence of which I regard as important. Further, I took this insufficiency as my clue in investigating where the radical bifurcation was introduced. This led me to the substance-attribute mode of thought, which I have subsequently traced to its origin in an order-mistake in the Cogito. The order-mistake is the possible fallacy to which I have earlier referred; it ceases to be a fallacy only if we grant the rationalistic principle of a parallelism between the realm of knowledge and the realm of being--together with its corollary, the representative theory of perception. However, this is far too much to ask us to grant, only in order to maintain the duality of body and mind. It is preferable to reject provisionally the Cartesian dualism and go on to examine new paths that are more promising.

It is much more probable that the grand assumption demanded of us cannot be justified, than the opposite. There is no better indication of the break-down of rationalism, than to have to accept a rationalistic premiss on faith.

At this point it may be objected that we are not doing full justice to Descartes in this short chapter. Perhaps this is inevitable, but we must remember that in this thesis Descartes is mentioned for background purposes only. He is treated more especially for two reasons. First, he gave a clear and incisive exposition of the duality of mind and matter; he saw the problem though he did not succeed in solving it. Descartes was the one who defined the problem for subsequent philosophy to this day, so that developments in the philosophy of mind can best be traced by first returning to Descartes for general orientation. His importance is easily seen from the single fact that Santayana, writing close to 300 years after him, seems in many places to be answering him point by point, treating Cartesian problems in Cartesian language in order to be able to break the deadlock and disagree. Of this we shall give illustrations later on.

Descartes is also important from another point of view. It was the new physics that prompted him to exclude all mind, all intelligence, all purposiveness from nature. As we have seen, the great advance of physics resulted in a denudation of the physical world; what was left was a colorless, soundless mass of mechanical matter--measurable but not qualitative. In this sense Descartes was the first great philosopher to struggle with the

problems which the new conceptions of science were creating. The problems are still with us today; we still have not succeeded in solving them, and in incorporating the conflict which underlies them into a coherent and satisfactory picture of the world. The trouble is that we feel the moral life, moral values as endangered by the scientific advance. The fear is legitimate because the danger is present. And it is present to a much greater extent tody than it was in Descartes' time. The development of the biological sciences (including psychology) represent both a hope and a threat--especially in view of certain tendencies which press for a biology that is to be the extension of physics into the sphere of life. Possibly the greatest problem of modern culture is the effort of philosophy to digest the violent new perspectives of science, and its renewed attempts to facilitate an infiltration of moral values into the world of scientific research. Here we have a conflict between two sides of the human personality. From the point of view of this conflict Descartes must be considered a **pioneer**.

Thus our main concern is not Descaftes but the problem of consciousness. So far I have described some of Descartes' most important views on the body-mind problem, and have tried to point out where he has fallen short of giving a satisfactory solution of the problem and why. His fallacy may be called (in Santayana's terms) a fallacy of substantialization. If I would go a step further at this point and say: " substantialization of the insubstantial ", I would be well on my way towards epiphenomenalism.

However, that is not my intention at the present.

There is a simple mistake in Genesis 1, when the text is taken literally, to which we can profitably compare Descartes' fallacy. To the best of our knowledge it is the case that the origin of all the light (except for the negligible amount which originates in other fixed stars) reaching the earth by daytime is the sun. Astronomers and physicists could give us the details of how this light-energy is generated; they could supply us with figures and explanations in terms of the transformation of hydrogen atoms into helium. They would tell us that there is no light we know of which does not originate in stars (the sun, of course, is a star) or in a conglomeration of stars or in galaxies which are in the process of crystallizing into stars. Now Genesis 1, on the other hand, maintains that God has created light on the first day, whereas the sun, moon and the stars were created on the fourth day only. In this way the product (light) is held to be independent of its originating organ (the sun). The constant correlation of light and sun is explained by stating that the sun 'rules' the day and the moon the night. Note the similarities of this situation with the theory of parallelism. There the mind is seperated from the body; consciousness from concomitant brain-processes. Consciousness is an entity in need of no other entity for its generation (Descartes admits no corresponding brain processes for the mental process of pure thought); in need of only God in order to exist. The mind is not considered a function which becomes meaningless and empty (even passes out of existence) when it is seperated from its organ. The

organ-function relation only seems to hold between brain and mind; if the latter seems to be causally dependent on the former, it must be emphasized that this is only an apparant dependence and the trained eye of the mind easily observes that the two are really quite distinct -- which excludes the possibility of causal interaction between them. The brain is not an antecedent condition of the possibility of mental states just as the sun is not an antecedent condition of the existence of the God-created and substantially independent light that seems to emanate from it.

By means of this analogy we can see how much artificial pipe fitting of concepts and abstract entities would be neccesary to give some sort of an explanation of what is immediately observed. A physics or an astronomy that started out with the premiss that all correlation between light and its origination in hydrogen - burning stars is purely illusory would come up with strange results - if it could reach any results at all. I am using this analogy not to refute occasionalism or parallelism and through them the radical originative disjunction of matter and mind; I am using it only to show that just as a physics based on a similar premiss would be an improbably strange physics, so a metaphysics based on this is an improbable and strange metaphysics. There is an important use which we shall be able to make of another aspect of the light analogy in connection with epiphenomenalism. We shall talk about that at a later place.

CHAPTER 11

T. H. HUXLEY: CONFLICT OF EXPERIMENT AND INTROSPECTION

Let us now exchange the word "mind" for the word "consciousness". Mind is a term both vague and ambiguous; it is most advantageously employed as a collective noun for all mental states in opposition to their usual contrary, body. This is why I have used it in connection with Descartes, but now we will have to be more specific. And there are problems even in the case of Descartes; how should we render the Latin "cogitare" and the French "penser" in English? The accepted translation used to be "to think" and "cogito ergo sum" was translated " I think therefore I am (exist) ".

But this translation is faithful neither to the letter nor to the spirit of Descartes. Let me illustrate. In Principles of Philosophy[1] 1. 9 he says: " By the term cogitatio I understand everything that takes place within ourselves so that we are aware of it---And so not only acts of understanding, will and imagination, but even sensations, are here to be taken as instances of ' cogitare' ". In what sense can we say that for Descartes a sensation is a 'thought'--which seems to be the implication of the old translations?

---

1. Writings, p.183

2

In more recent translations this situation has been remedied by the introduction of the English word "consciousness" (and sometimes "experience") to stand for cogitatio. We can easily enough draw up the linguistic equation " mental states = conscious states ". There is, even in sensations as Descartes correctly remarks, a mental, cognitive factor---this is what we usually call perceptual consciousness today. Concepts as recognitional capacities are also involved, and these are more apt to be taken as the cognitional side of perception; but we are not concerned with these now. One reason why the rendering of cogitare as " to be conscious " is useful, is that it brings out a distinction which the word "thought" usually tempts us to overlook, namely that between consciousness and the conceptual factor in cognition.

3

According to N.K. Smith , Descartes uses 'consciousness' in two senses. In one sense it is employed as a general name for all states of consciousness, including the contents and (objects) of consciousness. This is the referential totality of conscious experience. In a more restricted (and, for Descartes, more proper) sense, however, consciousness is distinguished and separated from its contents as an ultimate unanalyzable simple force or light. It is in this conscious light that experience-contents appear---in this sense conscious-ing is the ultimate characteristic of the self as totally different from extended substance. In the light of consciousness (in

---

2. A good example of this is the E.Anscombe, P.T. Geach translation I am using.

3. " Studies in the Cartesian Philosophy" pp.90-91 (Macmillan: London 1902).

the second sense) the mind's eye examines the object of knowledge, viz. the ideas presented to it. This is the origin of the introspectionist tradition in modern philosophy for the effective formulation of which the terminology of consciousness is especially adapted. I may remark at this point that Descartes' contention that of all the things we know, we know the mind or consciousness best is in one sense true and in another false. This is due to the ambiguity of Descartes' use of the word consciousness. It is certainly the case that in knowing anything I know only my own conscious states---thus in this sense I know consciousness best because that is the only thing I know. However, from this it does not follow that I also know consciousness in the more restricted sense of a simple light, apart from its contents. Although it certainly seems to be the case that it is necessarily involved in the knowledge of anything, from this we can by no means infer that in knowing anything we also know it. For as we have already pointed out, a necessary condition is not necessarily revealed in the product it makes possible.

The above remark was in self-defence, as I must admit I have no idea what consciousness in this narrow sense is, despite the fact that I have been relying on it unquestioningly all my life -- to enable me to sense, to think, in fact to do anything in which I can say that I (as apart from my body, at this point) am involved. Indeed, some people would insist that consciousness is so close to me that it could not be closer--- that it is, in fact, me. If I would accept this statement

(which without examination I cannot do), I would have to admit that I just do not know what I myself am. And this is not such a surprising admission to make.

It is also not surprising if I say that though I do not at present know what consciousness is, I very much want to find out something about it. I am reminded of G. E. Moore's famous statement:

> "...the other element which I have called "consciousness" --- that which sensation of blue has in common with sensation of green--is extremely difficult to fix. That many people fail to distinguish it at all is sufficiently shown by the fact that there are materialists. And, in general, that which makes the sensation of blue a mental fact seems to escape us: it seems, if I may use a metaphor, to be transparent---we look through it and see nothing but blue; we may be convinced that there is something but what it is no philosopher, I think, has yet clearly recognized." 4

I think I can understand what Moore was trying to say. Our cognitive faculties are strictly functional; they are, so-to-speak, ' out-directed'. They look out beyond themselves at the object (this is what it means to be object-directed) and in the act of knowing they do not know themselves. The realization of this in one form or another goes back at least to St. Augustine. It is indeed a mystery (as he thought) that while we know so much about the external world, we know nothing about our own faculties of knowledge; the physiology of sensation, the neurology of brain processes, - and most of all about consciousness, which is certainly a necessary condition of knowledge in at least one very important sense of that word. The animal only acquires interest in itself when something

4. G. E. Moore: Philosophical Studies (Kegan Paul, Trench, Trubner: London, 1922), " The Refutation of Idealism", p. 20.

does not function well, when it falls sick. Otherwise, the cognitive organs disregard themselves and are submerged in, completely taken up with, the object. In other words, in knowing we don't know knowing or the conditions of knowing, in seeing we don't see the eye; the intermediaries are readily overlooked, just as in looking through the window we do not see the glass but the birds outside and in the microscope we disregard the lens and see only the cell-structure it reveals. This essential ignorance has too often been covered up by the systematic guessing-game of epistemologists and psychologists. It is extremely hard to make consciousness its own object, to make it turn in (180^{o}) upon itself; but I feel it is not impossible. In a sense this whole thesis may be looked upon as an attempt to find some sort of meaning for the notion of consciousness. It is not easy to do this, and in the end it may well turn out that we have not succeeded in making the concept intelligible. But it is worth our while to make the attempt, because the phenomenon of consciousness is fascinating and the concept of consciousness is a suggestive one.

This leads me to a consideration of epiphenomenalism. I hope that in the course of the argument it will become clear how closely the epiphenomenalist position is connected with a certain conception of consciousness. I will try to show what this conception is and subsequently I will attempt to give an estimate as to how far the position that consciousness is the distinctively mental activity, can be maintained.

I will start up with an exposition of the views of T. H. Huxley, the most prominent exponent of the theory in the 19th century. His most important work on this topic is the essay " Animal Automatism"[5].

Epiphenomenalism maintains that human beings (as well as most higher animals) are conscious automata. For this reason it has been alternatively called the Theory of Conscious Automatism. Its origin goes back to Descartes' statement of the problem of body and mind. Nothing will illustrate this better than the fact that Huxley himself quotes Descartes in " Animal Automatism"[6], and in another essay[7] traces two trends, both derived from Descartes, the logical developments of which seem to conflict and thus have to be reconciled. The solution, so it is thought, lies in epiphenomenalism.

As Norman Kemp Smith has pointed out, there is an essential inconsistency between Descartes' physics and his metaphysics. The physics inevitably terminates in a materialism. The metaphysics, on the other hand, leads to Berkely and Kant and an essentially subjective -- at best critical-- idealism. Epiphenomenalism is an attempt to reconcile these antithetical views---which amount to opposing tendencies of development and which, to a great extent, underlie the dilemma of the 20th century. In other words, epiphenomenalism is a hybrid theory originating in the desire

5. T. H. Huxley: Animal Automatism and Other Essays, (Fitzgerald:New York 1884). Referred to as "Animal" hereafter.
6. Animal, pp.1ff
7. T.H. Huxley: Lay Sermons, Addresses and Reviews, (Macmillan:London 1870) "On Descartes' Discourse", pp.351-378. Book hereafter referred to as " Lay Sermons ".

to reconcile the conflict between scientific experimentalism and moral introspectionism. It is essentially a half-way position; its value lies in the indecision which bred it and in the fact that it involves the recognition of a basic problem which many other theories tend to overlook.

Descartes held that brain and mind stand in the relation of interdependence in some cases and independence in others, in the following manner.

1. Some neural processes do not produce corresponding conscious processes---reflex actions, for instance (which is not to say that we cannot grow conscious of them).
2. Some processes of the brain and the nervous system necessarily produce conscious states. Sensations, originating in sense organs and terminating in the conscious mind, would be an example of this. Actually, Descartes made a threefold division of experiences within this category into appetites (hunger), passions (anger is an example) and sensations.
3. Some conscious states can exist without corresponding brain processes. Descartes held that cognition and volition belonged in this category of spontaneous, purely spiritual but causally active states.

An alternative theory would be as follows:

1. Some brain processes do not produce corresponding conscious states.
2. The brain is the organ of consciousness; in this way all conscious states have corresponding nervous (neural) processes.

But this does not give us epiphenomenalism, though it establishes a naturalistic basis for that theory. For the epiphenomenalistic view we have to add a further proposition.

3. All conscious states depend for their existence on antecedent nervous states, and can in no way causally affect these states.

Epiphenomenalism insists on the causal impotence of consciousness.

Huxley was fully aware that the localization of brain functions had been one of the great aspirations of neurologists since the beginning of the nineteenth century, and he took it as evident that the organ of consciousness is the brain. In accordance with the belief current in his day, he designated the 'anterior division' of the brain as the seat of consciousness. Today the region most often quoted as the seat of consciousness is the mid-brain or diencephalon[8]; mainly because of new neurological investigations with some of the most important work being done at McGill by Dr.Penfield and his associates, but experimental

8. Dr.Wilder Penfield and Theodore Rasmussen: The Cerebral Cortex of Man, (Macmillan: New York,1952).

evidence, even to this day, warrants no final conclusions on the topic. At any rate, it seems clear that human consciousness is closely bound up with the human cortex; similarly in all other animals it is connected with the higher regions of the brain.

Now operate on a frog; sever the lower part of its spinal cord. This makes any sort of connection with the brain impossible; the lower segment of the spinal cord is independent of any 'conscious control'. If you then expose the frog to a number of tests, surprising facts will come to light. When an irritant is applied to the leg-- a leg which cannot occasion the feeling of pain, because its nerve connections to the brain and hence to any possible conscious mind are cut off---the frog will apply a foot, and if that is held back, the other foot! - to rub off the acid. The nerve and muscle coordination necessary for such an action is truly amazing. And it has been accomplished solely on a mechanical stimulus-response basis by the lower spinal cord.

When the spinal cord is left intact in the frog and only the foremost half or two-thirds of the brain is taken away, the extent of the nervous coordination manifested by the animal is even more surprising. The central nervous apparatus is sufficient to make the frog swim, retain complete balance, jump, walk, eat and perform a number of other complex actions. When only the anterior part of the brain is removed, the frog hears nothing and sees nothing and is presumably without consciousness-- and yet it performs actions which we had previously thought only animals

in full conscious control of their body could perform. By placing obstructions in its way (which the frog will carefully avoid) we may even conclude that although it cannot see, visible bodies seem to act through its eyes in the form of nervous impulses penetrating its brain, and through its brain affecting the body's motor mechanism of movement.[9]

Now add to this the evidence of complex operations performed by human beings in a purely mechanical manner; for example the phenomena of somnambulism and habit-actuated actions. Huxley realized that he could make a much better case for the Cartesian doctrine of the automatism of animals than Descartes himself could. Today, we can make a still better case for the automatism of human beings and animals, than could Huxley. The experimental evidence of science seems to be in favor of an underlying, basic biological mechanism. We may call it an ' organic mechanism' if we wish, so long as we realize that it is a mechanism. It is evident that the vital functions, such as the beating of the heart and breathing, are automatic. We also know that most of the actions a human being performs are recurrent, repetitive actions; habit-actuated. And as for the acquisition of habit and the intellectual functions: we are told that at the basis of most of our learning is Pavlov's conditioned reflex. For the vital functions it is reflex action, for action it is habit, for learning it is the conditioned reflex; none of these involves

---

9. "Animal", pp.7 ff.

consciousness as an antecedent condition of its possibility. Machines have been constructed (and we wouldn't be willing to admit that machines are conscious!) which imitate the conditioned reflex. Machines have been built which learn by trial and error. Machines have been built which manifest complex patterns of social behavior. If all this can be done without the intervention of consciousness--is it not possible that everything is done without it?

This, then, is the way the doctrine of epiphenomenalism was born. The scientist-philosopher (a strange combination, which unfortunately rarely works without injury being done to at least one of the two fields ----and often to both) finds himself in a situation in which he has succeeded (or he expects he will succeed) in explaining everything in mechanical, material terms. In effect there is nothing left that could be assigned to the functioning of consciousness. The scientist feels forced into the position of asserting that on the basis of his findings and expectations consciousness has no causal efficacy. The machinery of the body is sufficient to do all the work.

There is, however, an additional diagnosis of consciousness that is made by the epiphenomenalist. The scientist admits the existence consciousness, but confronted with its basic, overriding subjectivity, does not know what to do with it. What else can the scientist do but state, with Huxley, his belief that " we shall, sooner or later, arrive

at a mechanical equivalent of consciousness, just as we have arrived at a mechanical equivelant of heat "[10]? Now the epiphenomenalist goes a step further; he states--and this is the vicious aspect of the theory-- that since we can make nothing of consciousness on a scientific basis, its causal efficacy must be nil. It cannot be useful for the organism, and on the basis of the economy of nature we must maintain that what is not useful cannot be necessary. Consciousness is a waste product.

