

## Self-Consciousness

## Preliminaries

In order to avoid premature arguments, I will begin by saying that at present I have no reason to believe that what I mean by “self-consciousness” is essentially different from what other persons now living in the Western world who have an interest in philosophy mean by it. In philosophy, the term “consciousness” is often used to mean the same thing as “self-consciousness”. Other phrases used to describe self-consciousness are: “the awareness of oneself as being distinct from the rest of the world”; “the state that one is in when he asserts, ‘I exist’.” We will avoid other premature arguments if we agree at the outset that there are degrees of self-consciousness. Thus, we will say that a person has very little, if any, self-consciousness during a deep and dreamless sleep; that he has more self-consciousness when absorbed in some activity, e.g., watching a movie which he enjoys, and still more when, e.g., he is thinking about how he should spend the rest of his life.

Julian Jaynes’ *The Origin of Consciousness in the Breakdown of the Bicameral Mind* suggests that self-consciousness as we in the Western world know it, may not be a universal phenomenon among all present human cultures, or in Western and Middle Eastern European culture over the period a few thousand b.c. up to the present.

My initial purpose is to explore the question, “Are there propositions about self-consciousness?”, i.e., sentences:

- which can be shown to require the existence of self-consciousness for their meaning;
- which are either true or false, and for which there exists, at least in principle, a way of determining which is the case?

The importance of this question occurred to me as I read Hofstadter and Dennett’s *The Mind’s Eye* (Hofstadter, Douglas R., Dennett, Daniel C., *The Mind’s I*, Basic Books, Inc., N.Y., 1981), which includes several interesting stories, thought-experiments and discussions concerning self-consciousness. Since most of the often conflicting arguments seemed equally convincing, I began to suspect that this might be another case of literature dominating science. In such circumstances, our first step should be to ask if there are any propositions applicable to the subject in question. If we are unable to discover such propositions, that simply means that, at least at present, the subject cannot be studied scientifically or logically. It in no way implies that the subject is not important.

I will frequently, but not exclusively, use the computer as a means of minimizing the vagueness that, at this time (1988), is inevitably part of such a discussion.

## Are There Propositions About Self-Consciousness?

A “proposition” is a sentence which is either true or false, and for which there exists, at least in principle, a method of deciding which. Thus, for example, “Water boils at 212 degrees Fahrenheit at sea-level” is a proposition. Our space probes show that machines can be used to decide the truth or falsity of some propositions. For example, some of the data transmitted by the Viking Mars lander can be interpreted as asserting, “There is conclusive evidence of organic life everywhere on the surface of Mars’ is false.” (Can propositions be ranked in accordance with the sophistication of the machine needed to decide them?)

As soon as we state our desire to know if there are propositions which apply to a subject, we are in effect asking a question about the ontology of the subject, specifically, we are asking if the subject is one that can be subjected to scientific (or mathematical) study, i.e., if the concepts of the subject are Objects, as that term has been defined elsewhere in this book. “The discovery of the

logical type to which a puzzle-generating idea belongs is the discovery of the rules governing the valid arguments in which propositions embodying the idea (or any other idea of the same type) can enter as premises or conclusions.” — Ryle, Gilbert, “Philosophical Arguments”, reprinted in (Ayer 1959, p. 338). (Thought in passing: philosophy of science has typically based its theories on existing sciences, attempting (among other goals) to establish just what can be known (meaning, “known scientifically or logically”). Perhaps its results can be applied before the fact at the beginning of a new scientific field, if indeed the study of self-consciousness turns out to be that.)

What are some candidates for propositions about self-consciousness?

One class of such propositions is:

(1)  $x$  is capable of being self-conscious, where  $x$  can take as value any proper noun, including one that names a fictional character, as well as the name of the person, machine, or other entity asserting (1).

Let us consider the case, first, where  $x$  does not refer to the being asserting (1). We now confront the question, “Is it possible to decide the truth or falsity of (1)?” One way of approaching this question, is by asking if there are tests for self-consciousness.

## **Are There Tests for Self-Consciousness?**

In 1950, Alan Turing wrote a paper titled, “Can a Machine Think?” He proposed to answer the question by what has come to be known as the “Turing test”. In this test, there is an interrogator in one room, and a man (A) and a woman (B) in another room. The interrogator communicates with each via computer terminal only. The test consists of two parts: in the first, the interrogator attempts to determine, by asking any questions he wishes, which of A or B is the man, which the woman. This part is repeated several times, with the assignment of “A” and “B” being randomly chosen each time. In the second part, the man is replaced by a machine and the process repeated, only, of course, in this part, the interrogator attempts to determine which is the machine. If the interrogator decides wrongly in the second part about as often as he does in the first, then we say that the machine can think. Turing provides a sample of the dialogue that might take place in the second part:

“Q: Please write me a sonnet on the subject of the Forth Bridge.

“A: Count me out on this one. I never could write poetry.

“Q: Add 34957 to 70764.

“A: (Pause about 30 seconds and then give as answer): 105621.

“Q: Do you play chess?

“A: Yes.

“Q: I have K at my K1, and no other pieces. You have only K at K6 and R at R1. It is your move. What do you play?