This, then, is where we are led by the development of Descartes' position to its logical conclusion. If everything material, including the human body, conforms only to mechanical laws, what place shall we find in the world order for consciousness? Whereas Descartes started out with the substantiality of the conscious mind, the position is now reversed and the epiphenomenalist ends with the conclusion that consciousness is altogether insubstantial. But underlying all this is the basic paradox of mind and body. All our knowledge is couched in conscious terms and may ultimately turn out to be only "knowledge" of states of consciousness. The material world is apprehended only under the forms of the ideal or mental world. When I speak of matter and impenetrability, for instance, and say that the two are inseperable, this statement can be reduced to the much less pretentious sentence, " my consciousness of matter and my consciousness of impenetrability are always conjoined." My consciousness reveals to me a world of matter and force, and I naturally regard the mind as an inhabitant of this world. But,

---

10. Lay Sermons. p. 372

on the other hand, this material realm only appears to me in the context of conscious experience; I have no proof that it exists apart from my experience of it. It pays off if I believe that it does; the animal cannot live without this belief, but even the fact of the utility of this belief is only another appearance in the conscious manifold. The conscious mind exists in the natural world; but the natural world exists only as an appearance in the conscious mind. I am a creature in the world; but the world is my creation. We must ask the question: Can this paradox be resolved? Can the epiphenomenalist resolve it? Is it not strange that everything of which we are conscious is interpretable only in terms of physical science (in the wider sense, in which physical science includes biology), whereas consciousness itself is not[11]---- and without the light of consciousness nothing could appear and there would be no self-subject to which anything could appear. " Les extremes se touchent "; this is quite correct in this case of naturalism and idealism, if by this meeting we mean an encounter in conflict, a tension---almost on Hegelian principles--which it is extremely hard to resolve. Indeed, we have this added disadvantage, that if we were Hegelians we could at least be sure that a resolution of the conflict is possible; as it is, however, at the present we are not even convinced of this. And even if such a reconciliation is in the long run possible, epiphenomenalism may not be the theory on which this will be based.

---

11. This, I believe, is one of the reasons why philosophers have occasionally been tempted to reject consciousness as a factor separate from the flux of experience contents. See esp. William James:" Does Consciousness Exist?" -In:" Essays in Radical Empiricism,"(Longmans, Green & Co: New York,1912),pp.1-38.

Let us now have a closer look at the tenets of the Theory of Conscious Automatism, as it was put forward by Huxley. We should not be deterred by the metaphorical language we shall occasionally have to use; epiphenomenalism has never been very widely held and consequently nobody has ever made an attempt to state the theory systematically--- to set forth the assumptions and the arguments in a logically coherent order, to evaluate the evidence and then to draw the conclusions. In the place of explicit statements we will have to rely on hints from the few philosophers who have held this theory, and often invent our own analogies to make the writers' points evident. Perhaps it is true that all philosophical systems ultimately depend on spatial metaphors and analogical images transferred from everyday experience to serve as explanations of difficult metaphysical insights. Not many men can have these insights; and often they can be communicated only with great difficulty, and almost always very imperfectly. This need not, however, prejudice us against their probability-value.

The first assumption of the epiphenomenalist is no different from the assumption of any other human being; he holds that men as well as higher animals are conscious organisms. The doctrine of the privacy of the conscious field has been viciously attacked on a number of different grounds in the recent past. However, it seems at least as indubitable as any empirical matter of fact can be, that my consciousness is private to me and that I have access to no other conscious field beyond my own, though I willingly infer that such conscious fields do in fact exist.

We can grant this assumption without further discussion; for even though it should turn out that no other being beside myself is in any way conscious (though I cannot think of a way in which this could be proved), the epiphenomenalistic arguments would still apply---in the single instance of a conscious field with which I am acquainted and the existence of which I cannot deny.

If it is evident that nervous processes are the antecedents of mental states, we may now ask the question whether it is equally evident that there are cases in which conscious states are the immediate antecedents of neural processes. We are immediately tempted to answer in the affirmative. Consider the phenomenon of volition. Wishing for something and then deciding to do it: this certainly seems to be an instance of a conscious state issuing in action---an action which, we prefer to think, would not have been done had we not decided to do it.

The epiphenomenalists hold that the belief that the conscious self has access to the effector mechanisms of the body is based on an illusion. At a later place we shall discuss Santayana's account of the origin of this illusion---for now it is sufficient if we note that it is difficult to see how a volition could cause (or occasion) the occurrence of a brain process (eventually terminating through the appropriate nervous messages in the execution of an order), if we hold that the brain is the organ of mind, that consciousness is a function and not an entity (as

with Descartes) and that accordingly there is a one to one correlation between brain processes and mental processes. It seems to me that if we hold that conscious events cause neural events, we can no longer say that brain is the organ and consciousness the function, because the relation of organ to function is one of logical antecedence and temporal priority.

Huxley holds that there is only a one-way traffic of causal influence between the brain and consciousness. Neural events can and do cause states of consciousness; conscious events cannot and consequently do not cause neural events. Consciousness is simply a collateral product of the functioning of the nervous system - brain circuit; a shadow that accompanies but in its insubstantiality cannot influence the body. It is a silent witness---the analogy of the witness is an important one for epiphenomenalism--- a spectator that cannot join in the game. Volition, the mental state in which a conscious effort seems to be directed at the accomplishment of an action, is an emotion, or feeling, which indicates that physical changes in the brain have taken place. It is a sign <u>of</u> neural events that have taken place, not a sign <u>for</u> such events to take place. Conscious experience is the way <u>we</u> are informed by the nervous system and the brain, of what action has been taken. We may, indeed we will have to, ask the question who or what is this 'we' that is informed in this way of the actions that have been taken? Is it not the case that in some very important sanse <u>I</u> originated and <u>I</u> performed the action in question? What conception of the self and of personal identity is implicit in the

epiphenomenalist position? This is an interesting and revealing question and at a later point we shall discuss it at length.

The notion of a one-way causal connection between neural events and consciousness is a difficult one, and for the present I think it would be well to try to elucidate it. I will do this by means of two short analogies.

The first analogy compares the nervous system and the brain to a correspondent. The body is a complete machine, a closed circuit, so to speak. It receives and evaluates stimuli, calculates and effects the correct response (though there is, of course, a margin of error). In this at least the three basic factors of sensory input, built-in tendencies of organs and stored memory-information are involved. Consciousness, however, is not involved. The action is selected and subsequently arranged for by an autonomous central process, which is a physical process (or group of processes) in the brain. Generally speaking, the brain keeps us informed of what is going on, but this information must be conceived of as information which is transmitted in a letter; it never reaches the recipient prior to writing or mailing. The temporal difference between, say, a sensory stimulus registering in the brain in the form of a brain event, and the same registering in consciousness as a sensation, is, of course, very little. But it is sufficient to make all effective causal action on the part of [redacted] consciousness impossible. For if there is a short span of time between the brain event and the conscious state, and the brain event is sufficient

in itself to select actions as well as to carry them out by way of its control over the effector organs, then the conscious field is even temporally unable to interfere. The cortex writes the letters when the action is being taken; but by the time the letters reach the conscious self the action has already been done. The conscious self is half a step behind the brain and it can never hope to catch up to it. What it could act on (if indeed it could act at all) would be something that has not as yet been done. But what has not as yet been done the conscious manifold knows, as yet, nothing about. What it knows about, however, has already been done. In this way consciousness can act only on the past, and that is an absurdity. If we followed this up we would find the conscious self in a situation analogous to that of an uncle who keeps up a vigorous correspondence concerning the advisability of an action which his rash nephew, without the old man's knowledge, has done long ago.

This analogy is not an argument; we should neither press it nor insist on its inadequacies. Its function is one of elucidation--it is helpful in pointing out at least one sense in which we can conceive a natural event to be causally inert. In an attempt at a further clarification of Huxley's position I will now introduce the second analogy, which is more apt to do justice to the essential privacy of conscious experience. Let us conceive of the conscious field as the diary of the nervous system and of the brain. Every word, or every

sentence in the diary stands as a sign for a stimulus or a process of registering information or a selection or an order of the brain complex. But not every action of the brain registers in consciousness; just as not all actions of a man are entered in his diary---some are not important enough, are a matter of mere routine; others can't be noted for a lack of space; others again are just given passing mention which one is apt to overlook. Now, if we say that the conscious manifold is the diary of the brain, it is not clear whether we should extend the analogy and posit a mysterious decoding agency in the self to read this diary, or else be satisfied to maintain that the self is not different from the conscious manifold--that the diary is somehow both intelligible and intelligent. From the awkwardness of this last sentence it is evident that the analogy is, at this point, stretched too far. But the problem of the nature of the self is raised once again.

Consciousness, then, is a witness of what is taking place in the brain. It is its own candle, its own luminosity, in the light of which some of the multitude of events taking place in the brain, are revealed. What shows up in the light of consciousness is symptomatic of what has taken place in the brain a fraction of a second earlier. The phenomenological description of experience contents stands for or signifies corresponding sets of physiological neural events. The former gives an account of something that cannot act, and it cannot act because it is and end-product---a function of animal life, with no survival value.

It is insubstantial: there is no native continuity belonging to it. Furthermore, what seems as a mental process is in reality only a series of conscious events, discrete and discontinuous, which must be re-created at every moment by the substantial process going on in the brain. Mental events in this way do not form an autonomous series; one conscious event cannot even cause another conscious event. All conscious events are caused by physical events and any continuity which may be observed in the mental series is due solely to the solid physical continuity of the underlying brain processes. In this way the Cartesian notion of temporal discontinuity is preserved; but it is restricted to the conscious series. The difference is that for the epiphenomenalists the renewed creation of the conscious manifold, from moment to moment, is accomplished not by God but by brain processes. Conscious events give us the appearance of a substantial process; in fact they amount to no more than a discontinuous series of events with central brain processes contributing the apparent continuity.

Huxley does not make an explicit statement to the above effect, but it is a corollary to the epiphenomenalist position. If the human being is compared to a cash register, we have another useful analogy, which will shed light on this point. If I punch $5 on the machine, a complicated series of events takes place and then the sign "$5" appears in the viewer. This sign, which simply shows what I have registered, in no way affects the money in the till ---- that neither increases nor diminishes. The amount of money in the till changes only when with a physical action

I add to it or take away from it. If I now think of the sign "5" as a conscious event, it is evident that it only registers the changes in the till (if I suppose myself to be honest in punching no more and no less than I add, etc.) and cannot in any way make the merchant rich or occasion his bankruptcy. The case is similar to Kant's interesting analogy of the hundred thalers.[12] Kant, in arguing against the ontological proof of the existence of God maintained that existence is not a predicate; in effect the issue was one of distinguishing between essence and existence and of stopping the practice of illegitimately inferring the latter from the former. The similarity between the two analogies is not accidental; its importance will become clear in my chapter on Santayana's ontology.[13] Essence and existence are different realms or kinds of being. Consciousness which Santayana calls "spirit" and existence are likewise different ontological realms.

Returning to the cash register, let us note another aspect of the analogy. Each sign in the transparent viewer (viz. consciousness) stands for a complex series of events. It would take pages to describe even the little we know of the series of physiological events taking place immediately prior to or even at each instant at which we have the sensation of blue; if we knew everything, the description would be considerably longer. We would have to speak of the activation of sense-receptacles, impulses sent

---

12. Critique of Pure Reason, p.A599 (N. K. Smith trans. Macmillan, London, 1953 - p.505)
13. Chapter III, below.

along nerve-fibres, synaptic junctions, coordinating centres, events in the optic lobe, in the diencephalon, events in cells and groups of cells, and so on. Of all this I am not aware when I am aware of blue. Of all the innumerable physiological events taking place when I am consciously writing these lines, or of the physiological events which enable me to think consciously of physiological events, I am not conscious. The conscious manifold is the simplest, most transparent thing in the world. The epiphenomenalist would agree with Descartes that mind is better known than body. As we have previously remarked, in a sense nothing but consciousness (including, of course, the contents of consciousness) is known. Consciousness is clear and distinct and precise and enormously simplified. If we supposed for the moment that the mind is in possesion of the power of causation, it seems to me that the epiphenomenalist would attempt to press home the reductio ad absurdum argument that it would not know how to employ it. Even if the conscious manifold knew what it wanted, it would not know how to get it; it could not manipulate the body (the means to that end) because it knows so little of the mechanism of the organic machine.

The body does everything there is to do; and as the signs change on the cash register (the flux of appearances in the conscious manifold), all they do is inform the witness of how much has been deposited. Without the smooth functioning of the machinery of the cash register the signs could not appear--without the actual transactions taking place the signs would be virtually meaningless. By themselves

the signs can neither influence the merchant's finances, nor occasion the appearance of other signs. The "$5" sign appears by itself and can in no way bring about the appearance of the "$10" sign, even if "$5" was, so far, always followed by "$10". The signs go up if and only if the appropriate mechanism sends them up.

But to the analogy of the cash register there is this strange corollary. The signs on that machine are functional, inasmuch as they are important for keeping the merchant informed of transactions completed and, in the long run, of his financial situation. Huxley would deny that the same holds of the human being. The sign on the human cash register does not inform the merchant; the cash register itself is the merchant---it does its own business. The self, the "I" which I know primarily---and possibly only---as a conscious being is not really the merchant it thinks it is. To modernize the famous driver-analogies of the Cartesians, we may think of the self as the pilot of a rocket. The pilot sits in the cockpit and manipulates the switches he finds, works what he thinks is the controlling mechanism, adjusts the fuel supply and the rate of acceleration---and he is completely satisfied because all is going well. He is in complete control of the machine. And, then, it turns out that the rocket is a guided missile. It is fatally set on a course---the 'pilot' is a spectator in the cockpit, who attributes the fine functioning of the machine to his own sporadic efforts at the 'controls'.

Needless to say, this is a strange view of the nature of consciousness; but not on this account false. First of all it involves the view that the conscious self is suffering from an illusion if and when it holds that it can exert influence over the body. Secondly, it creates a tension between Huxley's essential scientific and naturalistic attitude and the conclusion to which he is forced, namely that consciousness, a function in a world of universal causal interaction, is a causally inert waste product of the animal machine. We have already pointed out that epiphenomenalism is a hybrid theory which results from the tension between a mechanistic experimentalism and the remnants of the introspectionist tradition of modern philosophy. It represents the clash of science and philosophy, not only in Huxley the scientist and the philosopher, but also in Western culture. The basis of the conflict goes as deep as the paradox underlying all knowledge--which is the paradox of the place of consciousness in the world and of the world in consciousness. Huxley, insofar as he was a philosopher as well as a scientist, stopped short of drawing the final conclusion from scientific experimentalism, which is a behaviouristic type of psychology that involves at least a methodological and occasionally even a complete disregard of the conscious self. From this point of view we have good reason to maintain that epiphenomenalism is a halfway house on the road to behaviourism.

A second tension that is now added is the following. Consciousness

is found to be a something (Huxley never defines it) which isn't really a thing at all. All things, units, entities in the natural world are, in at least one sense, active. Even an organic waste products, in their essential non-functional passivity, give rise to some series of events in the natural world. Consciousness gives rise to no events whatever in the world of nature; it is essentially, while other waste products are accidentally, inert. Its origin is in the natural world, but it cannot be fitted into that world---it is in its nature different from anything that is found in the universe of mechanico-causal interaction. It may be urged against Huxley that in the attempt to escape the difficulties of an interactionist theory he denied all causal efficacy to the conscious manifold and in doing this was forced to postulate something that in no way fitted the scientist's material universe. In effect his epiphenomenalism required the introduction of a whole new category of being into the minimum number of ontological ingredients of the real, and this requirement held for epiphenomenalism in general. It was Santayana who realized this. In the following two chapters we will proceed to consider how by means of his ontological distinctions we can eliminate the second tension implicit in Huxley's epiphenomenalism, while the rest of the thesis will be reserved for a consideration of whether the primary tension can be eliminated---whether or not epiphenomenalism is sufficient to explain and account for the basic problem of consciousness (and through that, of knowledge) and to resolve satisfactorily a conflict of science and philosophy.

CHAPTER 111

SANTAYANA'S FOUR REALMS OF BEING

Occam's razor, the famous maxim " entia non multiplicanda praeter necessitatem " was devised to limit theory to the irreducible minimum. Economy in ideas and beliefs is to a certain extent desirable, but it is also dangerous and easily carried to an extreme. It presupposes that the pattern of things is simpler than that of ideas, that nature at large is also economical instead of being superabundant in her manifold forms and luxurious vegetation. The rigorous application of the razor in the field of the theory of knowledge leads to phenomenalism and eventually to a solipsism of the present moment where nothing beyond the immediately given is admitted. " If God or nature had used Occam's razor and had hesitated to multiply beings without necessity, where should we be?" [1] Santayana asks. In themselves all things (the fantastic as well as the apparently sensible) are equally possible but equally unnecessary. If we want a philosophy based on the " large facts " which we know in the daily process of living, we should not be afraid to postulate or to recognize as many kinds of being as the solid body of our experience might suggest.

This preliminary remark, pointing to the virtual irrelevance of Occam's razor to an ontological analysis of experience will help us to

---

1. George Santayana: Soliloquies in England and Later Soliloquies (Constable & Co. London 1922), "Occam's Razor" p.197

understand Santayana's position with regard to the four realms of being which he distinguishes. I should say that in what follows I am only concerned with the later phase of Santayana's philosophical development. This frankly ontological phase was ushered in with the publication of Scepticism and Animal Faith in 1923, which was followed by the four volumes of Realms of Being.[2] It has been argued that the doctrine of the earlier period differs in no essential respect from the later doctrine. For our purpose this question is incidental. It is sufficient for us to note that Santayana himself considered the four books of Realms of Being superior to the five volumes of The Life of Reason.[3] He thought the former represented a clear exposition of his mature philosophy.

Santayana, then, claims that his philosophy is the product of a systematization of ordinary reflection; a revision and a clarification of the categories of common sense.[4] Since the question is that of distinguishing the edge of truth from the might of imagination we must adopt Descartes' method of universal doubt. We will doubt everything that can reasonably be doubted, we will separate the wheat from the chaff, weed our beliefs until we are left with nothing but the hard and precious kernel of residual certainty. Criticism and doubt result from a conflict of dogmas, Santayana remarks,[5] and it will be well for us to

2. A good example is: William Ray Dennes "Santayana's Materialism" in: The Philosophy of Santayana, The Library of Living Philosophers Series ed. by P.A.Schilpp (Tudor Publishing Co.Second Ed.New York 1951)pp.419-443.
3. See for example his statement to this effect in letter to Mrs.C.H.Tooy (March 28,1941); Daniel Cory: The Letters of George Santayana,(Scribner's: New York 1955) p.346.
4. George Santayana: The Realm of Spirit(hereafter referred to as "Spirit" (Scribner's: New York 1940),pp.272-3. Also: Scepticism and Animal Faith, (Dover Edition: New York 1955) (hereafter referred to as "Scepticism")pp.v-vi.
5. Scepticism p.8

remember this. To the animal in action it would never occur to doubt the reality of his object of intent, or the relevance of his knowledge. On the plane of action there is no room for doubt or disbelief; the former is essentially an intellectual exercise and as such irrelevant to action, and the latter, as we shall see, since it is concerned with denying the conditions under which alone it can exist, is intellectually dishonest.