“A: (After a pause of 15 seconds): R-R8 mate.”

Turing’s is a behavioral test of the machine’s ability to think. It ignores all questions of what the machine is made of. We ask, now, if there is a behavioral test for self-consciousness? Although Turing excluded human beings from taking the part of the machine in the second part of the test, I think it is important that we *not* exclude humans, or animals, from the self-consciousness test.

Although, in his paper, Turing attempts to answer most of the major objections to the idea that machines can think, he does not deal with the question of major variations in results among different interrogators, or even among the same interrogator at different times. The same variation may

well occur in any behavioral test of self-consciousness. Turing's test explicitly excluded the interrogator's knowing anything about the physical construction of the machine. If we adopt this restriction in our case, tests such as the following are excluded: "Self-consciousness in animals has been studied experimentally. In an interesting series of experiments, Gordon Gallup established that chimpanzees can come to recognize themselves in mirrors — and they recognize themselves *as* themselves too, as he demonstrated by putting dabs of paint on their foreheads while they slept. When they saw themselves in the mirrors, they immediately reached up to touch their foreheads and then examined their fingers. See Gordon G. Gallup, Jr., 'Self-recognition in Primates: A Comparative Approach to the Bidirectional Properties in Consciousness,' *American Psychologist* (vol. 32, (5), 1977, pp. 329-338)" — Hofstadter, Douglas R., and Dennett, Daniel C., *The Mind's I*, Basic Books, Inc., N.Y., 1981, p. 471.

We must also keep in mind that it is always possible to "hard-wire" certain responses into a computer. E.g., if we, as designers of a machine — call it JOE — to be submitted to self-consciousness testing, could always include within the machine's program, statements equivalent to the following:

"If interrogator asks, 'What is your name?', reply 'JOE'."

"If interrogator asks, 'How do you feel today?', reply, 'Just fine, thank you.'" etc.

Of course, the skilled interrogator will attempt to learn, by other questions, if in fact such replies are merely the result of answers hard-wired into the program.

Our remark above about different interrogators coming to different conclusions about Turing tests of the same machine, must be kept in mind here also. Consider, e.g., Joseph Weizenbaum's program ELIZA, developed at MIT in 1966. This program engaged in a question-answer dialogue with the user, the program simulating the responses of a psychiatrist. The following is an example: the user's input is given in upper-and-lower case type, the program's responses are given in all capitals:

"Hello.

"HOW DO YOU DO. PLEASE STATE YOUR PROBLEM.

"I want a good example for my book.

"YOUR BOOK.

"Yes, my book about computers. Can you help me?

"YOU'RE BEING SOMEWHAT SHORT WITH ME.

"I didn't mean to be.

"YOU SAY YOU DIDN'T MEAN TO BE.

"That's right. What do you know about computers?

"DO COMPUTERS WORRY YOU?..." — quoted in (Raphael 1976, p. 199)

A number of users thought the program was capable of acting as a substitute psychiatrist, thus attributing to it, we may presume, a certain degree of self-consciousness. "Actually it all works by syntactic trickery, as a more careful interrogation of the program would show." — *ibid*, p. 197.

Now we consider briefly those sentences of type (1) in which *x* refers to the speaker. If we in fact had a test for self-consciousness, would there be any difference from what we have already said? E.g., suppose a typical 1987 computer were programmed to output, and did output, the sentence,

(1') I am capable of being self-conscious.

Would this be a proposition, in particular, a false proposition? Suppose the computer were programmed to output, and did output, the sentence,

(1") I am not capable of being self-conscious.

Is this a proposition, and, in particular, a true proposition?

Now consider (1') and (1'') being uttered by the author of this book. Do the answers to our questions change?

## **Other Sentences That May Be Propositions About Self-Consciousness**

Is Descartes's famous sentence,

(2) I think, therefore I am.

a proposition about self-consciousness? Is it a proposition at all? If so, it must be either a logical or an empirical proposition.

Among the very large number of discussions of this sentence in the philosophical literature, the reader should see Jaako Hintikka's "Cogito, Ergo Sum: Inference or Performance?" in (Sesonke and Fleming 1965).

Another sentence is:

(3) You understand this sentence.

Is it a proposition about self-consciousness? Is it a proposition at all? Again we resort to a computer model. A few lines of the programming language Pascal are sufficient to make a typical 1989 computer read in such a sentence, and variations of it, and to reply, "Yes, I do" or "No, I do not". Does that mean the sentence is a proposition? Either of the two computer replies seems to imply that the machine "knows" what the word "understand" means in this case, and whether it really does or not is a matter which must be decided, and not necessarily by the self-consciousness test.

Suppose a person who does not understand English is given a piece of paper containing sentence (3) and, below it, two boxes, one labeled "Yes, I do." and the other labeled "No, I do not.". He is told, in his own language, to "try to read and understand the sentence" (something which he cannot do); he is told the meaning of the sentences labeling the boxes and is then asked to mark whichever box he feels is "appropriate", feeling free to guess or pick one at random. Is sentence (3) in this case a proposition?

Consider an entire book filled with nothing but repetitions of sentence (3).