In his quest for the absolutely certain Santayana's first step is to deny the transcendent character of knowledge. We are acquainted with appearances only, and appearances are modes of our sensibility. We can easily doubt the validity of any causal theory of perception, since in its essentials it involves an illegitimate inference from effect to cause. We can never have immediate experience of the object alleged to stand behind its manifestation in sense or thought. In this way we are reduced to the conscious manifold: the self witnessing a continuous motion-picture of appearances, which unfolds itself in a real sequence. Thus I can still remember past changes and anticipate future developments in the fortune of the actors.

If we now push our scepticism a step further, we can go on to doubt the veracity of memory and the relevance of expectation. My remembering something that has allegedly taken place in the past involves belief in an absent and postulated object just as much as my claim to know something involves belief in a postulated substance independent of

my knowledge and underlying the flux of appearances. The persuasive force of memory and of expectation is very great: I am certain that something really has taken place in the past, that something is just at the point of happening. Even the groundless and ultimate fact of experience (when I admit nothing else) is saturated with change. And yet when I reflect on the suasive value of memory instead of remembering, it is at once evident that I can find no certainty there. Any given memory or expectation is by definition a present datum. Given now it seems to refer to the not-now; and when we take it to refer to that we inevitably load it with interpretation. In itself, when we refuse to take it as a sign for something absent, it is in no way different from other present data.

When in this way we take all appearances at their face value, the ultimate position at which we arrive is a solipsism of the present moment. As instinctive belief in an environment beyond the given scene and in a past and future beyond the specious present is removed, sensible events lose the urgency of actual motions. The appearance of motion becomes the illusion of motion; change becomes a picture of change, events ideas of events. The whole field of appearances crystallizes into one dateless, motionless complex, beautiful and brilliant in its self-identity. The perception of change is domesticated into the same, non-temporal appearance of change. This is the point where the advance of scepticism is at last checked.

Now let us consider, with Santayana, some peculiarities of this ultimate residue---of what is left of our conscious experience.[6] The most important feature of this plenum of appearances is that it cannot be said to exist. Nothing given exists, in the sense in which "existence" is used " to designate such being as is in flux, determined by external relations and jostled by irrelevant events."[7] When data are taken for what they are and not for what they mean, they reveal to us Santayana's first realm of being, the realm of essence. Nothing existent can be given in intuition, and nothing which is given in intuition can exist. " Intuition," in this sense, is awareness of the non-transitive, for Santayana. The realm of essence is an infinite continuum, made up of every conceivable (and inconceivable) quality and property and relation. It includes all the qualities of being which may be actualized in any possible world together with all the qualities and characters which cannot. It is what the ancients understood by pure Being and it is logically prior to all other realms.

Essences enjoy the ontological, though neither the cosmological nor the moral privileges, of Platonic ideas. They are eternally subsisting self-identical and self-sufficient universalia ante rem. From the point of view of value essences are neutral; they have no magical power by which they could exert control over the natural world of flux in order to facilitate their own actualization by that flux. Once they have been actualized they do seem to entrench themselves and stubbornly resist any change, but this resistance originates in the flux of the realm of

6. In this early exposition I am roughly following the argument of the first chapters of Scepticism and Animal Faith.
7. Scepticism p.42

matter which has momentarily and in a certain locale assumed that form and now refuses to yield. The principle of essence is identity, just as the principle of matter is change. The realm of essense is made up of the sum total of mentionable objects, but we must immediately note that none of these is an object of belief or an object of knowledge. They become objects of belief when they are taken not simply as essences but as signs for the presence of substances. Consequently an essence has no location in physical space and no position in real time: it is eternal and dateless, though (on occasion) immediately given in intuition and as such the ultimate term in scepticism, and the bedrock of certainty.

But this certainty will not take us very far. If we do not go beyond the realm of essence we are condemned to a speechless and senseless gaping --- and even in this gaping there is more involved than merely the realm of essence: it takes a living intuition to envisage essences. Consequently we embark on a critical reconstruction of belief taking us from scepticism, which is an intellectual exercise, to animal faith, which is the unquestioning credulity of the animal in action. This transition involves the gradual discovery and subsequent investigation of the minimum number of assumptions we implicitly make in action and in the everyday business of living. Santayana introduces a radical disjunction between the contemplative and the active life of man and his epiphenomenalism is based on the belief that the former

has no relevance to the latter and the latter functions perfectly and as a closed system of habits without the need of any causal interaction from the former. My growing conscious of an object is incidental and dispensable for my handling that object. The terms we are concerned with in the sphere of conscious acquaintance (as contrasted with know-how) are intuitions, and essences envisaged which may or may not be taken for signs of things. The units in the field of action are bodies clashing in the dark, modifying, changing, eating and killing one another.

In this way the great postulate of animal faith is the existence of substance, or of substances forming an environment continous[11] with the animal's body. The system of substances in flux forms the natural world of which the animal is a member. Santayana is a realist in the sense that he maintains that the object of knowledge exists independently of knowledge. When I become conscious of a change, the first phase of that change has already taken place and thence passed out of existence independently of and prior to my cognition of it; and the last phase of that change has not yet taken place but is anticipated by me. Intuition synthesizes the piecemeal but constant flux of substance into a specious perspective; translates self-annihilating change into a changeless picture of change. We will speak more of this later.

The realm of matter, then, is the powerhouse of nature. As to what

matter is, Santayana gives us no answer: he maintains that that is a question for the physicist, though this does not mean that even science will ever be in a position to provide a final and in all respects satisfactory solution to the problem. But his hints as to what he understands by matter are sufficiently clear for a general appreciation of his theory. Matter is the generative order of nature; the source of all activity and power. It is what Whitehead would call " substantial activity "; the universal flux of substance continuous in its modes and measurable in its phases.

It is essentially a dark irrational force, a principle of arbitrary choice and random actualization. Random, that is, when we regard it from the point of view of the realm of essence, from the consideration of which alone we can assign no plausible reason why one form should be actualized rather than another, or all, or none. The realm of essence is an infinite plenum where no member is granted priviliged emphasis at the expense of the others and where each universal enjoys an identical amount of reality; the geometrical triangle as much as Plato's Good, the essence of amoeba as much as the color green. From the point of view of matter, however, the forces or activities progressively unfolding themselves are by no means irrational in the sense of being uncaused. The realm of matter is precisely the realm of dynamic causal interaction between substances. This realm is conceived by Santayana in strictly mechanical terms, as defined by the repetition of dominant tropes,[8]

8. G. Santayana: The Realm of Matter (Scribner's: New York 1930), p. 107. This book is hereafter referred to as "Matter".

where by a trope we understand the essence of an event under the form of eternity.[9]

Just as the realm of essence enjoys logical priority, the realm of matter enjoys generative priority. An essence would still be though it could not exist without matter, but the mode of being proper to matter is existence and matter could not exist without existing in some specific form and therefore relying on the realm of essence from where alone forms can be borrowed. However, when I speak of "priority" here, this is not to be understood in a temporal but only in a logical sense. Now I shall proceed to a brief consideration of the two secondary realms of being, namely those of spirit and truth. I must remark that there is a sense in which spirit can be considered primary and this is well expressed in Santayana's famous comparison of his ontology with the Christian doctrine of the Trinity of God as formulated in the Nicene Creed.[10] We only know through the spirit, and in this way the realm of the spirit may be considered primary in the order of knowledge. But knowledge itself, and hence the order of knowledge, is secondary in Santayana's estimation, in consequence of which, I believe we are safe if we speak of the realm of the spirit as also essentially secondary.

This view is further suported by the fact that in its existence spirit presupposes both material substance and essence. Substance is the organ while essence is the object of spirit. Substance generates the spirit

9. Ibid., p.102
10. Spirit , pp. 291 ff.

and this spirit only lives in the intuition of essences. Spirit is a category of being and hence not any individual being; it is "divided into spirits by its organs, and into intuitions by its occasions".[11] Generally speaking, spirit is an imperceptible and unsubstantial cognitive energy. An intuition which is a mode of spirit, may be roughly defined as an act of consciousness---a moment when the light of awareness falls on an essence: in sensation, thought, hallucination or even in a dream. I will deal with this at greater detail in the next chapter, and note some difficulties in connection with it; for my present purpose it is sufficient if we remember that intuition is the direct and obvious possession of the apparent and hence not identical either with perception (which also involves belief in a non-given object) or with the spontaneous activity of the fancy (than which it is wider.)

Spirit is Santayana's term for what had before him been called mind. He maintains that the natural link between matter and mind is that " bodily life should excite feeling, and that perception, emotion and thought should report material events."[12] Such a report in no way changes the system of habits that is the animal's psyche and hence causal influence between the psyche and the spirit (body and mind) is limited to a properly epiphenomenalistic one-way flow. Santayana insists that consciousness, mental discourse cannot be explained in terms of a mechanistic science, or behavioristic psychology (which is the only form of scientific psychology for him), because it belongs to a realm of

---

11. Spirit, p.253
12. Ibid., p.90

being altogether disparate from that of matter. Psychology is a part of biology and biology a part of physics---and physics reveals to us the foundations of things. The foundations of all things are in the material world, and in this sense an essence is not a thing.

In the case of spirit we have to distinguish between external conditions and inner nature. The external condition of spirit as manifested in a stream of intuitions is a certain harmony in substance and complexity of material organization--it is the psyche or inherited seed of life which placed in the proper circumstances establishes and maintains a certain specific form and inner bodily equilibrium. The psyche (though in its unity a mythological notion)[13] may be studied by science: the conditions of spirit are material. The inner nature of spirit, however, falls outside the scope of natural science altogether because it falls outside the scope of nature; it is " a view of the world which is not a part of the world".[14] This is what Santayana means by the "transcendental character" of knowledge. Though externally considered intuitions certainly exist, considered from the point of view of their internal moral (this word is used by Santayana mostly as the contrary of "material") character, they cannot, strictly speaking, be located in space and time. An intuition is a dateless transcendental station for viewing all things, though the viewing itself is always in terms of essences. This is the way in which, by means of Santayana's ontology, the epiphenomenalist can eliminate the paradox of having an essentially inefficacious by-product of animal life involved in the natural world of dynamic causal interaction.

---

13. Spirit, p.15
14. Matter, p.151

The realm of truth is a certain finite part of the infinite realm of essence. Since Santayana's theory of truth is incidental to his epiphenomenalism, I will treat it only sketchily before I pass on to other matters that are more important for my purpose.

In any theory of truth it is sound practice to distinguish two questions:

1. What is the nature of truth?

2. What are the criteria of truth?

The first question is a request for a definition or descriptive analysis of truth; the second is a question about criteria on the basis of which we can assert the truth or falsity of a proposition. Now for Santayana the truth about any fact is the standard comprehensive description of that fact. The standard comprehensive description of a fact includes every member of the realm of essence that is somehow relevant to that fact, and hence it is eternal and immutable. Generally speaking, we can define the realm of truth as a complete record of the impingement of existence (matter and spirit) on essence. As a fact, so-to-speak " drops through " the realm of essence [15] it leaves an irrevocable track which is eternally preserved. Thus the truth about a fact is an essence, or a group of essences: unchangeable and infinitely extended through the realm of pure Being. It is ideal in the sense that in its totality it can never be completely described, thought of or divined. Hence any proposition is true or false (in answer to the second question posed) only insofar as it repeats or contradicts a part of the standard comprehensive description of the fact in question. In the realm

---

15. Scepticism, p.267

of truth is laid up the complete history of the world, and it has been laid up there from all eternity---truth is frozen history, which has the peculiarity of possessing essential reality even before the events which it describes have actually taken place.

Now that we are acquainted with at least the basic outline of Santayana's ontology, it will be in order to attempt a further clarification of the nature and role of the different realms and of their interrelation. First of all I have to remark that these realms of being are not different cosmological regions. They are ontological ingredients into which the one existing world can be analyzed. In this sense Santayana's system is to be considered naturalistic; he admits only one world, the generative order of nature which assumes form (from the realm of essence), creates and sustains the life of spirit and finally by the selection of the forms it exemplifies, determines what the truth shall be. The difference between the three realms that are our primary concern can be brought out with clarity if we carefully distinguish three different meanings of the word "is"--- one each to correspond to essence, matter and spirit.[16]

The most basic and proper meaning of the word is identity and as we have already remarked identity is the principle of essence. In this sense "is" denotes the singular and exclusive identity of each distinguishable character with itself. In the realm of essence everything eternally is

16. The following exposition is based in its outlines on " Some Meanings of the word 'Is' ", in Santayana's Obiter Scripta (Scribner's; New York, 1936), pp. 189-212

what it is. Essence possessing primacy in the order of being gives us the intrinsic ideal possibility of all things; the flux of existence could never wear any determinate aspect or character if essences were not eternally available to characterize it, to differentiate its parts, to enable it to change from one form to another.

Contrasted with this self-identical reality of essence, we may use the word "is" to designate the dynamic flux of existence. To exist means to stand in external relations to miscellaneous events and things in the field of action: existence is persistence of substance through change. Existents <u>are</u> only insofar as they change: essences on the other hand cannot change but are <u>ex</u>changed for other essences each time observable change occurs. Nothing could be more unlike essence than matter, which is the existing potentiality of specific things---- the primordial flux charged with the power of creation. It is interesting to observe that according to Santayana we never intuit substance; the terms of conscious discourse are ideal and universal though its occasions are real and specific. Intuition is just that act of cognition in which the ultimate is immediate--- it is always the intuition of essence. Substance is not intuited but posited and therefore theoretically always problematic, though necessarily presupposed in practical life. More about this important topic will be said in connection with Santayana's theory of perception.

Finally, we can use the word "is" in the sense of actuality, which defines the kind of being proper to spirit. We have already distinguished between the external conditions and the inner constitutional peculiarity of spirit.[17] In accordance with this distinction we must maintain that the reality of spirit is essentially different from the reality of matter.[18] It has been argued that the category of spirit is derivative in the sense of being composed of activity of a high degree of complexity (psyche) and characteristics intuited, or, in other words, of the two primary realms of matter and essence. However, I feel that it would be well to distinguish here between the embodiment of essences by matter and their envisagement on its own plane by spirit. Santayana's position seems to be that matter when it reaches a certain complexity of organization breaks through into an altogether different realm of being, which though certainly dependent on matter and essence for its emergence, is yet on a different plane from that of the incessant flux which constantly selects and embodies a certain finite set of essences. The spiritual perspectives open at a right angle to the material flux which houses their organ;[19] they are acts of seeing, luminous and essence-directed. Intuition is less and more than natural existence; internally considered it is a synthetic vision of change which is neither in space nor in time. Externally considered, however, the intuition is chained to its organ and in that sense spatio-temporally located and exposed (mediately) to the vicissitudes of a natural existence. Consequently intuitions can be

---

17. Supra, p.54
18. T. G. Henderson: Essence, Matter and Nature in the Philosophy of Santayana, Harvard University Ph.D. thesis:1939; pp.47-57 esp. p. 56
19. George Santayana: " The Realm of Essence,"(Scribner's: New York 1927) p.134. Referred to as " Essence " hereafter.

ordered in the relation of contiguity or of succession only by their organs or by their existent objects of intent. The flux of nature is an infinitely divisible process; a moment of spirit, on the other hand, is an activity in the Aristotelian sense of the word--it is complete, self-contained and self-centred at the every moment of its existence and for that reason indivisible.

Thus, though it certainly seems to be the case that a union of substantial activity and complex organizational form underlie spirit as its organ, and the realm of essence stands open to it as its field of objects, we must nevertheless note that spirit is more than its organ and more than its objects, and more than its organ and objects combined, and for precisely this reason forms a third, emergent realm of being which though existent is addressed to essences, and though a product of flux, arrests the flux in a unitary vista, gives rise to the world of appearances and <u>is</u> in the intensity of an incandescent moral actuality.

I will now proceed to a discussion of Santayana's theory of perception which is well adapted for an expository juxtaposition of the realms of essence, matter and spirit. In order to be as concise as possible I am omitting all reference to the realm of truth, which would only introduce unnecessary complications into the argument without any compensatory advantages. We must immediately at the outset guard against committing the fallacy of equating intuition with thought or perception; it is both more and less than these. An intuition never reveals anything but

essences and no matter how these essences are revealed--by sensation, memory or hallucination--the act which is the condition of the possibility of their appearance is still the act of intuition. In a solipsistic specious present to which scepticism reduces our experience, nothing but essences are given, and these essences are taken for what they are: all meaning , all reference of the given to the not-given is absent. Even the act of intuition is absorbed in its objects, and if our animal haste did not force us to treat essences as signs we would never realize that appearances cannot be groundless: that essences have no inherent power of their own to appear and that consequently in rendering them present, an animal activity is inevitably involved. However, scepticism is not a form of life, for just this reason; namely our animal predicaments make it virtually impossible for us to disregard the urge to consider appearances as signs for the presence, desirable or dangerous, of substances. We cannot enjoy the aesthetic quality of the envisaged universals because the animal psyche, below the level of consciousness, is busy interpreting them, approving or rejecting them, judging them and devising ways for tackling them.

These are the two levels of human life: on the one hand pure intuition, the joy of a disinterested conscious life, opens for us a non-existent realm of dream images. On the other hand, these images are used by our animal nature to describe the dynamic, substantial and material entities against which we inevitably run up in the field of action. From this two consequences are seen to follow. First, the terms of knowledge

are symbolic. Secondly, no knowledge is indubitable.

Nothing but the immediately given is indubitable, and the immediate is always some essence. Santayana holds that a good many of the epistemological difficulties of past philosophy can be led back to the mistaken conception that knowledge is or ought to be intuition of fact.[20] Pure intuition never reveals any other relation beyond that of immediate self-identity---and such tautologies yield no knowledge. When the act of consciousness becomes " intuition " of fact, it is no longer intuition proper but perception, which is intuition tempered by animal faith. The universals which are the objects of intuition, are, in perception, charged with intent and projected on the appropriate locus in physical space to become a description of the thing encountered in action and defined by bodily attitude. In this way substance is essentially unknown; " the object of every practical perception is the thing-in-itself "[21] which is nothing other than the natural thing as it actually exists. In perception we utilize the visionary data of intuition to describe the objects on which our life as animals depends. Since we are describing substances in terms of essences, we cannot hope that perception will yield a pictorially exact copy of its object, and since we never actually go beyond intuition, there is always room for theoretical doubt that all our belief in substance might be nothing more than an elaborate and horrible illusion; scepticism may always drive us back to the specious present and remind us of the precariousness of all knowledge.

---

20. Scepticism, pp. 171ff.
21. Soliloquies in England and Later Soliloquies; " The Progress of Philosophy ", p. 213. This work is hereafter referred to as "Soliloquies".

Santayana's position is, in effect, that in an animal's life action is instinctive and primary. In action faith in the pre-existence of the objects of intent, as well as in the capacity of the animal to affect them, is implicated. Intuition which is the necessary (though not sufficient) condition of knowledge, and knowledge itself, and finally criticism, which is a revision and reconstruction of knowledge, are all secondary and incidental to the labor of life--to action in the dark.[22] A dynamic relation of interaction between my substantial self (psyche) and surrounding substances is prior to the cognitive relation between my conscious self and its object, the appearance of substance. Thus disbelief while life is going on is profoundly insincere; while a methodological doubt, since it is nothing but a reversion, on an intellectual plane, to the immediate and the absolutely certain, is not only useful but also necessary for philosophy, though scientifically barren.

We can advantageously formulate Santayana's theory of perception in terms of Meinong's famous distinction between the act, content and the object of thought; and this, I hope, will throw additional light on the realms of being here discussed. The act of perceiving is an instance of spirit: a living intuition. The content of the intuition is a finite group of essences. The object of the intuition is a substance encountered. An intuition immediately <u>expresses</u> and mediately <u>reports</u> substances.[23] It expresses the state of the psyche which generates it. On the other hand, in appropriate cases it also reports the existential presence of other

22. Scepticism, p.172
23. Soliloquies; " The Psyche", p.218.

substances in the environment, to which the body responds. Santayana did not explicitly distinguish between signs and symbols- he used the words interchangeably. In his sense, then, we can say that perception is a form of cognition of the absent; essentially sign-cognition inasmuch as it is intuition in which animal faith takes the essences revealed for signs of substances. Knowledge is true belief in symbolic terms presupposing antecedent action, or at least bodily reaction in the animal whose intuitions carry that knowledge, on the substantial natural thing, of which this true belief is the knowledge. Thus to know is a secondary activity hiding behind the over-powering primacy of action. This in turn logically implies that conscious knowledge is unnecessary for intelligent behavior, which is the basic postulate underlying epiphenomenalism. Accordingly, I pass on to a closer examination of Santayana's notion of spirit and subsequently of his version of the theory of conscious automatism.

CHAPTER IV

THE REALM OF SPIRIT

We have so far been concerned with sketching the general background of Santayana's philosophy, on which his epiphenomenalism is grafted. This was a task both necessary and useful. It was necessary in order to show how T. H. Huxley's epiphenomenalism could be made consistent if a new category of being were introduced--- a category which is ontologically different from natural things encountered in common everyday experience. It was also useful, because it gave us a summary idea of at least one type of philosophical system that can be built around or made consistent with the epiphenomenalistic approach to the relation of mind and body. We have seen that two important characteristics of this system are: (1) that it is naturalistic, and (2) that though insisting on a cosmological unity of ground, it refuses to reduce all reality, and more especially spirit, to a similarly unified ontological status. This chapter will be devoted to a closer investigation of Santayana's treatmentof spirit. I take this opportunity to remark that Santayana's style is in places highly ambiguous and metaphorical. His impressionistic writing makes a close study of the text extremely difficult to carry out and very often futile. His language is far from being precise, in consequence of which though his main views emerge with relative clarity, particular single passages occasionally seem to conflict and to require reconciliation which is not in all cases possible. We will say more

about this and discuss examples at a later point in this chapter.

There is a sense in which Santayana's theory of spirit is in line with a great philosophical tradition going back at least to Plato's Symposium, which considers the human mind the true mediator between the endless flux of matter and the divine self-identity of pure Being. Yet we must be careful here, for Santayana insists that he does not want to assign any moral or magical prerogatives to essences---- these are possible characteristics and nothing more: morally neutral though ontologically primary. Nor does Santayana consider matter in any sense base; on the contrary, it is the existent matrix, the necessary condition of the actualization of any emergent value. But when all this is said, it still remains evident that spirit is a mediator (in the sense of a mean) between the realm of matter and the realm of essence. For though a product of material causes, it is a leaping flame which boldly disregards its conditions and strives for virtual union with the ultimately real, its Will being to know and to love all essences.

Santayana distinguishes between the material and the immaterial, the existent and the non-existent. These two pairs of contraries, in different combinations, spell out for us the realms of being with which we are concerned. Materiality implies existence and non-existence implies immateriality. The former defines the realm of matter, the latter the realm of essence. Spirit, however, introduces a new combination

of the contraries: it is both immaterial and existent. Since spirit, as the intuition of essence, is the counterpart of pure Being, it is immaterial, pure and disinterested. However, since in the long run an instance of spirit is a biologico-natural event (considered from the point of view of the parent organism), it is existent and not immortal as a disembodied spirit (if such could exist) would be. The price to be paid for existence is limitation and at times the human spirit is intensely conscious of these limitations. By its origin it exists and is a natural product; by vocation and inclination it is addressed to the whole realm of essence and hence is immaterial and "supernatural". Though a flux of existence underlies its production, spirit itself is not in flux: it is an act of cognitive synthesis. If matter is the existential flux of a multiplicity of substances in time with external relations holding between part and part in space, and essence the self-identical being of eternal characteristics in internal dialectical juxtaposition to one another, then the actuality of spirit is where existence approaches essence, for it is flux arrested by the synthesis of intuition in a unitary, specious vista.

By the very act of observing the flux of nature spirit escapes from that all-pervasive change---becomes the unchanging knowledge of change. In this way it transmutes and redeems the self-stultifying senseless rush of existence and confers value on it. It is something like a dividend which accrues to matter when a certain complexity of organization is reached; a beautiful flowering superadded as the gift of maturity.

Existence unfolding itself is a dramatic spectacle for the spirit ---a spectacle in which the spirit as spirit is not involved. Pure intuition (which is what spirit would be if it were free) is by nature disinterested, happy and contented. It sees everything in perspective: impartially as one examines chronicles of ancient wars. In our spiritual dimension we see everything sub specie aeternitatis: in terms of essences or, as Santayana often puts it, as episodes in another man's life. Thus matter has produced spirit which exists in a moral dimension of its own and is addressed to the realm of essence and to that alone. Spirit has no preferences, no idiosyncrasies: it is never a principle of choice. William James wrote an article entitled " Are We Automata?"[1] in an attempt to refute Huxley and Clifford who had just then expounded their epiphenomenalism. James argued that consciousness was essentially a principle of choice, of selective attention and discrimination.[2] Santayana, in contrast to this, maintains that the direction of mental discourse could never be determined by consciousness alone: the concept of direction is meaningful if and only if we assume antecedently determined ends. These ends are defined, when immediate by the fortunes and needs of the organism at a particular time, and when long-range by the single aim of self-preservation, or the preservation of the equilibrium of the inner environment of the body. This, by definition, is the task of the animal psyche, which is a system of habits; the specific form of bodily life, the sum total of all vital mechanisms. This material organ or system of organs, planning, selecting,

---

1. Article " Are We Automata?" by William James in MIND Vol.IV,1879,pp.1-22.
2. " Are We Automata?",pp.8ff.

distinguishing, acting from her central Operations-Room in the body, is the principle of choice. Matter has no ideals, but it has a certain impenetrable stubbornness[n/] to keep doing what it is doing--to retain past rhythms. It is probably this propensity to repeat established patterns that defines the task of self-maintenance for the psyche. At any rate, Santayana emphatically refuses to admit the possibility of any mental machinery.[3] Intuitions are not animals, they cannot breed other intuitions; and since an intuition cannot exist without an object, it cannot choose its own object for that would imply that it is, from at least one point of view, causa sui. " Intellect exists by intellection"[4]; it is a moment of spirit created and does not on any count pre-exist. Santayana's main criticism of James would, I think, be that he hypostatized consciousness into a spiritual substance (a contradiction in terms, for Santayana) which could exist independently of its instances and choose its own objects.

Santayana is intensely conscious of the fact that common sense, the implicit assumptions of which he claims to extract, investigate and systematize, conceives of spirit not at all as a final fruition of life. Accordingly, he distinguishes two meanings of "spirit"--one a creative wind or breath, a subte physical influence shaping the world, the other a moral witness and victim of existence.[5] The former conception is superstitious: a remnant of magic which attributes power to appearances and turns the pure spirituality of consciousness into a source of energy, which it is not. Substance is indifferent to cognition in the sense

---

3. Matter, p.140
4. Spirit, p.100
5. Spirit, pp.280ff. Also G.Santayana Platonism and the Spiritual Life (Scribner's:New York 1927), pp.47 ff.

that my knowing about something will in no way change that thing. This is recognized by common sense today, though it was not a few hundred years ago. Beliefs about the magical efficacy of the evil eye are a case in point---and books on anthropology abound in other examples. However, Santayana wants to go a step further and maintain that it is not only with respect to ambient substances that consciousness is inefficacious; it is powerless to act on the substance of its originating psyche as well. To think that ideas or ideals can act on or in any way transform the environment is sheer mythology. It involves a confusion of realms of being, the fallacy of substantializing the unsubstantial. It is the confusion of the two natures of man: and now we see that epiphenomenalism involves the conception that the human self exists on two levels. Man is a rational animal, Aristotle said--a compound: one person with two natures. This, too, is what the title of the book "Scepticism and Animal Faith" reveals; scepticism as the function of the cognitive, rational part of man, leading us back to the bedrock of certainty and the "supernatural" beauty of intuition; and animal faith--the investigation of the lower, substantial self (naturally as revealed in intuition, because consciousness is the necessary condition of any knowledge): a systematic expostion[1/] of beliefs implicated in action.

On the level of action man is an animal in dynamic interrelation with things and other animals. On the level of contemplation man is a transcendental centre for viewing all thing. "A man habitually identifies himself as much with his body as with his spirit: and since both are called "I", it is no wonder if what happens in each is felt to be also

the work of the other."[6] However, we should take good care to distinguish the two selves: they belong to two altogether different realms of being. As body or psyche, I am a part of the flux of nature: as spirit I stand above that flux and know it by synthesizing it in my intuitions. But here further caution is requisite, because the fact that as spirit I feel myself essentially elevated above matter to the plane of disinterested observation may suggest to me the treacherous notion of a disembodied spirit. Santayana in touching on this subject never neglects to emphasize that a disembodied spirit could not exist, would have no support, no theme, no foothold anywhere in the vast stretches of essense confronting it. A disembodied spirit, without an innate principle of choice (which, as will be recalled, depends on antecedently set goals), would have to consider the whole infinite realm of essence at the same time---an admittedly monstrous task. Spirit can only live and know if it is lodged in a particular body; the spark of intuition must reveal the world, light up a part of the realm of essence from a particular contingent point of view; and though this perspective may be considered detrimental and even violent by the spirit which strives to know and love all things indiscriminately, some perspective and some particular body is absolutely necessary for spirit to exist at all. And if spirit did not exist, it would just evaporate into identity with the realm of essence.

The illusion that my willing (a mental act) can influence the

---

6. Spirit, p.10

material environment through the activation of the body's motor mechanism is based on a harmony of spirit and matter.[7] This harmony is evidently grounded on the fact that it was the material environment with all its rich potentialities that occasioned the conscious act of "willing" (an emotion expressive of the psyche's condition). The spirit, being a product of immediately antecedent material involutions, finds itself moving in harmony with the further development of its generative substance. The observation of this harmony is interpreted by the spirit in its self-centred boldness as a sure sign of its mastery over all things. However, in order that the spirit be liberated from the anxiety of an animal life we must realize that its domination over nature is not physical but moral, not materially effective but intellectually redemptive.

We have now seen another point of view from which Santayana may be considered as belonging to an old philosophical tradition: he retains the radical disjunction between mind and matter introduced by Descartes, and his conception of the self as existing both on an active and on a contemplative plane goes back at least to Aristotle. The essential separation of spirit from psyche is brought out clearly in Santayana's distinction between behavioristic and literary psychology. The former is a science, and as we previously pointed out,[8] a part of biology. The latter is an art[9] and as such only an exercise of dramatic insight. Science relies on observation and experiment, and the life of the psyche

---

7. I am here developing Santayana's suggestion on p.214 of The Realm of Spirit. See also Spirit, pp.80-82, 162.
8. Supra p. 54.
9. Scepticism, p.252

is in principle completely observable. Mental discourse, on the other hand, is invisible and intangible: it is essentially immaterial and elusive--it can never be an object of intuition. Consequently it cannot be treated by physical science; spirit is not only ontological but also biological overflow. This moral dimension is mechanically non-existent[10]; on this point there seems to be an important divergence between Huxley and Santayana. Huxley hoped that science would one day discover the mechanical equivalent of consciousness just as it discovered the mechanical equivalent of heat.[11] Santayana, on the other hand maintains that as the mechanical equivalent of spirit we could at best find an organic sensitivity and responsiveness in the psyche; a heed of external relations which lies at the basis of the original evocation of conscious life.[12] I postpone the closer examination of this significant difference between the views of these two main exponents of epiphenomenalism until later. The question more properly concerns Santayana's attitude to science in general and so it will be considered at the end of this chapter.

Now, since spirit cannot be observed as an essence is observed, nor encountered as a substance is encountered, the study of animation in nature and even of my own past animation is transferred into the sphere of literary psychology. Spirit can only be enacted and since this act of consciousness cannot observe itself, no observation can in any way be immediately relevant to a study of intuition. As Santayana puts it,

---

10. Spirit, p.59
11. See, Supra p. 20-30.
12. Spirit, pp.59,106; Matter p.152; Soliloquies, pp.222-3

intuition is"not a link in any traceable process",[13] it is created and dissolved from moment to moment on a plane inaccessible to the senses. In the science of nature knowledge of things is mediated by essences, whereas in literary psychology the process is reversed (and thereby made more uncertain) as substance mediates the knowledge of essences present to a posited intuition.

[14] We have previously remarked that epiphenomenalism is a halfway house on the road to behaviorism. Now in Santayana we can see this suggestion thoroughly confirmed. He gives free reign to behaviorism and the science of nature, in the sphere of nature; but with the same forceful gesture he draws a borderline between matter and spirit, reserving for the latter an autonomous poetic function as a transcedental centre for viewing all things. By the very act of handing over the material sphere to an essentially mechanical science, he confines the advance of that science in an effort to salvage and establish a sanctuary for some of the moral values that have, since Descartes' time, been on the defensive, constantly losing ground and now finally within sight of land's end where they would be pushed into the sea. Santayana's solution of the problem is a compromise, sacrificing power in order to stay alive. The spirit has to renounce all claims to physical efficacy, all hope of directing its own life. Proper respect has to be paid to power; the spirit's very nature is to be humble, to practice piety and foster the inner treasures of a contemplative life.

---

13. Scepticism, p. 246.
14. Supra , p. 42.

Spirit is in its essence contemplative, Santayana maintains,[15] but the psyche engaged in action inevitably imposes animal faith, faith in the substantiality of things presumably underlying its visionary data, on the spirit. This results in the illusion of the inner man that it is the spirit that trades the blow in the heat of action and not the physically competent animal psyche which is continuous with its material environment. The impartial integrity of the spirit is violated in subjecting it to the anxieties of animal life. In this way when it is distracted the spirit is involved in a permanent paradox, for it tries to escape from the miscellaneous predicaments of its parent animal and pass into its proper essence-directed contemplative stage, whereas disregarding in this way the body, it is in effect rejecting the condition of its own existence, on the material of which it parasitically feeds.

By distraction Santayana means contrary and inescapable commitments, and of these he distinguishes three: distraction by the Flesh, the World and the Devil. The nerve of all distraction is that the spirit is burdened with issues alien to its vocation which is a super-material absorption in beauty. The most important and consequently most dangerous form of distraction is that by the Devil, by Mephistopheles or by Lucifer, as Santayana puts it in his picturesque manner. Here spirit rebels against the sources of spirit; sensing the oppression of material circumstances it either assumes the servility

---

15. Matter, p.158

of Mephistopheles yielding with resignation to all seduction, or the proud rebellion of Lucifer posing as magically omnipotent. The conflict underlying all the misfortunes of the spirit is that it has so far failed to recognize its proper nature and function and attempted the impossible. The condition of liberation is self-knowledge: the realization that the spirit is a transcendental witness of the spectacle of life, that though physically impotent, it is the source of all value and in virtue of its intelligence enjoys moral autonomy. Thus the liberation of the conscious life from the torment of distraction is dependent upon the recognition of its epiphenomenal character. The Will in spirit ("Will" is used by Santayana to mean the " observable endeavor to develop and preserve a specific form")[16] is precisely not to will, but to understand the lure and the sorrow in all willing.[17] By understanding the spirit transmutes its own suffering and prepares the way for union (based on sympathy) with the Good, and the love (based on charity) of the world.

On the one hand, spirit must be resigned that it cannot change the world: action should be left to the animal in us. This renunciation, Santayana tells us, is so much the easier because all power is physical and all love of power is psychical and not spiritual. On the other hand, however, by this voluntary renunciation of all power (a sphere not proper to the spirit), it attains to happiness in

---

16. Spirit, p.53
17. Ibid., p.177

prayer, love and charity; the virtues of the inner life. The psyche is concerned with finding her way around in a hostile world: with gathering food and keeping alive and reproducing. The spirit, in contrast with this, has no affinity with the will of mother Psyche: it is disinterested, has no preferences and abhors violence. And yet it must listen to the anxieties of the psyche, suffer with her in all her changes of fortune, be a running commentary in conscious terms on the vicissitudes of animal life. To have to care, to be forced to serve---this is the felt tragedy of the spirit. Santayana is not clear about how spirit (if it is inefficacious) could be made to serve when it can only observe. I presume he means the spirit feels it has to serve: it is constantly overwhelmed by the load of animal faith forced on it by the anxious expectancy of the psyche. If I interpret him correctly,his point is that distraction originates in just this felt necessity of the spirit to serve, which however is not its proper function. The result is dissatisfaction, and eventually the martyrdom of the spirit. For sensing its potential perfection in pure intuition, it can gain no repose in the subservience of its substance-directed stage.