Is the sentence, "This sentence is false," self-contradictory when no one is contemplating its self-contradictory nature?

Imagine that all self-conscious beings were eliminated from the universe, but that one of them had left behind a computer, powered by solar energy, which printed sentence (3) over and over and had no other capabilities. Would the sentence be a proposition? Is the sentence one that becomes true whenever it is perceived (understood), and is neither true nor false otherwise? What does "understood" mean here? If no self-conscious being speaks the language it is written in, but many of these beings see the sentence before them, and perhaps say, in their own language, "I don't understand that," is it true?

We may ask similar questions about the following:

(4) You are now reading this sentence.

(5) You are now sleeping.

The question whether (4) is a proposition when no one is reading it may remind us of the question whether the tree falling in the forest makes a sound when no one is there to hear it. Is (4) a proposition when someone who understands written English reads it? Is it a proposition when it is read by someone in the dream state described elsewhere in this book under “Hypnagogic Art”? In this state, one reads a text without being self-conscious — in fact, as soon as one becomes self-conscious, the text disappears. Is (4) false or merely meaningless when it is observed by one who doesn't understand English?

## Additional Thoughts

### Proving the Existence of Other Minds and of the External World

The obscurity and inconclusive nature of much Western philosophy seems to me in part to be due to the failure of each philosopher to squarely face the problem which self-consciousness presents to his thought. The importance of self-consciousness may well be philosophy's best-kept secret. Many philosophers write as though at times they regard man as a self-conscious being, at other times, not. Is it possible that Marxism would have been true if only man had not been self-conscious, i.e., if only he had been unable to understand Marxism's predictions about him?

In light of the above discussion, we may ask: is there a “rigorous” analysis of self-consciousness, so that we would be able to show that certain philosophical claims and questions were really reducible to fundamental propositions and questions about self-consciousness? Putting it another way, is there a correct, as opposed to an incorrect, way that self-conscious beings should write about self-consciousness?

For example, it seems to me that in many instances, Heidegger's term, “Being”, can be replaced by the phrase, “the World of the self-conscious individual”, where “World” is used in the existentialist sense of “a set of concerns” as, e.g., “the World of the mathematician”, “the World of the street poet”, “the World of the IBM executive”, as well as the World of a specific person, e.g., “the World of the author of this book”.

A traditional problem in philosophy is that of proving the existence of other minds.

What would a proof that self-conscious beings exist, or are possible, “look like”? Where would the proof itself begin, i.e., would it be meaningful to have “preliminary remarks” to such a proof? What kind of being would the preliminary remarks, still less the proof, be addressed to?

What kind of being is Heidegger's *Being and Time* addressed to? If the answer is “the kind of being which Heidegger calls ‘Dasein’”, then how much need be said about Being to such a being?

Why don't we precede every sentence we write with the sentence, “You will be able to understand the next sentence”? Why don't we precede every math proof with a proof that the given proof is correct?

“[Agrippa the skeptic] denies that anything can be proven, since every proof requires a previous proof.” (Quoted in Borges, “Avatars of the Tortoise”, *Labyrinths*, New Directions, 1964.)

In Lewis Carroll's “What the Tortoise Said to Achilles”, the tortoise argues that, in the following, A and B do *not* imply Z:

- (A) Things that are equal to the same are equal to each other.
- (B) The two sides of this Triangle are things that are equal to the same.

(Z) The two sides of this Triangle are equal to each other.

He argues that we must have an intermediate assertion, C:

(C) If A and B are true, Z must be true.

And then another intermediate assertion, and another, etc., with the result that we can never reach the conclusion, (Z):

(A) Things that are equal to the same are equal to each other.

(B) The two sides of this Triangle are things that are equal to the same.

(C) If A and B are true, Z must be true.

(D) If A and B and C are true, Z must be true.

(E) If A and B and C and D are true, Z must be true.

(F) If A and B and C and D and E are true, Z must be true.

.  
. .  
.

(Z) The two sides of this Triangle are equal to each other.

Here is an erroneous proof of the existence of other minds:

There exists a mind that can understand this sentence, and furthermore that mind is not the mind of the author of this sentence, and therefore there exists at least one mind besides the author of this sentence.

The error lies in the assumption that the sentence must necessarily have been created by another mind. (It could have been generated by a random selection of letters.) But even assuming that the sentence must have been created by another mind, a peculiarity of the above proof is that it must wait to be read in order to be true, whereas a typical proof in mathematics or formal logic is known to be true (barring errors) by its discoverer when he has written it down. The checking of the proof by other mathematicians or logicians is not normally *the proof*.

(In passing, I can't help pointing out that if any literate human being I know of, did come across such an erroneous proof as the above, whether in the sands of some uninhabited planet, or as part of the processing of a digital representation of noise, their first hypothesis would *not* be that the proof was the result of a random process, even though, logically, it could be. Certainly, in that case, they would have to attempt to estimate the probability that such a sequence of letters was produced by a random process. Then they would have to estimate the probability that such a process was in turn produced by a random process.)