The diagnosis of Santayana's epiphenomenalism is thus as follows. It is a calculated retreat into the inner conscious life of spirit -- a quiet submission of the mind to force, in its physical impotence only in order to affirm so much more forcefully its absolute moral independence. I do not presume that this theory was intended as a reaction against Hegelianism, but at any rate it seems clear that it amounts

to an explicit contradiction of Hegel's notion that reason governs the world. Reason cannot govern the world: it cannot even control the body in which it is lodged. It cannot direct the course of historical development, either as a grand impersonal principle, or as finite human rationality working piecemeal in the natural world. It is altogether impotent and its function is to find itself, approach the ideal of pure intuition as close as humanly possible and live a religious life of worship, of love and of charity. " If spirit were a power, its first concern would be indeed to reform this world"[18], Santayana admits, but it is evident that the world cannot be re-formed along moral lines. It is irrational to its very core, taking its own line of development quite apart from the moral fiction of intentions and rational purposes. " If any idea or axiom were really a priori or spontaneous in the human mind, it would be infinitely improbable that it should apply to the facts of nature."[19] The spirit has to renounce every form of the obsession of claim and of possession and concentrate on a detached contemplation of all things under the form of eternity. Spiritual equilibrium is reached when one's past (and even intended future!) actions are seen in a historical perspective, free from anxiety and remorse and hope; in general free from the rumble of the universal flux, treasured for what they intrinsically are--crystallized eternal essences. The spirit has a propensity to greet everything with equal enthusiasm: it is this power of appreciation that has to be developed in us and

---

18. Spirit, pp. 225-26
19. Scepticism, p. 289

freed from the psyche's persistent partiality. The intensity and scope of this "moral illumination" depend on physical circumstances, but the spirit, once liberated, accepts whatever is offered as one accepts free entertainment, and asks for no more: and when nothing more is offered and the light of intuition is extinguished with the parent psyche by the groundswell of physical forces, it is perfectly content to accept the inevitable and wait for another psyche to actualize it if it ever will be actualized again.

This leads me to the consideration of an ultimate implication of Santayana's moral idealism.[20] " We talk of 'life' ", he says, " as if it were unquestionably something precious or even divine. Perhaps... the vocation of the spirit may be to overcome this prejudice." This statement seems to me to carry the force of a revelation. It uncovers the essential direction of the development of Santayana's system. We have noted that the theory of epiphenomenalism was put forward as an attempted solution of difficulties which had arisen from a conflict of science and philosophy. Methodologically the conflict was one between observation and experiment on the one hand, and introspection and empathy, on the other. From the point of view of achievements, scientific experimentalism resulted in conrete discoveries facilitating a progressively wider and more efficient control over the environment. In contrast with this, introspectionist psychology did not in its achievements go beyond the development of a series of miscellaneous, and for the

---

20. Spirit, p.61

most part contradictory, theories. At best it led to the self-control and the inner self-contentment of the philosopher concerned with such introspection, and this contrast of the individual's intellectual self-control with public material control over the environment is one which we shall have to keep in mind. I am not at present concerned with tracing or evaluating these two contrasting trends of development; we shall talk about them at greater length later on. Here I am only concerned with pointing out that there is a conflict underlying the problem which epiphenomenalism attempts to resolve, and that this ultimate development of Santayana's philosophy seems to be directed not at the solution but in the long run at the rejection both of the conflict and of the problem. No conflict can be resolved by the radicalism of rejecting life altogether. Socrates once said[21] that the philosopher is constantly in the pursuit of death. I suspect he may have said this with tongue in cheek, but even if he did not that statement can be understood and appreciated on the basis both of old Pythagorean beliefs concerning the debasement of the soul in its union with the body, and of Plato's own theory about the immortality of the human soul. However, in the case of Santayana, who does not belive that anything is immortal,[22] the strange view that philosophy is a preparation for death instead of a discipline enabling us to live a better life, is much harder to account for or to understand.

I do not want to overstress this point--and similarly I do not

---

21. Plato: Phaedo, Stephanus, p.64
22. Scepticism, p.271

want to use it as a theoretical argument against Santayana's system, for the simple reason that it isn't one. But I do believe that considerations concerning the existential implications of philosophical theories (especially theories in the moral sphere, or "lay religions" as Santayana calls them) are important and revealing. At any rate it seems clear that Santayana pays a great price for saving the life of the inner man and the aesthetic and moral values that go with it--perhaps too great a price. The sphere of action is completely sacrificed, the spirit is made completely dependent. It is reduced to the status of a childish poet indulging in its fancies; an innocent bystander bubbling with the joy of life and yet essentially tragic, inasmuch as it is abused and compelled to swear allegiance to causes it knows and cares nothing about; to the powers that be. In places the spirit is even equated with the imagination,[23] moving on a level higher than the material flux and transmuting that flux ideally though sharing it existentially. The spirit is removed from the battlefield of competitive living to a contemplative distance within the inner man, where it matters little what we think, since a thought can never be the cause of an action---and in the substantial flux of nature only the body's activity can in any way affect or redistribute the energy-patterns prevalent at any one particular time. In this way the spirit teaches us to renounce life or at least shows us the meaning of sacrifice and through sacrifice, of liberation. While animal faith is directed at the realm of matter, pure intuition is directed at the realm of essence and as such knows no rules of

---

23. Spirit, p.161

preference. But we must also remember that essences are not food for the animal and hence could never have developed, just as they cannot sustain, life which is a necessary condition of the existence of the spirit.

Resignation, in its manifold forms, is the feeling that permeates Santayana's philosophy of spirit. Introspection is stripped of its scientific pretensions and admitted into philosophy as a poetic exercise of the fancy, the dramatic art of empathy only. Santayana's hope is that by his distinction between the realms of matter and spirit, literary psychology, however far scientific psychology may push it back, will always remain in possession of the moral sphere.[24] Thus, in a sense, the vicious aspect of epiphenomenalism is eliminated; for consciousness is no longer held to be a waste product. From the point of view of material efficacy, it is indeed a waste: it is pure biological overflow. But now a second aspect of this same single fact of the material production of an inefficacious by-product is added. From its own point of view, spirit, which is accustomed to disregard the conditions of its existence, is the result not of a biological overflow but of divine incarnation. There is a striking analogy between the duality in unity (the rational and the active selves) in man and the incarnation of the divine in animal form in the person of Christ.[25] In this way the spirit becomes the source of value

---

24. Scepticism, p.252

25. George Santayana: The Idea of Christ in the Gospels, or God in Man (Scribner's: New York, 1946) pp.227 ff. Hereafter referred to as "Idea"

in the world: the precious actuality which natural processes conspire to create, and which though it cannot react back upon its source (a first sensation is, physiologically, a last event!) observes that themeless flux, comments on it and in a synthetic picture preserves it. Thus we have to distinguish, with Santayana, at least two types of function and of activity, one of which is non-causal. It would be false to say that the spirit has no function or is completely inert, unless we mean by "function" causally effective performance and by "inert" the incapacity for causal functioning. The Aristotelian meaning of the word "function" would be more proper here, since its connotations are not necessarily causal (the notion of cause is restricted by Santayana to Aristotle's efficient cause). Thus the spirit being the first non-material actualization of form can be quite properly said to have a "spiritual" function--that of transcending the flux by raising it to the specious unity of knowledge.

I will now revert from this generalized discussion to the consideration of some more specific issues in connection with Santayana's epiphenomenalism. I shall begin with some observations on his notion of the relation of mind and body.

"Such is the natural link between matter and mind, that bodily life should excite feeling, and that perception, emotion and thought should report material events" says Santayana,[26] and this seems to be a good summary of his thesis. Mind is a successive hypostatization

---

26. Spirit, p.90

of the tensions of animal life into spiritual terms; it is a transcript of physical changes.[27] The mind at any one moment is expressive of the condition of the psyche at the immediately preceding moment: hence consciousness is a function of the organism, on which it depends for its evocation and continued support from moment to moment. Aristotle held that the intellect had no special organ and hence it came to the body from the outside. Santayana, on the other hand,[28] maintains that the whole man is the organ of his intelligence, and it is for precisely this reason that the mind is not an immortal stranger in the body, but the son of the house (though admittedly strange in his habits). Mind enhoes the life of nature and presupposes a complex order of animal life to generate it. Its generation itself seems to be conceived by Santayana in the following manner. An animal in the natural world is completely dependent for its survival on the environment. Food must be found and ambient destructive forces must be averted. The thrust of the environment has to be parried and in order to do so the psyche, alternately defensive and aggressive, weaves a web of sense organs around herself, acquiring a physical sensibility to outlying things and distant events. By this sensibility of the self to the not-self the psyche has set the stage for the appearance of spirit. When physical sensibility reahhes a high degree of articulation consciousness is generated "automatically";[29] though not by physical derivation. Reflex actions evoke images,

---

27. Essence, p.x
28. Idea, p.235
29. Spirit, p.59

voices acquire meanings; we grow conscious of stimuli and of our reactions to stimuli. Our growing consciousness of material events is not an added fact in the physical world; it is an added cognitive fact, which is on a different plane from that of its objects, as well as from that of its terms. When Bacon said that knowledge is power, if by knowledge he meant cognitive awareness, he was confusing two distinct realms of being. Substances are independent of knowledge, though they are capable of being affected by action. The psyche is a mode of substance---and since in the natural sphere one motion can only be traced back to another motion, consciousness can in no way affect the psyche's habitual,automatic functioning.

Santayana does not always seem to be certain and unequivocal on this point, though it is undoubtedly the crucial issue in determining whether a theory of mind is epiphenomenalistic or not. Consider this statement: " We have a more compact mind and a stronger will when the themes of our intuitions are relevant to our action."[30] The meaning of the sentence hinges on what we understand by the word "relevant" ---relevant as knowledge <u>about</u> an action, or as knowledge leading to a more efficient execution of that action. I do not have to point out that the latter is the interpretation which would be generally accepted by laymen as well as by most philosophers. However, for the sake of fairness I admit that by and large Santayana's position emerges with relative clarity.[31] He insists that the psyche for her

---

30. <u>Spirit</u>,p.93
31. <u>Platonism</u>, p.48; <u>Spirit</u>,p.59; <u>Soliloquies</u>,p.223 etc.

proper functioning has no need of the spirit, and even if it would, consciousness could not effectively redistribute the physical flow of energy. We can formulate the question in the following manner. If the spirit is only a stream of intuitions superadded to the teachably reflex machinery of the body and can in no way affect or modify the functioning of that machinery, would the psyche continue to act with the same intelligence and effectiveness if the spirit were removed? This would seem to be a good way of testing whether consciousness is truly epiphenomenal or not; for if the unconscious psyche would be erratic or deficient in its functioning, we could no longer maintain that the spirit is a mere inert spectator of the dance of life and of action in the dark.

However, the question is not as easily decided as it would seem at first sight. Surprisingly, Santayana does not hold that the functioning of the psyche would be the same with the spirit absent as it is with the spirit present.[32] Evidently we do things when we are conscious which we would or could never do unconsciously. But, Santayana adds, this does not imply that it is the causally efficacious presence of consciousness that makes the difference. Rather, the absence of consciousness is expressive of the fact that the physical organs (or functions) whose concomitant effect it is, have for some reason not been properly activated. These physical organs are certainly efficacious----and if they do not function (as in sleep, for example) the observable behavioral difference in the animal is

---

32. Spirit, p.281

considerable. But it has to be emphasized that this difference is due to bodily organs and not to the spirit---the absence of consciousness is only a sympton of suspended activity, and in no way the cause of it. The psyche functions automatically, reacts to stimuli received through the senses, and shapes and changes the world in accordance with her interests through her impulses. That consciousness accompanies these furious activities is incidental to the activities themselves as well as to the interests they subserve.

This leads me to another apparent ambiguity in Santayana's theory of mind. So far we have used the words "spirit" and "consciousness" interchangeably, and a good case could be made for the view that this actually is the way in which Santayana uses the terms. However, consider this statement: "....I shall find it impossible, when I come to consider the realm of spirit, to identify spirit with simple awareness, or with consciousness..."[33] A few pages before this[34] he says "....even if spirit mean nothing but pure consciousness..." implying that it does mean more than that. On the other hand in The Realm of Spirit he identifies in no uncertain terms intuition (a moment of spirit)[35] with consciousness, ("....spirit...might be called consciousness...")[36] going even so far as to maintain that literally the phrase "unconscious mind" is a contradiction in terms.[37] Once again, his position seems to be clear in outline, though (possibly on account

---

33. Scepticism, p.275
34. Ibid., p.272
35. Spirit, pp.59, 93, 106, 281, etc.
36. Ibid., p.vii
37. Ibid., pp.37-38

of his highly poetic style of writing) lacking in precision. Consciousness is the mental activity par excellence for the epiphenomenalist; there is no clearer statement of this than in a hint of Santayana's[38] to the effect that the objects of the mind are not a part of the mind. "Mind", strictly speaking is nothing but pure activity, non-causal of course---the activity of consciousness. But Santayana, if I understand him correctly, introduces a distinction between pure consciousness (or, passive intuition) and consciousness as it actually is in everyday experience.

Pure consciousness, which Santayana sometimes calls pure intuition, is something like the act of envisagement implied by the presence of essences. It is the self-transparent activity of the transcendental ego, involved in the specious present to which the sceptic is reduced. In it appearances are taken at face value; meaning and interpretation are absent. A word pronounced, for example, is taken for the sound it is---and written, for the graphic shapes making it up on paper. Such reduction of consciousness to its purest form, however, is essentially an abstraction; it can be approached and in a supreme moment even possibly attained, but it can never be sustained. Consciousness, as we actually find it is charged with intent, thickened into belief; it is heavy with expectation and the psyche's urge to take the given for a sign of the not-given. In belief and understanding, mind is still nothing more than the activity of consciousness; but now this activity is considerably thickened--into

---

38. Scepticism, p.179

memory, expectation, perception; in short, into experience. This distinction is made by Santayana explicitly, though not in all the detail one would like, in the last chapter of The Realm of Spirit. " If spirit, taken abstractly, might embrane all essence impartially...spirit in the concrete, as it actually exists, is directed upon order, and upon a definite and selected order, beyond which it is swamped, lost, tortured and maddened."[39] Spirituality consists in approaching this pure form of consciousness as closely as it is possible for a human being, by resigning the psyche's interests and compensating for the spirit's violent perspective (from the body as centre) by a universal readiness to know and to delight equally in everything that may be presented. If spirit were free to develop without being thwarted by the constant solicitation of its attention to the miscellaneous issues of animal life and fortune, it would be pure intuition--happy in itself and pledged to nothing further, never obsessed with fear about the not-given. Spirit functioning with ideal perfection would thus be manifested in a stream of pure intuitions.

There are, then, two forms of spirit; one infected by anxiety and animal faith, the other pure, concerned only with the intuition of essences. The degree of perfection a particular spirit achieves is judged by determining how closely it approaches its ideal of disinterested speculation, pious love and eternal joy. If I understand Santayana correctly, this ideal can never be fully realized in practice:

---

39. Spirit, p.298

but it is the end proper to spirit, in the direction of which it naturally tends. If the spirit, being merely the animal psyche grown conscious, would become completely disinterested and sacrificial, this state (implying suspension of the psyche's life-sustaining activity) lasting for a moment, would be equivalent to physical death. However, a greater or lesser degree of spirituality can be reached while physical life goes on without the final disaster of death, and the road towards this liberation of spirit from anxiety leads through, first, the resignation of physical interests and second, the recognition that even the order of logical discourse or of music, even the spirit's wish to love and to understand all things is based on the prim functioning of the psyche. The end result can be compared with the Nirvana of Indian philosophy: it is a form of self-forgetfulness in which intuition is absorbed in its object, like a poet in the contemplation of beauty, or a child in the enjoyment of carefree play. Santayana holds that both the existential and the essential forms of spirit are epiphenomenal, the distinction between them being, from this point of view, that while the former may still have magico-superstitious notions about the physical efficacy of ideas or of mental activity, the approach to the ideal limit of the latter involves the realization and acceptance by spirit of its own physical impotence and the consequent acquisition of intellectual dominion over all nature. Even though the influence of animal faith " the toys of sense become the currency of commerce "[40] such vehicular use of the immediate for

---

40. Scepticism, p.82

purposes of descriptive symbolization is still without any effective relevance to the world of action or to the substantial objects so described.

Let me now briefly consider the conflict of will which subsists between the psyche and the spirit. Will, as we shall recall, is used by Santayana to denote " the observable endeavor in things of any sort to develop a specific form and to preserve it."[41] Now the will of the human psyche is to organize matter into the specific form of the human body and then to safeguard at all costs the inner bodily equilibrium, which preserves that form. The will of spirit, on the other hand, is precisely the opposite: its essence is to discredit all bodily interests---occasionally even to regard them as the greatest evil standing in the way of the proper development of its nature. But now it is evident that spirit is nothing but the psyche become conscious: a new moral dimension added to the machinery of the body. Any given change in the spirit can be led back to its cause----a change in the psyche. Thus in order that the spirit be disinterested, the psyche also has to be, at least to some extent, disinterested, and that goes against her nature: in the long run it may even prove to be suicidal. Thus there is first a conflict between the will of spirit and the will of the psyche; and then this conflict is led back to one within the psyche herself.

---

41. Spirit, p.53

Santayana recognizes the problem and devotes a short remark to it in The Realm of Spirit.[42] It is evident that there is a conflict between spirit and psyche, and that this conflict is in its origin internal to the psyche. Hence it follows that the liberation of the spirit, involving at least a partial satisfaction of its will, is primarily a shift within the psyche.[43] These Santayana considers facts; and from his general attitude it seems clear that he holds it is not impossible that the evocation of spirit, expressing an organic change within the psyche, may finally prove a fatal mutation for that psyche. However, he believes that even though his system points to such a self-contradiction within the psyche, the two conclusions an antagonist might want to draw do not follow.

1. It is not the case that a self-contradiction in the psyche necessarily reveals a self-contradiction in the theory which postulates it: since the psyche is the form of a changing natural organism that fights and precariously survives in a dangerously unstable environment, she is certainly liable to err, or at least to explore possibilities of development which lead to disaster.

2. It is not the case that in generating spirit the psyche was false to her own interests. If an antagonist could prove the opposite, suspicion would once again fall on the theory that postulated such a bungling natural organism, constantly contradicting itself, ruthlessly

---

42. p.64. In what follows I am elaborating this remark.
43. Spirit, p.208

undoing in a moment of frenzy the labor of centuries. However, notions such as " to be false to one's interest " are moral fictions in terms of which the spirit conceives of the natural world; in the material flux itself no such conceptions are operative. Matter is concerned only with new avenues of exploration or with new adventures embarked upon and new experiments tried. It does not hesitate to contradict or to annihilate itself in its modes: it does so all the time. Nothing generated is immortal---and it may well be that man having reached the only source of value in the actuality of spirit - in virtue of a biological mutation in the psyche - is destined to move irresistibly in this direction and ultimately to resign life and enjoy a supreme moment of spiritual dominion over all things. Humankind cannot bear much reality: there is nothing contradictory about the distinctly conceived possibility that the great infinity of the realm of essence having opened up before us in a conscious perspective, the violence of animal life will now be gradually repudiated - not indeed by the spirit but by the mutant psyche - and the soul more and more taken up with the indescribably beautiful, but fatal, vision of being in its purity.

This leads me to the discussion of my final topic in this chapter: the relation of Santayana's epiphenomenalism to natural science. I have already in various places remarked on Santayana's conception of the nature and function of science especially on his integrating psychology as a part of biology and biology as a part of physics. I

have also pointed out his important distinction between literary and scientific psychology. In general it would be fair to sum up Santayana's attitude to science as passive and unassuming, epitomized in his statement: "In respect to the facts I am ready to accept anything that the experts may tell us for the moment, to accept it as I do the weather, without cavil but without excessive confidence."[44] However, there is one important exception to this general modesty toward natural science: his hard and fast distinction between two kinds of psychology, one dramatic and the other scientific. This amounts to the designation of the ultimate boundary of science, in at least one direction, on the basis of ontological analysis: for Santayana holds that science can and will never be able to treat of consciousness. The distinction between literary and behavioristic psychology is central to Santayana's epiphenomenalism.

The essential difference between Santayana's and Huxley's version of epiphenomenalism is this: the latter holds that science can and most probably will find the mechanical equivalent of consciousness, while the former maintains that this is impossible since consciousness is mechanically non-existent. The difference could perhaps be traced back to the fact that Santayana is a moralist, and Huxley a scientist, and consequently the latter is concerned with the advance of scientific inquiry, while the former mainly with the preservation of aesthetic and moral values. At any rate, it seems clear that Santayana thought

---

44. Spirit, p.275

the scandal of the scientifically unexplained existence of spirit was a permanent one and confidently postulated the realm of spirit as completely free of any possible interference from the side of physiology or psychology. The distinction he draws between science and literature is only apparently clear-cut; as soon as we go beyond his immediate statement of it, we encounter difficulties in trying to classify certain problematic fields of study or methods of investigation. A prime example of this is Santayana's own classification of psycho-analysis, which is spoken of as a " school of literary psychology " in one book,[45] whereas in another it is listed along with embryology and physiology as a branch of science which opens " a trap-door into the dim carpentry of the stage "[46] viz. goes well beyond a dramatically empathetic study of the conscious manifold.

In this, I feel, we are putting our finger on a real difficulty for Santayana: he has limited the scope of the sciences without consulting them first and seing whether or not they could find a mechanical equivalent for consciousness. Now if consciousness does have such a mechanical equivalent expressible in some mathematical formula, epiphenomenalism is not thereby refuted---quite to the contrary. But Santayana's version of the theory would collapse, for a decisive blow to the distinction between literary and scientific psychology would be sufficient to destroy the core of his ontological distinction between the realm of the spirit and the realm of matter. Also, the moral values he set out to restore to eminence in the sphere he thought

45. Scepticism, p.260
46. Matter, p.143

proper for them would once again be in danger of being evicted. Questions such as " Is consciousness essentially hidden to sensation or is it only accidental that we do not have a sense-organ by which to "tune in" on other conscious manifolds?" (sensation should here be taken to include any possible form of extra-sensory perception) at this point acquire great importance and may in fact prove decisive for the issue. For the object of perception is always a material thing or event: whatever is perceived can, to some extent, be treated by science. It may turn out that Santayana's elaborate ontological distinctions have been founded on current ignorance.

To all this Santayana might reply that it is undeniably true that there is a private, immediate and causally inert element involved in our experience. If this has a mechanical equivalent which can be expressed mathematically, handled psychologically and explained physiologically: in a word treated by natural science, we are not saying more than that consciousness has two aspects. The subjective aspect certainly cannot be treated scientifically: it is the condition of having any knowledge at all---and it is this that he refers to when he speaks about the spirit.

I am not at present concerned with challenging this claim. Here I only wish to point to an important general consideration which is suggested by the above discussion. Epiphenomenalism, in any of its possible forms, is dependent for its verification or falsification on

natural science. An epiphenomenalism based on ontological analysis only is, to my mind, of little value---especially if we admit, as Santayana does, that the foundations of things are revealed to us by physics. There are speculative ontological reasons for and against epiphenomenalism, but these by themselves are never enough; they have to be supplemented by all available factual information which general human observation and the sciences provide. At the beginning of our discussion of Santayana we admitted that there was no reason why one should minimize the number of ontological factors involved in the world at the cost of inaccuracies - which in the long run may make a crucial difference - and an unwarranted high level of abstraction only in order to carry Occam's principle to its logical conclusion and arrive at a unity of ground for all phenomena. Similarly, however, now I must say that I can see no reason for postulating different ontological realms when it may turn out that several can be reduced to a single one, even though this may seem incredible to my present way of thinking. Epiphenomenalism is based on science because the only tolerably meaningful notion of how a material event can produce a mental event is scientific. To speak of harmonies in matter resulting in the emergence of spirit is agreeable mythology but makes very poor sense when taken literally. To speak of brain events and neural circuits and synaptic connections between chains of nerve-cells may in the end also turn out to be a myth, or good sense couched in mythological terms, but here we are in an altogether different region, where experiments can be devised, practical results reached; where facts

can explode theories by refusing to be handled under defined and controlled conditions in ways predicted. Once this essential dependence of epiphenomenalism on science is summarily seen, we may proceed to a closer examination and the eventual evaluation of the theory.

CHAPTER V

CONSCIOUSNESS

It may be well, at this point, briefly to summarize what we have established so far. For background purposes we have started this discussion of epiphenomenalism with a short review of Descartes' theory of mind. Descartes introduced a rigid dualism of body and mind in human beings, but held, with the Port Royalists, that animals are unconscious machines. The former of these views was accepted by Huxley, the latter modified and extended. On the one hand, according to him, epiphenomenalism is based on the belief that the body-mind disjunction is justified and correct. On the other hand, while Descartes maintained that animals are unconscious material machines, Huxley amended this to the effect that at least some animals as well as human beings are conscious machines. However, if we are conscious machines, consciousness can do nothing more than observe the functioning of the machine. This, roughly speaking, is the logical origin of the epiphenomenalist's witness-analogy, in terms of which the mind is conceived as a concomitant effect of certain types of physical activity and hence as a mere spectator of such activity---a spectator which has not the least chance of influencing things or animals in the environment or even the brain which presumably produced it.

Thus we have seen that consciousness is, for the epiphenomenalist, not a thing in the usual sense of that word--its ontological status,

involving an essential causal impotence, is altogether different from that of any other function or entity in the natural world. Accordingly, we had to introduce Santayana's ontological distinctions in order to render Huxley's unsophisticated epiphenomenalism consistent. In other words we may say that Santayana, with his distinctions between different kinds or realms of being, supplies the metaphysical background for the theory of conscious automatism. We examined these realms of being, and in reviewing Santayana's attitude toward science we found that his distinction between literary and scientific psychology was ultimately arbitrary and may on that account be proved wrong by the advance of science. Also, it involved a contradiction of Huxley's hope that one day science will discover the mechanical equivalent of mind, and treat of consciousness as it now treats of heat or motion or animal tropisms.

One of the most important points in the argument was that for the epiphenomenalist "mind" is roughly equivalent to "consciousness", which is not an entity in the natural world but a function, an activity dependent for its generation on certain, in principle observable, material involutions in that world. Mind is nothing but a set of discontinuous conscious events and hence the only immaterial waste product of the animal organism is consciousness itself, which somehow transforms quantitative changes in the brain into qualitative "mental" events. I wish to examine some aspects of this apparently unique activity of consciousness.

Without consciousness I can have no thoughts and no sensations. But

now I ask, is this because without consciousness there is no "I" (self-subject) or because without consciousness there are no thoughts and sensations? Is it the case that *esse est percipi* holds for sensations and thoughts and that thus there can be unconscious sensation and thought? Linguistically speaking the answer must, of course, be in the negative, since by a sensation *we mean* something like the perceptual awareness of an object and by a thought the cognitive awareness of anything at all; and awareness and consciousness are in these cases, synonymous. However, physiologically speaking there can be unconscious "thoughts" and unconscious "sensations". I will jump out of bed, even if I was sleeping soundly, if somebody sticks a pin in my arm. The sensation of intense pain in my arm will wake me up, bring me back to consciousness in a second or two---but should I say that the sensation was not a sensation before I grew conscious of it? Then what was it that woke me up---for if we limit the word "sensation" to its strict denotation ("conscious sensation" is redundant but emphatic), it could not have been the pain itself, however intense it seemed?

Then would it be more correct to say that consciousness supplies the unity of my sensations and thoughts? There is an observed continuity in my thoughts and my desires and aspirations; even in my pains. The continuity may be expressed by stating that all these thoughts, desires, etc. are *mine*: I can remember them, I feel that I have changed a great deal or very little, as the case may be, and I feel that despite these changes I am still the same person I was yesterday or five years ago.

Thus if consciousness supplies the continuity of my sensations and thoughts, it is not any different from me : it is identical with the self.

But when the continuity of conscious experience is radically broken, as in amnesia, consciousness (the present self) evidently no longer has access to information stored in brain patterns, which properly decoded would be memory of past thoughts and experiences. Thus even on this view consciousness can only be identical with my present self...but this self has the peculiarity of riding the wave of time, of being always present as the condition of anything being present in the sense of appearing, though it is momentary and presumably passing out of existence and being re-created with a slightly modified content at each instant. The past self, and thus also the continuity of the present self with the past self, is preserved, enshrined in the memory patterns of the brain. But in what sense can I say that I am still "I", myself, when I am not conscious? Am I ever justified in saying: " There he is, asleep"? Is he in fact there? His body is there and presumably his memory patterns are also there (though this latter is, to make an understatement harder to verify); and yet there is something missing. I may call what is missing a decoding agency which would translate those physical memory traces into feelings and conscious discourse, and which could add further items to this store of information gathered from past experience, or at least grow conscious of and witness the addition of new information. I know that this something extra is missing in the case of the sleeping man ----- I know this from my own experience. Consciousness is something like

the total inner difference between being asleep and awake---and the difference is enormous. In a sense one lays down one's life every night on going to bed: when I fall asleep, in some very important sense, I am no longer I. This is the enormously intense intimacy of consciousness--of the conscious self--which is temporarily absent in sleep. I am inclined to think that the fear of death is fear of the destruction of this intimacy. For considered from this point of view consciousness is either present or absent, and the difference between the alternatives is absolute.

Now note that one can get tired of being conscious, or rather, consciousness itself can get tired. Consciousness is a readiness for the world---if not to act or to modify impending actions in the light of long range considerations, then at least to perceive and to understand such actions. However, in certain cases the world can just go past you and you are no longer a step ahead, ready for what may come next. Events, actions go past the man, consciousness can no longer move along with them, comprehend them as they come in waves assaulting the mind's eye. Attention is hard to focus (and sometimes this is reflected in the eyes): it remains stationary as motion floats past the sense organs and thought remains imperfectly understood. Everything seems strange and unreal: this experience is quite common after a long period of sleeplessness. From this point of view consciousness is a matter of degree: we can be more or less intensely conscious.

In this connection we would do well to introduce a distinction between

consciousness and attention. Attention implies consciousness, but consciousness does not imply attention. We are well acquainted with the curious phenomenon of looking at a thing and yet not seeing it. It would be more correct to say that we look but do not notice, since properly speaking we see the object in question---for instance, St. Peter's beard in the stained-glass church window. By this I mean that if we could secure a photograph exactly reproducing the conscious manifold taken in by the eye and lit up by the mind's eye at the moment in question---a careful study of this would indubitably reveal the beard which did not register in consciousness when our photograph was, so-to-speak, only a single frame in an extended motion picture. Consciousness may be thought of as a conscious "field", in which case attention falls usually only on a small part of that field:--occasionally the intensity of attention is more diffused and stresses a wider section of the manifold, but often it is altogether absent. It is absent especially when the conscious manifold reveals familiar objects in familiar surroundings. The familiar is accepted unquestioningly--its peculiarities are not seen, they are lost under the sugar-coating of apparent intimacy which is the most efficient disguise of essential ignorance. Not that we wouldn't notice it if the familiar (or quasi-familiar) object were missing,- for then it would be MISSING; but as it is, it is simply taken for granted like the bed in one's bedroom or the doors of which one grows truly conscious only when the keys have been lost.

This is the biological economy of consciousness. When something is seen for a long time it acquires an aspect of familiarity and is no

longer noticed. The object has been judged reliable, normal and dull: for all practical purposes it is assumed that no dangerous, life-destructive forces lurk behind it in the background waiting only for an opportunity to attack in unsuspected and violent ways. If the object has for a long time failed to manifest any irregular behavioral features, the perceptual radar of consciousness no longer stops to consider it, but skims over it habitually, finding everything in order. If the object would suddenly reveal the barrel of a gun directed at me from behind it,on the other hand,I would in a moment become intensely conscious of the object, the gun and the whole situation. If I am an active man, I would in a second be engaged in vigorous action directed at finding out what the whole thing is about, and if it is not just a friendly hoax, in instituting immediate counter-measures as best I can.

This situation suggests two observations, one favorable, the other unfavorable for epiphenomenalism. On the one hand it seems clear that to look and not to notice depends on a pre-conscious sifting of objects of attention, a physical selection of what should be considered and what not, in which the criterion of elimination is, often or always, insufficient biological interest. On the other hand, however, there is an apparent slant to this pre-conscious selection: it seems to be invariably the case that problematic situations, where actions of a high degree of complexity and intelligent articulation are necessary, are selected as objects for the most intense scrutiny of attention. We know what is meant when people in an unfavorable position are said " to think furiously" about a way out of the difficulty, or"to be intensely conscious" of the danger to which they are exposed. This is what we may understand by the biological economy of the organism:most actions are done as a matter of habit, and only the most difficult situations and operations are selected for conscious consideration (for here habit seems

nn longer sufficient), in order to facilitate their correct solution or proper execution.

The aim of philosophy is, in a very important sense, the intensification of conscious experience. The intensification is, however, not brought about by biological means such as the administering of drugs or narcotics. Philosophy begins in wonder. Wonder is conscious attention paid to an object, when the motive for such attention is not immediate biological interest. It is this faculty of bringing to explicit consciousness the fulness of an object (the object in all its peculiar incongruity and ramifications), despite the fact that biologically its importance is negligible, it is easily handled by habitual, automatic action and has for a long time been classified as familiar, " the same old (unimportant) thing "; it is this faculty that has to be cultivated, stimulated to activity by philosophy. In other words, philosophy should refuse to take anything for granted. By a conscious effort the philosopher should attempt to peel ordinary, everyday objects of their sham familiarity and see them for what they are.

Certain types of ennui can also be explained in terms of the lack of intense conscious experience. For a man living a sheltered life the intensity of attention arising from the necessity to cope with life-destructive forces, is altogether lost. At the same time, as consciousness is, so-to-speak, becoming lazy, the intellectual stimulation of philosophy to wonderment is also lacking. Thus attention has nothing to which it could attach itself externally and no internal incentive toward a revision of the past assessment

of familiar objects, viz. no urge to see the world from a new perspective. As there is no food for the attention, boredom inevitably follows. It is thought that there is nothing worth our awareness, that nothing which is close to home is truly interesting. This is a point where we have to disagree with Santayana concerning the source of happiness of spirit. Spirit is not made happy by the disinterested speculation of anything in heaven or earth; of any essence. The source of its happiness is the contemplation only of familiar essences--the fabric of joy is woven of wonderment which is nothing but seeing old objects in a new light. The two typical ways of escaping from tedium are also very suggestive and seem to confirm this general view. People bored with the familiarity of everything around them are either concerned with looking for dangerous situations, or hope to relieve the uniformity of life by long and frequent travels. The former propose to do away with the boredom of an inwardly empty life by reinstating biological interest---this is the origin of the concept of thrill. It also explains the great number of murders committed each night in miscellaneous third-rate television plays. The latter attempt to rekindle the intensity of consciousness by providing new, as yet unfamiliar objects to contemplate. The two suggested escapes from the stupor of an idle consciousness are similar inasmuch as in both the stimulation of consciousness is effected externally--by means of new situations or of new objects---and for this reason only imperfectly and temporarily.

It is interesting to note that in time of stress, and especially in

cases of emotional excitement, consciousness often becomes blurred--- it is upset and no longer mirrors the circumstances with clarity, as the lake no longer reflects the sky and the sun when pebbles are thrown into it. The eyes flood, you have the feeling that you can hardly control yourself. Afterwards one says: "I hardly remember what happened. I can remember that I went up to the speaker's rostrum, started reading in an uncertain voice---and the next thing I know I was walking down and the chairman was shaking my hand..." It seems as though in crucial moments consciousness could no longer fuction with the usual clarity; as though it were simply switched off and the body put on automatic control. We can take the case of an airplane as our model here; when the pilot blacks out the plane may be put on radio control--and such automatic control is much less liable to make errors or to fail at the crucial moment than the pilot.

" I no longer had control of the pitch of my voice, of the movement of my hands (which occasionally rose in explanation and gesticulated)-- in general of what I did...", our shy public speaker says. Is it not a contradiction to say that I had no control of what I did? The answer, I think, is: not necessarily. I may mean that I (the conscious self) had no control over what I (the body) did. The more basic reflex arcs took over direction of the body's functioning. So far the epiphenomenalist may agree with us, for the fact that consciousness can at a certain point no longer believe that it is exerting control over the body's functioning does not imply that it did ever in fact exert such control. But how is it that the speaker cannot remember anything about the period during which

he was on "automatic control" ---just as one cannot remember anything that happens to one's body in deep sleep? It seems to be the case that one can only remember explicitly that of which one has been, to some extent and at some time or other, conscious. Penfield says "....the original formation of the memory pattern must be carried out from a high level of neural integration..."[1] Is it possible that to become conscious of an event is a necessary condition of the preservation of that event, in the form of a memory pattern, in the brain? There certainly seems to be a correlation between the intensity of the original apprehension of an object and the intensity of the memory of that object. However, here it is once again debatable whether the consciouness of the object is only an accompanying phenomenon or sign of the physical formation of a memory pattern in the brain, or is conversely the cause of it.