Another traditional problem in philosophy is to prove the existence of things in the "external world":

"Kant calls it 'a scandal of philosophy, and of human reason in general' that there is still no cogent proof for the 'Dasein of Things outside of us' which will do away with any scepticism." — Heidegger, Martin, *Being and Time*, tr. by John Macquarrie and Edward Robinson, Harper and Row, N.Y., 1962.

"...The 'scandal of philosophy' is not that this proof has yet to be given, but that *such proofs are expected and attempted again and again*. Such expectations, aims, and demands arise from an ontologically inadequate way of starting with *something* of such a character that independently *of it* and 'outside' *of it* a 'world' is to be proved as present-at-hand. It is not that the proofs are

inadequate, but that the kind of Being of the entity which does the proving and makes requests for proofs has *not been made definite enough*. This is why a demonstration that two things which are present-at-hand together, can give rise to the illusion that something has been proved, or even can be proved, about Dasein as Being-in-the-world. If Dasein is understood correctly, it defies such proofs, because, in its Being, it already *is* what subsequent proofs deem necessary to demonstrate for it.” — *ibid*, p. 249.

“We cannot significantly question the reality of the external world, or deny that there is evidence of external objects in the testimony of our senses; for, to do so is simply to dissociate the terms ‘reality’ and ‘evidence’ from the very applications which originally did most to invest those terms with whatever intelligibility they may have for us.” — Quine, W. V., *The Ways of Paradox and Other Essays*, Harvard University Press, Cambridge, Mass., 1979, p. 229.

Our first question, in response to the demand for a proof of the existence of the external world, should be, *For whom is such a proof intended?* A being in the external world? But then why should the proof be necessary? All that should be necessary is to state that such a proof is desirable. QED. One might reply that the proof (or disproof) could be “private”. But a private proof is no proof at all — it is a misapplication of the word “proof”. We give proofs in order to convince others, as well as ourselves, of the validity of an assertion. So, a valid proof that the external world exists, is unnecessary. A valid proof that the external world does not exist, is merely a sequence of words which have been given the wrong label.

In any case, if we take it for granted that propositions can only apply to Objects, then the question whether there are propositions about self-consciousness boils down to the question whether self-consciousness is an Object. Some readers may be inclined to say that the above quotations from Heidegger constitute a negative reply to this question. If that is the case, does it follow that computer science efforts to design self-conscious machines are doomed? Or is self-consciousness rather a property of a machine in the same way that “beauty” is?

To a student of mathematics or symbolic logic, the centuries-long failure to find convincing proofs in philosophy — proofs which are as valid as mathematical proofs — not only for the existence of things outside of us, but also, e.g., of the existence of God, lies in the simple fact that neither the entities concerned, nor the rules of inference, are ever defined in a mathematically acceptable way. (Whether they can be so defined is another story.) A proof is a piece of mathematics or symbolic logic, no more and no less.

### **Sentences and Questions Which Merely Announce the Existence of Self-Consciousness**

*Thesis*: there is a class of sentences and questions which, despite their syntactic appearance, do no more than announce the existence of self-consciousness.

Consider, e.g., “Mathematics is nothing but the manipulation of symbols.” We ask, first, “Is this sentence a proposition?” If so, is it a proposition that becomes false whenever it is uttered by a person who understands its meaning? I.e., we may think, informally: “To a typical 1989 computer, mathematics may indeed be nothing but the manipulation of symbols, but of course the computer doesn’t know that (as far as we know)!”

Or consider the philosophical question attributed to Leibniz, “Why is there something rather than nothing at all?” We ask, first, “If there were nothing at all, could the question be asked?” Next we ask, as in the previous paragraph, “Could the question be asked by a non-self-conscious being?” The answer to the latter would seem to be no if we agree that one necessary constituent

of self-consciousness is the awareness of the difference between the self and the world (we should say, the awareness that two entities, self and world, exist).

(Digression:

“If there was a creation event, it had to have had a cause. This was Aquinas’s whole question, one of the five ways he established the existence of God. If you can find the first effect, you have at least come close to the first cause, and if you can find the first cause, that to him was God. As astronomers you can’t say anything except that here is a miracle, what seems almost supernatural, an event which has come across the horizon into science, through the big bang. Can you go the other way, back outside the barrier and finally find the answer to the question why is there something rather than nothing? No, you cannot, not within science. But it still remains an incredible mystery: Why is there something instead of nothing?” — Allan Sandage, quoted in Ferris, Timothy, *Coming of Age in the Milky Way*, Anchor Books, N.Y., p. 351.

End of digression)

Or consider the sentence, “Being is All.” Again, varying our wording, we ask, “What kind of being is required to utter such a sentence?.” Suppose Being weren’t All: what sort of being could conceive of that idea?

Or consider the sentence, “The world is nothing but a dream.” We ask: if it were nothing but a dream, would we know it?

Or consider the question whether we can know the thing-in-itself. We ask: is the question anything more than an announcement of self-consciousness? Or, putting it another way, what would be the nature of a being that *could* know the thing-in-itself? Would it be able to ask the question?

Or consider the endless discussions in philosophy over the validity of (Western) metaphysics. We ask if perhaps the “deep meaning” of metaphysics is simply the announcement, “Self-conscious beings exist!”, or, in other words, what lies beyond physics is the awareness that physics exists, i.e., that there is something else besides physics, i.e., that there are self-conscious beings. Thus the discussions about the validity of metaphysics can be interpreted as discussions about whether the subject of self-consciousness is a legitimate one for self-conscious beings, and if not, why not.

Someone may reply to the above that I am merely sloughing off certain difficult questions onto self-consciousness. But this reply already assumes a knowledge of what is properly “in” self-consciousness and what is not.

We may ask (and should ask), “Is there anything in existentialism other than the assertion that man is not an Object?” And in thinking of a reply, we should realize that the mere ability to understand the question speaks volumes about the nature of the being that asked it and which is attempting to answer it. (Redundant versions of some of these volumes have been laboriously produced by existentialist philosophers.)

Some philosophers (e.g., Hegel, Heidegger, Sartre) have gone to extraordinary lengths in order to analyze self-consciousness in such a way that its non-Objective nature is presented to the reader (does the reader need such a presentation?). Yet poets have been far more successful at such a presentation than any philosopher I know of. Consider, e.g., the eighth of Rilke’s *Duino Elegies*, which can be viewed in part as a meditation on the difference between consciousness in animals and human self-consciousness:

“Mit allen Augen sieht die Kreatur  
das Offene. Nur unsre Augen sind  
wie umgekehrt und ganz um sie gestellt  
als Fallen, rings um ihren freien Ausgang.  
Was draussen i s t, wir wissens aus des Tiers  
Antlitz allein; denn schon das fruehe Kind  
wenden wir um and zwingens, dass es rueckwaerts  
Gestaltung sehe, nicht das Offene, das  
im Tiergesicht so tief ist. Frei von Tod.  
I h n sehen wir allein; das freie Tier  
hat seinen Untergang stets hinter sich  
und vor sich Gott, und wenn es geht, so gehts  
in Ewigkeit, so wie die Brunnen gehen.  
Wir haben nie, nicht einen einzigen Tag,  
den reinen Raum vor uns, in den die Blumen  
unendlich aufgehn. Immer ist es Welt  
und niemals Nirgends ohne Nicht:  
das Reine, Unueberwachte, das man atmet und  
unendlich w e i s s und nicht begehrt. Als Kind  
verliert sich eins im Stilln an dies und wird  
geruettelt. Oder jener stirbt und i s t s.

Waere Bewusstheit unsrer Art in dem  
sicheren Tier, das uns entgegenzieht  
in anderer Richtung —, riss es uns herum  
mit seinem Wandel. Doch sein Sein ist ihm  
unendlich, ungefasst und ohne Blick  
auf seinen Zustand, rein, so wie sein Ausblick.  
Und wo wir Zukunft sehn, dort sieht es Alles  
und sich in Allem und geheilt fuer immer.

Und wir: Zuschauer, immer, ueberall,  
dem allen zugewandt and nie hinaus!  
Uns ueberfuellts. Wir ordnens. Es zerfaellt.  
Wir ordnens wieder und zerfallen selbst.  
Wer hat uns also umgedreht, dass wir,  
was wir auch tun, in jener Haltung sind  
von einem, welcher fortgeht? Wie er auf  
dem letzten Huegel, der ihm ganz sein Tal  
noch einmal zeigt, sich wendet, anhaelt, weilt —,  
so leben wir und nehmen immer Abshied.”

“With all its eyes the creature-world beholds  
the open. But our eyes, as though reversed,  
encircle it on every side, like traps  
set round its unobstructed path to freedom.  
What *is* outside, we know from the brute’s face  
alone; for while a child’s quite small we take it  
and turn it round and force it to look backwards  
at confirmation, not that openness  
so deep within the brute’s face. Free from death.  
We only see death; the free animal  
has its decease perpetually behind it  
and God in front, and when it moves, it moves  
into eternity, like running springs.  
We’ve never, no, not for a single day,  
pure space before us, such as that which flowers  
endlessly open into; always world,  
and never nowhere without no; that pure,  
unsuperintended element one breathes,  
endlessly knows, and never craves. A child  
sometimes gets quietly lost there, to be always  
jogged back again. Or someone dies and *is* it.

.  
.  
.

Did consciousness such as we have exist  
in the sure animal that moves toward us  
upon a different course, the brute would drag us  
round in its wake. But its own being for it  
is infinite, inapprehensible, unintrospective, pure, like its outward gaze.  
Where we see Future, it sees Everything,  
itself in Everything, for ever healed.

.  
.  
.

And we, spectators always, everywhere,  
looking at, never out of, everything!  
It fills us. We arrange it. It decays.  
We re-arrange it, and decay ourselves.  
Who’s turned us round like this, so that we always,  
do what we may, retain the attitude  
of someone who’s departing. Just as he,  
on the last hill, that shows him all his valley  
for the last time, will turn and stop and linger,  
we live our lives, for ever taking leave.”

— Rilke, Rainer Maria, *Duino Elegies*, tr. by Leishman, J. B., and  
Spender, Stephen, W. W. Norton & Co., Inc., N.Y., 1963, pp. 66-71.