We may think of the human nervous system-brain-mind network as a system of response-centres of ascending complexity. Certain actions, such as the beating of the heart, are taken care of automatically; others which must be periodically repeated become habitual and require no attention for their correct and efficient performance. Unconditioned reflexes may perhaps be considered to be at the bottom of this hierarchy---and consciousness at the top. Accordingly only the most difficult tasks are referred to the mind: cases where careful deliberation seems necessary. This most complex of all the response centres has the added privilege of being able to grow conscious of all actions, though not always for the purpose, or even with the hope, of controlling them. It would be disastrous if attention

---

1. W. Penfield: Cortex , p.234

would have to be paid to every little action of the body---to every heartbeat, to breathing, to the digestion of food, to raising and moving the feet in the right direction in walking and a hundred others. On the other hand, it could prove equally disastrous, on this view, if we were not in a position to deliberate on the advisability of certain actions: and for such deliberation consciousness seems essential. A careful study of maps, of enemy strength, of the morale of his troops and other factors by the commanding general of an army, can be taken as a case in point. Basically, the issue is one of selecting an action. But such is the complexity of the considerations involved, that nothing but long conscious deliberation---a series of conferences of the general with his advisers and himself---can result in the final selection and subsequent execution of the necessary action. If the general would not deliberate and carefully consider all aspects of the situation, he would be regarded foolish or unconscientious, the latter meaning something like " lacking in conscious reflection on the consequences of his actions ". Tasks which require the best of us, then, seem to be referred to our highest response centre, the conscious mind.

However, the epiphenomenalist could here object that the consciousness of deliberation is nothing more than a sign that selective brain activity is taking place. There is no necessary connection between thoughts (in the sense of non-, or pre-conscious sets of brain processes) and consciousness. Work on the localization of brain functions shows, the epiphenomenalist would argue, that the "thought elaboration centres" do not coincide with the seat of consciousness. Thus in certain centres complex brain activity

can go on without the concurrent activation of attention. The selection of any course of action is effected by groups of autonomous central processes; any consciousness of these material (electrical?) phenomena is purely incidental, superadded to these processes, but in no way enhancing their efficacy when present or reducing it when absent. This conception of "thinking" is not to be confused with the so-called Motor Theory of Thought, according to which central brain processes only serve as transmitters in the basic stimulus-response relation. Autonomous central processes can in various ways modify and amplify the sensory imput before it issues at last in motor output. Thus there are at least three main stages involved in our becoming conscious, say, of a sensation: (1) The stimulus as it proceeds from sense organ to brain; (2) The ordering, modification and amplification of the nervous impulses in regions of the brain especially concerned with the functioning of the sense organ in question---but this is, figuratively speaking, ordering in the dark; and (3) The transmission of the "sensation" to the seat of consciousness (if we wish to put this in spatial terms, as in physiology it seems proper), where in the light of attention it is made explicit: I grow conscious of it.

This may not be a physiologically exact account of what happens; and it is certainly not exhaustive. But it is well adapted to strees the epiphenomenalist's point, which I take to be, that to grow conscious of a group of autonomous central processes which select actions and control the organism's motor output at any given time is incidental and dispensable for those processes. "Thoughts" and "sensations" insofar as these are physilogical processes and neither conscious in themselves nor conscious-ed

by me, are quite sufficient to effect all the intelligent, as well as the erratic and misplaced, responses of the animal. Mental events are, roughly speaking, the way in which physiological processes appear in consciousness.

It is difficult to determine whether in comparing and contrasting the physiological and phenomenological descriptions of a sensation, we should admit only one event or two events. It is indubitable that in my conscious field I am aware of events---- and by an event we may understand a minimum noticeable change. The problem is whether the conscious event I experience is the "same" event as the physiological one which is, in principle, observable in my brain. The question may be interpreted as a request for criteria of identity which, in this case, we do not possess. However, in some very important sense the two cannot be the same event.... at best they can be different aspects of the "same" event in which case they are not one event but at least two events, one observed in my conscious manifold, another observable as a physiological change, and possibly three events, the third being the "neutral" original change of which the other two are appearances.

I have already mentioned the peculiarly 'selfy' character of consciousness. The relation of consciousness to the self is a difficult problem to state and an even more difficult one to solve. Consciousness seems to be at least the necessary condition of the existence of the conscious self---but this proposition amounts to little more than a tautology. Consciousness seems to be that indefinable (unique?) something which

includes my conscious field and "me"; this, however, is not meant to imply that object-consciousness is in all cases accompanied by self-consciousness. It is not clear whether I am different from or more than, my consciousness--whether what might be called the transcedental ego, though in our experience it is inextricably bound up with consciousness, is yet logically a different factor from it or not. The feeling of my continued existence and self-identity do seem, to a great extent, to be dependent on my memories and thus to have a solid physical basis. Without memory of the past and expectation of the future the use of the personal pronoun "I" becomes meaningless---consciousness is, as Santayana suggested, completely absorbed in its object. In this way the "self-feeling" seems to be very closely connected with time, and consciousness of the passage of time, and consciousness of the identity of the self-subject through the passage of time. This raises the question as to what I am conscious of in self-consciousness. Am I conscious of the transcedental ego, which is perhaps nothing else than the unity of the conscious manifold? However, if the transcedental ego is not logically different from consciousness, then in self-consciousness I must be conscious of consciousness, or in other words, of the fact that conscious-ing is going on. Or finally, it is possible that in self-consciousness one attends to the unity of consciousness which, however, is only the mental equivalent of the physical unity of the generative organ of consciousness. I believe the epiphenomenalists would probably insist that the last of these three alternatives is closest to the truth.

Now it seems that we will have to modify our previous statement that

consciousness is a necessary condition of the existence of the conscious self, to the effect that though certainly a necessary, it is not in all cases also a sufficient condition. This self, it is to be emphasized, is only the conscious self---but it is basic inasmuch as everything else is classified as self or not- self only in relation to it. In actual experience there is a very close correlation between the self and consciousness: for example, consciousness never leaves me, for I leave with it: the "I" never survives the loss of consciousness in sleep, or in a coma, or in death. It is revealing to note the ambiguity of the personal pronoun "he". We say indifferently: " There he is, talking to his wife", " there he is, asleep " and " there he is,dead ". The last of these,at least, is illegitimate, for in death we can refer neither to his substantial biological, nor his conscious self. Even when we say: " There he is, talking to his wife ", we don't refer to his conscious self alone, but to him as a dark man 5'10" tall, standing in a corner of the room talking and at the same time thinking about what he is saying, and perceiving, recognizing, seeing us in a conscious perspective. This is the crucial difference between the "him" and the "I": for me he is still there even when he is asleep, and for him I am still myself, even though I faint or am etherized on the operating table. But for myself I am no longer there, I am as good as dead when I have inhaled enough ether to lose consciousness. If the totality of my observable behavioral characteristics is the me-to-him, consciousness or the conscious self, not observable to anybody but myself, is the me-to-me. It is the miner's lamp on top: a complete circle--I can no more escape the necessity of having to see everything in the light of this lamp than I can step out of my skin. Consciousness is the

epistemological lightbulb without which we cannot see the furniture of the world---and since there is no knowing without it, there is no way of telling whether it distorts or not, or whether we are not only having seemly dreams in a cold unfurnished room.

There is a possible argument against epiphenomenalism from the functional character of pain. It is evident that at least for pain esse est percipi: a pain, in the nature of the case, cannot exist without somebody being conscious of it. To speak of an unconscious pain (or unconscious pleasure) seems to be a contradiction in terms, for a pain is not a pain unless it hurts someone---unless it solicits the attention with a stubborn and imperious obstinacy. But it is at the same time generally accepted that pains are biological deterrents to action: one will not touch an electric wire because of the extremely disagreeable sensation of shock---occasionally one will even check one's innate tendency to be a glutton if the memory of the previous night's gruelling indigestion is vivid enough. Thus pain is biologically important and causally effective; and consciousness is a necessary condition of the existence of pain. A conscious sensation is in this way involved as an indispensable link in the causal chain culminating in actions directed at the maintenance of bodily equilibrium. Similarly, pains seem to be signs for remedial action to be taken. If I cut my finger the pain seems to be a desperate signal of the part concerned, a bid to capture my attention in order that I may see what had happened, stop the bleeding and bandage the wound. Pain is equivalent to the cry " Do something! " and as such it is ultimately directed to the conscious self.

Now the epiphenomenalist would probably accept the first premiss of the argument, but would indubitably deny the second. It is quite possible, he would admit, that pains can only exist if we are aware of them. However, there is no reason for us to believe that pain, always involving consciousness of itself, is in any way biologically significant. Pain is bad and useless; bad because it results in distraction,[2] and useless because it can occasion no other remedial action than the nervous impulses which cause it can and usually do. Biologically pain is not an added factor beyond the nervous impulses of which it is the mental equivalent. It is not a sign <u>for</u> the organism to shun a certain object or a particular course of action: it is rather the sign <u>of</u> such evasive activity on the part of the organism. The same mechanism that functions with pain present would function even if we were not <u>conscious</u> machines and thus the quantitative distribution of our nervous impulses could not be "mentalized", translated into the feelings of pleasure and pain. Consciousness, even as the consciousness of pain, cannot be a part of the animal machine. The machine is physical and everything physical is, in principle, observable. The basic assumption of epiphenomenalism is that every sensation and thought and feeling can be broken down into two factors: the nervous processes involved, and the mentalizing activity of consciousness. All causal influence of these thoughts derives from the nervous processes which, of course, are physical. In this way, though consciousness itself has a physical basis, it is biologically a waste, and causally the equation " nervous system ≡ nervous system + consciousness " holds without reservation.

---

2. George Santayana: <u>Spirit</u>, p.126

CHAPTER VI

EPIPHENOMENALISM AND SCIENCE

" Theoretically, if we could build a machine whose mechanical structure duplicated human physiology, then we could have a machine whose intellectual capacities would duplicate those of human beings." [1] This is a good summary of the attitude of many scientists to the human mind. The " great engineering dream " is now slowly becoming the dream of constructing machines which duplicate more and more facets of human behavior with ever increasing fidelity. In many fields, of course, the machines are already vastly superior to the human mind: for instance, they can calculate or count better and faster than we can. But these operations may be called "mechanical": they are repetitive and uninteresting---- this is not the field where human beings excel. In general, I think it is fair to say that most machines that have so far been constructed lack the adaptability of living animals and imitate only a very few, selected human behavioral traits. We are indubitably still a very long way from the ultimate aspiration of being able to construct robots approaching the organizational intricacy and behavioral characteristics of animals and men. At this point, it is even questionable if we ever will be in a position to construct such machines, or if such machines can be constructed at all. However, the rapid advance of cybernetics, the recent invention and subsequent development of electronic machines of hitherto undreamed of complexity, should not be lightly dismissed. It is indicative of the current tendency

---

1. Norbert Wiener: The Human Use of Human Beings (Doubleday Anchor Books: New York, 1954), p.57

of science to blend the technician with the medical doctor, the engineer with the physiologist.

The concepts of cybernetics have lately infiltrated the field of physiology. Certain basic similarities have been observed between the functioning of the nervous system in animals and the control mechanisms of electronic machines. The similarities are thought to be expressed in concepts such as "homeostasis" and "negative feedback". Homeostasis may be described as a state of stable equilibrium: a good example is the animal's body temperature maintained roughly at the same level throughout the internal environment. Negative feedback is an automatic control mechanism which regulates performance at any one moment by taking account of results reached by the performance of the immediately preceding moment. The concept is applicable to physiological as well as electronic mechanisms. It may be considered purposive inasmuch as it is directed at the maintenance of homeostasis; if output exceeds or falls below a certain definite level, information of this deviation is fed back to the controlling centre, where proper adjustments are made to compensate for the increase or reduction by slowing down or accelerating the rate of activity. For instance, when I reach out my arm to take a book from the shelf, there is a continuous and very rapid feedback of information to the brain centres, concerning the actual position of my palm and fingers at each moment, from which the controlling organ (or process) can calculate and determine how much further I have to stretch to reach the book, how solidly I have to wrap my fingers around it so that it won't slip out of my hand and how far I have to take it in order to be able to place it on the table.

Negative feedback is error-actuated or error-compensating, and as such

it is believed to be the basic principle underlying the functioning of the nervous system. A negative feedback mechanism always involves effectors and receptors; the former execute tasks, the latter relay information of the actual performance of the effectors, which is then immediately converted into data utilized in the adjustment and subsequent control of those effectors. In a human being the effectors are evidently the muscles and the receptors the various sense organs. From a different point of view every sensory stimulus may be considered a temporary disturbance of the organism's homeostasis eliciting, via a negative feedback mechanism, instantaneous response. This is sometimes offered as an explanation of motor reflex actions, and with certain modifications even of much more complex animal activity.

I hope these few words will suffice at least to indicate the trend of contemporary science, in general, and the importance of cybernetics for this development, in particular. Cybernetic models are used in an ever widening context in physiology: and advanced work in these two sciences seems to point in the direction of the final thesis that human beings are enormously and wonderfully complex machines. This, of course, is the view of epiphenomenalism: and in the light of these new scientific developments it is surprising that the theory is receiving so little attention today. If science will ever be in a position to prove conclusively that men are machines, it will be extremely hard-if not altogether impossible- to escape the epiphenomenalistic conclusion.

By an organism being a machine I mean that the three factors of sensory

input, built-in tendencies and stored memory information available at any time, determine the motor output at that time. All these factors are physical: consciousness, being an immaterial function, can neither be a part of the machine, nor can it influence the machine. We have talked about "moral values" and "inner life" in this thesis without once attempting to state what these values are, or in what that life consists. This is understandable because there are important and extensive disputes about the nature of value and of the good life, and we did not want to get involved in any of these. However, we may say at least this much, that the danger of science is that when and if men finally turn out to be animal machines, then it will be possible to control them as machines are controlled. If the truth of this proposition is seen it will also be evident that under such conditions all moral life, conscious joy and true individuality is made impossible.

This should, of course, not be taken as an a priori, moral argument against science: it is only an assessment of the situation and a brief glance at the possibilities. Once again, it should not be regarded as implying that scientific inquiry will ever arrive at the conclusion that human beings are automata---or that it would ever be in a position experimentally to verify such a conclusion. However, it is indubitable that science seems, at present, to be developing in a direction ultimately favorable for an epiphenomenalistic interpretation of the relation of mind to body. It is also evident that the tendency to regard human beings as machines is a serious threat to the moral life, and in the end to the existential destiny of the human race. The conflict is apparent in the

fact that ordinary men don't think of themselves as automata only--they speak quite readily of responsibilities, aspirations and of an intense subjective intimacy which is considered so alien to all actual and possible machines. The danger, on the other hand, is nowhere more apparent than in the fearfully thorough brainwashing-which is perhaps a form of conditioning- that can be accomplished even with our very limited knowledge of psychology.

This leads me to a consideration of the relation of epiphenomenalism to science. Properly, this topic should be discussed in the general context of the relation of science to philosophy. However, such a generalized discussion is certainly beyond the scope of this thesis. We will have to be content with a short examination of some aspects of the question as it more particularly concerns the theory of epiphenomenalism.

It is generally accepted today that philosophical and scientific questions are not of the same type. The problem of whether or not perceptions are representative of objects, for example, or whether the world is one or many, cannot be decided by observation and experiment. Possibly such questions cannot be decided at all, not indeed because they are pseudo-questions or muddles but because they are enigmas underlying our very existence. There is something paradoxical about brute givenness. When we ask basic questions either we have a deep-seated feeling of vacuity or dizziness which makes us wonder if we are asking anything at all, or we have a feeling of supreme confidence in the significance and importance of the issues. We may

be closer or farther to the solution of philosophical problems and for various reasons it may turn out that certain problems are scientific rather than philosophical. When man first began to reflect on his circumstances, to think about the world and about himself in the world, all questions were philosophical. Early cosmologies were made of the stuff of dreams and of wide guesses; the spontaneity of the imagination embroidered on the observed but sometimes ludicrously irrelevant conjuctions[7] in the world to construct fabulous theories and a speculative physics. There was no way of verifying what was right and what was wrong in this web of miscellaneous hypotheses. Now it seems probable that even if all possible methods of verification were available to us, there would still remain a residue of insoluble, but not for that reason unreal, philosophical problems. However, as new methods of verification are invented, more and more problems become, in principle at least, soluble. Artificial sense organs are produced, which enable us to transform hitherto unobservable phenomena into a form which can be easily observed by one or more of our five senses. The radio is a good example of such a development. The statement: " The stars emit mysterious unobservable signals which run through the bodies of animals and men without these being in any way conscious of them or significantly affected by them," would have been considered totally unverifiable eighty or a hundred years ago. Today the new science of radio-astronomy is concerned with just these signals rendered observable by Marconi's invention, and the above proposition has been conclusively verified.

There is, then, a gradual transference of problems from philosophy to

the field of science. Philosophy has at least two functions with respect to science: first it has to examine critically the conceptual framework on which a science is built, and secondly it has to attempt to correlate the results of various fields of scientific inquiry in order to provide a unified view of the current state of our knowledge, and supplementing this by speculative theories, eventually a unified view of the world. The first task may roughly be described as analysis, the second as generalization. Now it seems to me that a number of problems connected with the theory of mind are at present in the process of being transferred from the sphere of philosophy to that of science. The rapid development of biology in the last hundred years and the pronounced emphasis lately on the scientific aspect of psychology are at least two indications of this trend. Physiological psychology is slowly beginning to flourish, and we can expect, I think, further interesting information from neurology and the application of the concepts of cybernetics in the field of physiology. This is why I have previously maintained against Santayana that this is no time to insist on a hard and fast distinction between literary and scientific psychology, of which only the former can, by imitative sympathy, approach the phenomenon of consciousness. It is not impossible that an artificial sense organ may be invented by which we would be enabled to "tune in" on the mental discourse of others. It is not impossible that some of us actually have such a strange (sense?) organ active in the problematic phenomena of extra-sensory perception, which facilitates a direct awareness of another's conscious manifold without the benefit of observing the mediating factor of the person's behavior. It would, of course, not do

to employ these imaginative points as arguments against epiphenomenalism in general, and Santayana in particular, but I believe that if we want to arrive at an honest estimate of the situation, we must certainly take them into account and attempt to appreciate their significance.