“What Rilke thinks when he thinks the word ‘the Open’ can be documented by a letter which he addressed in the last year of his life (February 25, 1926) to a Russian reader who had questioned him about the eighth elegy. Rilke writes:

‘You must understand the concept of the ‘Open’, which I have tried to propose in the elegy, in *such* a way that the animal’s degree of consciousness sets it into the world without the animal’s placing the world over against itself at every moment (as we do); the animal is *in* the world; we stand *before it* by virtue of that peculiar turn and intensification which our consciousness has taken. [Rilke goes on,] By the ‘Open’, therefore, I do not mean sky, air, and space; *they*, too, are ‘object’ and thus ‘opaque’ and closed to the man who observes and judges. The animal, the flower, presumably *is* all that, without accounting to itself, and therefore has before itself and above itself that indescribably open freedom which perhaps has its (extremely fleeting) equivalents among us only in those first moments of love when one human being sees his own vastness in another, his beloved, and in man’s elevation toward God.’” — Heidegger, Martin, “What Are Poets For?” in *Poetry, Language, Thought*, tr. Hofstadter, Albert, Harper Colophon Books, N.Y., 1975, pp. 107,108.

There are philosophers who are so in awe of the mere fact that something exists rather than nothing at all, that they devote their entire lives to the attempt to bring the rest of mankind into the same state of awe as they are in. Hegel and Heidegger are examples.

But to be in awe of this fact is merely to be in awe of their own self-consciousness. For the un-self-conscious creature, something certainly exists, but there is no awareness, hence no awe, of the fact.

### **Being and Not Being Self-Consciousness**

As far as we know, a goldfish does not know it is a goldfish. We may imagine its life to be something like, “...Now everything is hungry...”; then food is seen and — we know, but the goldfish doesn’t — is eaten; “...Now everything is less hungry...But everything is growing dangerous!...”; we know, but the goldfish doesn’t, that a larger fish has threatened it, and that the goldfish has darted into a hiding place; “...Now everything is less dangerous...Now everything is growing dark...”; we know, but the goldfish doesn’t, that night is approaching.

“The discovery we have made that we exist...is called the Fall of Man.” — Emerson

“By [objectification] I mean the thing that is also frequently called ‘the hypothesis of the real world’ around us. I maintain that it amounts to a certain simplification which we adopt in order to master the infinitely intricate problem of nature. Without being aware of it and without being rigorously systematic about it, we exclude the Subject of Cognizance from the domain of nature that we endeavor to understand. We step with our own person back into the part of an onlooker who does not belong to the world, which by this very procedure becomes an objective world. This device is veiled by the following two circumstances. First, my own body (to which my mental activity is so very directly and intimately linked) forms part of the object (the real world around

me) that I construct out of my sensations, perceptions and memories. Secondly, the bodies of other people form part of this world...” Schroedinger, Erwin, *What Is Life?* and *Mind and Matter*, Cambridge University Press, Cambridge, England, 1944, 1958, 1989, p. 127.

“The material world has only been constructed at the price of taking the self, that is, the mind, out of it, removing it; mind is not part of it; obviously, therefore, it can neither act on it nor be acted on by any of its parts.” — *ibid.*, p. 128.

“For do not let us forget: to say...that the becoming of the world is reflected in a conscious mind is but a cliché, a phrase, a metaphor that has become familiar to us. The world is given but once. Nothing is reflected. The original and the mirror-image are identical. The world extended in space and time is but our representation (*Vorstellung*). Experience does not give us the slightest clue of its being anything besides that — as Berkeley was well aware.” — *ibid.*, p. 146.

“No single man can make a distinction between the realm of his perceptions and the realm of things that cause it since, however detailed the knowledge he may have acquired about the whole story, this story is occurring only once not twice. The duplication is an allegory, suggested mainly by communication with other human beings and even with animals...how on earth shall we decide that a common feature of all our experience is due to the constitution of our mind rather than a quality shared by all those objectively existing things?” — *ibid.*, p. 156.

We say: a self-conscious being knows that it exists (though it may not think of this all the time), whereas a non-self-conscious, or “minimally” self-conscious, being — e.g., a goldfish — does not know it exists, as far as we know.

We are inclined to attribute little or no self-consciousness to the lower animals. But perhaps *all* animals beyond a certain level are self-conscious — it is just that the selves of which they are conscious are not very advanced!

“Mammals, if they survived at all [assuming that the dinosaurs had not become extinct], would probably still be small creatures no larger than rats, and small size precludes self-conscious intelligence.” — Gould, Stephen Jay, “The Confusion over Evolution”, in *The New York Review of Books*, Nov. 19, 1992, p. 54.

(In passing, we have to ask why Gould believes that size alone precludes self-conscious intelligence, and not, e.g., size of brain relative to size of body? Or in other words, is there a theoretical limit to the *smallness* of a self-conscious being? An answer to that, plus an answer to the question how small a computer memory unit capable of storing a bit of information can be, would give us an answer to the question, What is the smallest amount of memory, as measured in bits, for self-consciousness?)

We can teach animals many things, but one thing we apparently cannot teach them is how to be self-conscious. Why?