Thus one possible reason why epiphenomenalism is ultimately dependent on science is that questions about consciousness, intelligence and the relation of conscious events to neural processes seem to be becoming scientific questions. Another reason is the general consideration that philosophical theories, even when frankly speculative, should take account of the factual information of the sciences. Ontological analysis should incorporate, as far as possible, the well-attested results of scientific inquiry. If these results, or the theories which they are usually taken to confirm, are not put to good use in the analysis, reasons for this should be given. Also, no philosopher should hesitate to plead ignorance when ignorance is honestly felt; and the possibility of further scientific advance in the field should be acknowledged and research leading to such possible advance at all times encouraged.

There is, however, another more compelling reason for the dependence on science of this particular doctrine of the relation of mind to body. Epiphenomenalism is also called the theory of conscious automatism. The concept of "automatism" is a scientific one. It is not enough that we can conceive of the animal organism as a machine: here positive evidence is needed. For if we can conceive of the animal as a machine, we can also conceive of it as an oligarchic society of cells, a teleological system of

parts each of which presupposes for its existence all the others and thus the whole, and in a number of other ways. We are using the concept of a machine and applying it to animals and men: but the concept is a borrowed one and thus before we can go a step further we must produce sufficient evidence to the effect that the human body is a machine. This evidence can only be scientific in nature---based on observation and experiment.

I must remark that this does not invalidate Santayana's ontological epiphenomenalism. Such an epiphenomenalism amounts to the resolution that from now on we shall regard the sphere of power as essentially (though not existentially) distinct from the sphere of conscious acquaintance or knowledge. It involves, as we have already pointed out, an analysis of each conscious event into, roughly, a set of physical processes which are causally effective and the mentalizing factor of consciousness which is not. The analysis seems to be logically a legitimate one, even though it carries Descartes' radical disjunction of mind and body and his conceptual atomism to their final conclusion, where not even the power of causation can belong to the mind once it is assigned to body. There are two criteria by which we could evaluate such a purely ontological epiphenomenalism, and Santayana's theory fails on both these counts. If a theory relies on no scientific evidence or sophisticated argumentation, then I can judge its value and estimate its probability by determining to what extent it does justice to my experience. Being of the highest possible degree of generality, it should account for the characteristic features of experience by means of a minimum number of assumptions. Epiphenomenalism

is inadequate in this respect because, though claiming the authority of common sense, it disregards the evidence of common experience. My everyday experience, if I do not wish to go beyond it, will never reveal to me the putative fact that I am an automaton, or suggest that my conscious self is in no way involved in the selection of my actions. Quite to the contrary: I feel that I (conscious self) can and do react on the world, that I can consciously control my actions, execute my consciously contrived plans and sometimes to some extent even realize my dreams and my aspirations. This is direct evidence against epiphenomenalism, and ontological analysis can give us no serious ground for believing that the evidence is merely illusory.

The second criterion which we could apply to the ontological form of epiphenomenalism follows from the analytic task of philosophy with respect to science. It is the following: Can we expect any new results, in experimentation or amounting to a revision of the conceptual framework of any of the sciences concerned with some aspect of the body-mind problem, if we grant any or all of the assumptions and conclusions of the theory of conscious automatism? The answer has to be in the negative. Epiphenomenalism can lead to no new experimentation because new experimentation led to the formulation of epiphenomenalism. Similarly, this theory cannot lead to the revision of the concepts of physiology or psychology, because the introduction of new concepts into these sciences, especially into physiology, gave the initial impetus to advance such an analysis of the experience situation. The theory is essentially derivative and based on the influx of scientific concepts into philosophy. Ontological epiphenomenalism may

represent a consistent though imaginative analysis of experience into logical factors, but without the cumulative evidence of science it carries little or no force of conviction. Epiphenomenalism is a philosophical theory that is built, and can only be built, on a foundation of scientific fact, though it involves a speculative extension of all such fact.

My next task is to try to elucidate in what sense epiphenomenalism can be considered to result from a conflict of science and philosophy. In at least one sense there can be no conflict between science and philosophy. The methods and the objectives of the two are altogether different, and when they are in an ideal state of equilibrium science and philosophy mutually complement one another. The philosopher is concerned with conceptual generalization and the scientist with factual investigation. No philosopher will dispute, say, the correct laboratory method of analyzing the chemical constitution of a compound, and similarly no scientist will try to argue about the nature of the good or examine the problem of universals. However, a conflict of these two distinct fields of human endeavor is possible under certain conditions. I have previously remarked that as our observational methods improve, certain questions which were originally thought to be philosophical are transferred into the sphere of science. Such a transference of problems is usually accompanied by a philosophical upheaval and the wholesale destruction of speculative theories advanced as the putative solutions of these problems. There is a confusion within philosophy because it is not clear how much of the subject in question can be treated by science and how much of it (if any at all) is left for philosophy. Also, scientific data newly made available are ill-digested and poorly integrated by philosophers. As a result we have hybrid theories and immature

generalizations reflecting the confusion of philosophers and of scientific theorists alike.

The great period of scientific advance from the seventeenth century to our own day is perhaps the best example of what I have called " the transference of problems". I will not bring elaborate historical examples to illustrate my point: such examples abound in any historical work touching on a comparison of Medieval with, say, Seventeenth Century science. Also, we may remind ourselves of the revealing fact that physics used to be and is sometimes still called "natural philosophy". Much of modern philosophy since Descartes is vitiated by the fact that somehow philosophers have been unable to find the right attitude toward science, assuming now a patronizing air of superiority, then the servility of uncritical admiration, at one point attempting to imitate its methods, at another rejecting it altogether---never quite succeeding in discovering the proper function of philosophy and in this way reaching a healthy balance. Epiphenomenalism is thus not only the direct result of the conflict of philosophy and science in the field of psychology and the theory of mind; it is also a symptom of the more general conflict of these two disciplines, revealing a deep rift within contemporary civilization and twentieth century man. When, to cite a crude example, we think of the hydrogen bomb with a sense of general accomplishment and yet go on to describe its horrors in a war which we say is possible only because of our moral immaturity, we are referring in a practical context to the same basic conflict, deploring our sad state in which control over the environment is not accompanied by self-control and ultimately by control over the uses to which this environment-control is put.

More specifically epiphenomenalism is the result of the conflict of science and philosophy with respect to the correct method of investigating the phenomenon of consciousness. It was formulated as a consequence of the introduction of the concept of automatism into philosophy. In the seventeenth and eighteenth centuries automatism was conceived on the model of clocks and clockwork music boxes; in the nineteenth century on the model of the heat-engine burning a combustible fuel. Today it is conceived on the more sophisticated pattern of negative feedback systems and electronic "brains". This is not to say that we have conclusive proof that the human body is a machine and that the central control system of the brain functions exclusively or even predominantly as a feedback unit. This proposition has not to date been verified, even though it is often believed that the cumulative evidence of biology and allied sciences points in the direction of its eventual confirmation. Epiphenomenalism assumes its truth at the outset and then attempts to make room for the old philosophical discipline of introspection in this new context. Automatism is taken to imply that consciousness, being "subjective" and to the best of our knowledge unobservable, cannot be part of the machine. Of the hierarchy reflex-response-conscious response, it is assumed that from the point of view of causal efficacy and biological survival value, the last term is dispensable. In this overhasty crystallization of views in an essentially fluid and developing scientific situation, the theory of conscious automatism can do justice neither to introspection nor to experiment. Both philosophy and science are forced into a compromise: the effectiveness of philosophy is restricted to the inner life and observational methods are denied the possibility of ever treating of the "subjectivity" of consciousness on

scientific principles. The doctrine of conscious automatism is truly a hybrid theory, and it implies a general notion of the relation of philosophy to science which is both confused and unsatisfactory. It is confused because it leaves undetermined how much of the experimental results of science should be taken into account by philosophy, and how what is thus acknowledged should be integrated. It is unsatisfactory because the theory imputes more to science than science has managed to establish, denies what research at a later date may establish and finally rejects the immediate evidence of common experience.

The epiphenomenalist claim can be stated in the following manner. There is no stimulus x of any complexity whatever, such that my becoming conscious of x would have to be a necessary condition of my appropriate response to x. Once again, we may phrase the claim as follows: The circuit sensory input- autonomous central process - motor output is complete on any level of complexity without the intervention of consciousness. Consciousness is unnecessary for the functional completion of any reflex arc.

Now that we have seen that epiphenomenalism is essentially dependent on science not only for the corroboration of its assumptions but also for the verification of its conclusion, let me proceed to investigate whether or not such verification is, even in principle, possible.

The epiphenomenalist holds that consciousness is a waste product of the biological organism. Waste products are inevitably present wherever organisms are present, and consciousness is the waste product, presumably,

of high grade brain activity. On the basis of fairly conclusive scientific evidence we may say that though indirectly dependent on the whole brain, consciousness is thrown off as a waste product of the functioning, more especially, of a certain limited part of the brain. Let us say, in accordance with recent theories about the localization of function, that this brain-segment is the diencephalon. The production of consciousness will not, of course, make a vestigial organ of the diencephalon; it may, on the contrary, put it in a position analogous to that of the kidneys, which are concerned with the elimination of liquid wastes. The epiphenomenalists would not, and I suspect could not, further specify the manner of production of a presumably immaterial waste by a material organ, nor again the peculiar constitutional difference there must be between the diencephalon and other parts of the cortex, so that it alone, and not some other brain segment, generates the by-product. However, we shall not insist on these difficulties: it is sufficient for us to note them.

Now evidently the easiest way of showing that the epiphenomenalistic equation " nervous system = nervous system + consciousness " holds would be by observing the action of the nervous system with consciousness present and then with consciousness absent, and subsequently comparing the two sets of observations. This is the only way we can do that sort of thing in science, and if it can be done at all, we will certainly be able to do it because we are concerned with matters of functional order and disorder and causal efficacy observable in action and behavior. Now let us remove a part of the frog's brain in order to eliminate its consciousness

(in the absence of contrary evidence I am assuming that a frog has consciousness, but this is incidental to my point for the hypothetical experiment could equally well be carried out on a human being). A change in the frog's behavior inevitably follows.[2] If we now wished to argue that the observed difference in behavior is due to the fact that consciousness is no longer present and actively engaged in the direction of the animal's actions, the epiphenomenalist could retort that this is by no means the case. The behavioral difference is due not to the elimination of consciousness, but to the elimination of the controlling centres of certain vital functions together with consciousness. The "seat of consciousness" is not a vestigial organ with no other function besides that of producing consciousness. Most probably it is also the controlling centre of a number of important nervous functions.

Thus we are confronted with the following situation. Whenever we eliminate consciousness and note a difference in the subject's behavior (and such a behavioral change is inevitable), the epiphenomenalist can claim that because of our imperfect experimental methods we have concurrently destroyed or at least impaired the functioning of some physical control centre. The epiphenomenalist, demanding the translation of all statements about consciousness implying causal efficacy, into statements about physical neural processes, seems to have not a strong or convincing but at least a consistent and well-contrived position. But now on further reflection we find that this argument which seemed originally to protect

---

2. This change may not amount to more than the loss of initiative: in the absence of any gross stimulus the animal sits completely passive and would starve sooner than feed itself. Huxley admits this (Animal Automatism, p.9); the difference between post-and pre-operation behavior is undeniable. Huxley's argument for epiphenomenalism depends on a conditional inference from "some" to "all" ("...the machinery which is competent to do so much.., might well do all"-p.9) and not any claim to the effect that the elimination of consciousness does not entail changes in the animal's behavior.

epiphenomenalism against objections based on experimental evidence, reveals a crucial shortcoming of the theory. For the epiphenomenalist maintains that consciousness is a waste-product of the perfectly functioning nervous system-brain unit. To eliminate consciousness, then, in order to show that its presence or absence is irrelevant to the functioning of the nervous system, we will necessarily have to impair the functioning of that nervous system. No matter how refined our experimental methods become, we will still have to interfere with the functioning of the nervous system-brain unit in order to eliminate consciousness, for so long as that unit functions perfectly consciousness will be present. But if the brain-balance is thus disturbed, it is inevitable that functional disorders and behavioral differences will be manifested. In this way it is clear that no experiment can possibly be devised which should verify the epiphenomenalist thesis. If consciousness is held to be the causally impotent concomitant effect of a perfectly functioning brain, then it is impossible in any scientific experiment to separate the two in such a manner that with consciousness absent the nervous system will still operate perfectly.

The above, I feel, is sufficient to show that epiphenomenalism cannot look to science for confirmation of its conclusion. We have already seen that its most important premiss, namely that of the biological automatism of animals, is inadequately supported by scientific evidence. However, the theory is dependent on scientific concepts for its original formulation. We did not claim that it is a scientific hypothesis: it is indubitably a philosophical theory. But without the backing of science, as an ontological theory it carries no force of conviction. In addition we have now established

that it can expect no direct support from observation or experiment. Our conclusion is that even though ontological epiphenomenalism is not self-contradictory, it is an unsatisfactory solution of the problem of mind and body. Evidence for the theory is scarce and the arguments advanced to support it (when such arguments are presented at all) are far from convincing. Its value lies in its recognition of the problem which underlies it and in the indecision which bred it. Being a symptom of the times, it reminds us of something more important than itself, namely of the fact that ours is a scientific age and that before philosophy finds its place in the modern world, its pretensions and its just claims must clash with those of science. The clash reveals a split within the human personality.

BIBLOGRAPHY

ASHLEY, W.R. — Design for a Brain, Chapman and Hall, London, 1952.

BALZ, A.G.A. — " Dualism in Cartesian Psychology and Epistemology ", in Studies in the History of Ideas, Vol.II, Columbia University Press, New York, 1925.

CORY, D. — The Letters of George Santayana, Scribner's New York, 1955.

DAVIDSON, W.L. — " Definition of Consciousness " in Mind, Vol.VI, 1881, pp.406-412

DELAFRESNAYE, J.F. (ED.) — Brain Mechanisms and Consciousness, A Symposium organized by The Council for International Organizations of Medical Sciences; Blackwell, Oxford, 1954.

DESCARTES, R. — Philosophical Writings, a selection translated and edited by E. Anscombe and P. T. Geach, Nelson, Edinburgh, 1954.

ECCLES, J.C. — " Hypotheses Relating to the Brain-Mind Problem," in Nature, Vol.168, 1951/2, pp. 53-57.

FEIGL, H. and SCRIVEN, M. — The Foundations of Science and the Concepts of Psychology and Psychoanalysis, Volume 1 of Minnesota Studies in the Philosophy of Science, University of Minnesota Press, Minneapolis, 1956.

HEBB, D.O. — The Organization of Behavior, John Wiley & Sons, New York, 1949.

HENDERSON, T.G. — Essence, Matter and Nature in the Philosophy of Santayana, Harvard University PH.D. thesis, 1939.

HOLT, E.B. — The Concept of Consciousness, Macmillan, New York, 1914.

HUXLEY, T.H. — Animal Automatism and Other Essays, Fitzgerald, New York, 1884.

Lay Sermons, Addresses and Reviews, Macmillan, London, 1870.

JAMES, WILLIAM — " Are we Automata?" in Mind, Vol.IV, 1879, pp.1-22

- Essays in Radical Empiricism, Longmans, Green and Co. London,1912.

- Principles of Psychology, Vol.1 Holt & Co., New York,1890.

JEFFERSON, G. — " The Mind of Mechanical Man," in British Medical Journal, Vol.1, 1949, pp. 1105-1121.

LASLETT, P. (ED.) — The Physical Basis of Mind, a series of broadcast talks; Blackwell's, Oxford, 1951.

LOCKE, JOHN — An Essay Concerning Human Understanding, Clarendon Press Oxford,1894 (ed. by A.C. Fraser), Bk.11, Ch.1, #9 - 25,Ch.27

MONTAGUE, W.P. — " Consciousness a form of Energy," in Essays Philosophical and Psychological in Honour of William James, Longmans, Green & Co. New York, 1908.

MOORE, G.E. — " The Refutation of Idealism", in Philosophical Studies, Kegan Paul, Trench, Trubner & Co. London,1922.

MUNITZ, M.K. — The Moral Philosophy of Santayana, Columbia University Press, New York, 1939.

PENFIELD, W. and RASMUSSEN, T. — The Cerebral Cortex of Man, Macmillan, New York,1952.

PETERS, R.S. — Brett's History of Psychology, George Allen & Unwin, London,1953.

RUSSELL, B. — The Analysis of Mind, Allen & Unwin, London, 1921.

RYLE, G. — The Concept of Mind, Hutchinson's University Library, London, 1949.

SANTAYANA, G. — The Idea of Christ in the Gospels or God in Man, Scribner's, New York, 1946.

Obiter Scripta, Scribner's, New York, 1936.

- Platonism and the Spiritual Life, Scribner's, New York, 1927.

- The Realm of Essence, Scribner's New York, 1927.

- The Realm of Matter, Scribner's, New York, 1930.

- The Realm of Spirit, Scribner's, New York, 1940.

- Scepticism and Animal Faith, Dover, New York, 1955.

- Soliloquies in England and Later Soliloquies, Constable and Co., London, 1922.

SCHILPP, P.A. (ED.) — The Philosophy of Santayana, The Library of Living Philosophers Series, Tudor, New York, 1951.

SCRIVEN, M. — " The Mechanical Concept of Mind," in Mind, N.S. Vol.62, 1953, pp.230-240.

SLUCKIN, W. — Minds and Machines, Penguin Book, London, 1954.

SMITH, N.K. — Studies in the Cartesian Philosophy, Macmillan, London, 1902.

TURING, A.M. — " Computing Machinery and Intelligence," in Mind, N.S. Vol. 59, 1950, pp.433-460.

VERSEELD, M. — An Essay on the Metaphysics Of Descartes, Methuen, London 1940.

WIENER, N. — Cybernetics, John Wiley & Sons, New York, 1948.

— The Human Use of Human Beings, Doubleday Anchor Book, New York, 1954.

WISDOM, J.O., SPILSBURY, R.J., and MACKAY, D.M. — "Mentality in Machines", a symposium, in Men and Machines, Aristotelian Society, Supp. Vol.26, Harrison, London, 1952.

WISDOM, J. — Problems of Mind and Matter, Cambridge University Press, 1934.

YOUNG, J.Z. — Doubt and Certainty in Science, Clarendon Press, Oxford, 1951.