We human beings, on the other hand, forget many things, but one thing we never forget is how to be self-conscious. Why? What would it be like to wrack your brain, trying to remember how to *do* this? Or would the whole problem really be that you had merely forgotten what the *word* “self-conscious” means?

“...between every given degree of light and of darkness, between every degree of weight and of absolute lightness, between every degree of occupied space and of totally void space, diminishing degrees can be conceived, in the same manner as between consciousness and total unconsciousness (the darkness of a psychological blank) ever diminishing degrees obtain. Hence there is no perception that can prove an absolute absence of it; for instance, no psychological darkness that cannot be considered as a kind of consciousness, which is only outbalanced by a stronger consciousness.” — Kant, Immanuel, *Prolegomena to Any Future Metaphysics*, “Second Part of the Transcendental Problem: How Is the Science of Nature Possible?”, sect. 24.

Suppose a drug or machine were invented which enabled a person to be tuned in and out of self-consciousness, so that, at the one extreme, he was in a dreamless sleep, and at the other he was fully self-conscious. At some point, might he not be in the state which Zen Buddhism strives for, i.e., the state in which one is fully aware of one’s surroundings, but not at all of oneself (as, for example, when one is fully absorbed in something, e.g. a film)? When a person is in this state, his World is exactly as it is when he is in any of his “normal” states, except that he does not know that his World is *his* World: there is no “I” present in this state. It is not that he no longer knows *who* he is, but rather that he no longer knows *that* he is. This state might also be the animal, or at least the pre-human, state, and the question must occur to us whether Buddhism may not have been an attempt to return man to that dimly remembered state.

If such a drug or device existed, would it be correct to speak of “propositions” about the states experienced?

It may be meaningful to speak of an individual as being highly self-conscious, but is such an individual necessarily “skilled” at entering and leaving, at will, various degrees of self-consciousness? Such a person would, e.g., be good at playing the game of not thinking of a white bear all day — he could in fact “win” this game most of the time. He would also be a good Zen Buddhist.

We are eyes on eye-stalks that are attached to inanimate objects, looking down and wondering if the object has free will, if we have free will.

## **Exercises**

*Exercise 1:* try, for brief periods of time, to get rid of your Self (the you which is at the center of your self-consciousness experience). Consider it an intellectual, not a religious, exercise. Your goal is to subtract your Self from the world while leaving everything else *exactly as it is*. Perception without a Perceiver. Try to do this without reading yoga books. You might begin by concentrating on the sound of your breathing while lying in the tub. Try to listen to it the way you would listen to surf on an idle afternoon, as something eternal. Or try concentrating, with eyes closed, on the sounds of your everyday world — dripping water, birds outside, traffic. Try to shut out words altogether (very difficult!). *Try to continually hear all (not just some, but all) the sounds around you.* Or try imagining that your surroundings are a scene in a movie in which no persons

## *Self-Consciousness*

are present. Try to lose yourself in the surrounding world so that you later realize that for a few moments your Self had disappeared.

*Related exercises:* (1) while sitting alone at home without radio, stereo or TV on, or while walking alone on a city street, try to perceive all sounds as having already occurred, as being played back on a recording made a fraction of a second earlier. (2) Try to see the world as if it had quote marks around it, i.e., try to experience the world as if it were what it is.

(Digression: Suppose a pill were developed that could immediately produce the Zen state of enlightenment, or *satori*. Would most Zen priests advocate the use of the pill, and if not, why not?)

*Exercise 2:* what, if anything, is wrong with the following “definition” of self-consciousness? Definition: (1) Think of yourself right now, as you read this — think of the being, the person, that is reading this. (2) You are in a self-conscious state. End of definition.

*Exercise 3:* Discuss (e.g., from an ordinary-language philosophy point of view) the following sentences: “I merely occupy my self-consciousness.” “My self-consciousness occupies the world.”

*Exercise 4:* We know what it means to say, “My soul occupies my body”, but attempt to find a meaning for “I occupy my soul”. What are the characteristics of a being that can *occupy* now this soul, now that one?

Suppose it were possible to become someone else for a few minutes or hours or days. Would the other person be able to tell that there were now two people who were being him or her, and, if so, wouldn't that destroy the purpose of the experiment? What would be the nature of a report on the experience of being someone else?

*Exercise 5:* Recalling Heidegger's definition of Dasein, “an entity which in each case I myself am”, imagine that there is an I which travels from personality to personality throughout the human population as well as throughout the population of other intelligent beings in the universe, and whenever the I occupies a given personality, that personality becomes self-conscious, considers itself unique, fundamentally alone in the world, etc. (At present, the I is occupying you.)

*Exercise 6:* try to view your own self-consciousness as an object, no more different from other self-consciousnesses than your hands are from other hands.

*Exercise 7:* will it ever be possible to assign “coordinates” to each human being’s self-consciousness, so that, given the coordinates of your own self-consciousness, you could look up the names and personal data of other persons whose self-consciousnesses were “close” to, i.e., very much like, yours? (“Who am I?” = “Where am I?”) If the question is nonsensical, state why. Otherwise, consider what it would mean to know that a person you had just met had a self-consciousness very much like yours. E.g., would the two of you frequently anticipate what the other was going to say?

*Exercise 8:* imagine that you knew there was an exact physical and mental duplicate of you living in an exact duplicate universe. (He or she, of course, would imagine this also.) How would this change your life? Would you feel less lonely, and if so, why?

*Exercise 9:* think of yourself as being a separate, possibly unique, individual. There is the you that is that individual but there is also a you that is thinking about that you. So there are at least two you’s. But there is also a third you that is thinking of those two you’s, so there are at least three you’s. But there is also a fourth you...

This exercise occurred to me in my youth — not as an exercise, but as a terrifying thought which seemed capable of making me lose my mind, because at times I could barely keep track of which of the you’s I really was. Are there other thoughts and questions which are capable of making us lose track of our everyday “I”? We recall Heidegger’s definition of Dasein: “Dasein is an entity which in each case I myself am.” — Heidegger, Martin, *Being and Time*, tr. John Macquarrie and Edward Robinson, Harper & Row, N.Y., 1962, p. 78 (p. 53 in original). Our goal here is thoughts and questions which remove from a particular Dasein, the “which in each case I myself am”.

“Consciousness is never experienced in the plural, only in the singular. Even in the pathological cases of split consciousness or double personality the two persons alternate, they are never manifest simultaneously. In a dream we do perform several characters at the same time, but not indiscriminately: we *are* one of them...

“How does the idea of plurality (so emphatically opposed by the Upanishad writers) arise at all? Consciousness finds itself intimately connected with, and dependent on, the physical state of a limited region of matter, the body...Now, there is a great plurality of similar bodies. Hence the pluralization of consciousnesses or minds seems a very suggestive hypothesis. Probably all simple, ingenuous people, as well as the great majority of Western philosophers, have accepted it.

“It leads almost immediately to the invention of souls, as many as there are bodies, and to the question whether they are mortal as the body is or whether they are immortal and capable of existing by themselves. The former alternative is distasteful, while the latter frankly forgets, ignores or disowns the facts upon which the plurality hypothesis rests. Much sillier questions have been asked: Do animals also have souls? It has even been questioned whether women, or only men, have souls.

“Such consequences, even if only tentative, must make us suspicious of the plurality hypothesis, which is common to all Western creeds. Are we not inclining to much greater nonsense, if in

discarding their gross superstitions we retain their naive idea of plurality of souls, but ‘remedy’ it by declaring the souls to be perishable, to be annihilated with the respective bodies?

“The only possible alternative is simply to keep to the immediate experience that consciousness is a singular of which the plural is unknown; that there *is* only one thing and that what seems to be a plurality is merely a series of different aspects of this one thing, produced by a deception (the Indian MAJA); the same illusion is produced in a gallery of mirrors and in the same way Gaurisankar and Mt Everest turned out to be the same peak seen from different valleys.” — Schrodinger, Erwin, “Epilogue: On Determinism and Free Will”, in *What Is Life?*

“In ‘Does Consciousness Exist?’ [William James] states that *pure consciousness* is ‘the name of a nonentity’ because, ultimately, it *cannot be experienced*. The experiencing ego is always ahead of itself; at the moment it is experiencing it is itself not-yet-experienced. It cannot be objectified because it is the *objectifier*. And for this reason, James claims — on the surest phenomenological grounds — that it does not exist as an *entity*, as a thing among things. But, clearly, it does exist, he admits as a *function* (or a process).

‘There is...no aboriginal stuff or quality of being, contrasted with that of which material objects are made, out of which our thoughts of them are made; but there is a function in experience which thoughts perform, and for the performance of which this quality of being is invoked. That function is *knowing*.’ (*Essays in Radical Empiricism*, pp. 3-4)

“Consciousness *as a function* is prior to the dichotomy of subject and object; it is subjective and objective both at once; it is ‘entirely impersonal ... pure experience’. One feels that what James primarily lacked for a more adequate solution to this problem was precisely the Husserlian and Sartrean distinction between the pre-reflexive (pre-personal, not-yet-reflected) consciousness and the fully reflexive, judging consciousness.” — Edie, James M., “Notes on the Philosophical Anthropology of William James”, in *An Invitation to Phenomenology*, Quadrangle Books, 1965, Chicago, Ill., pp. 128-129.

“I am going to close my eyes, stop my ears, extinguish one by one the sensations that come to me from the outer world. Now it is done; all my perceptions vanish, the material universe sinks into silence and the night. — I subsist, however, and cannot help myself subsisting. I am still there, with the organic sensations which come to me from the surface and from the interior of my body, with the recollections which my past perceptions have left behind them — nay, with the impression, most positive and full, of the void I have just made about me. How can I suppress all this? How eliminate myself? I can even, it may be, blot out and forget my recollections up to my immediate past; but at least I keep the consciousness of my present reduced to its extremest poverty, that is to say, of the actual state of my body. I will try, however, to do away even with this consciousness itself. I will reduce more and more the sensations my body sends in to me; now they are almost gone; now they are gone; they have disappeared in the night where all things else have already died away. But no! At the very instant that my consciousness is extinguished, another consciousness lights up — or rather, it was already alight: it had arisen the instant before, in order to witness the extinction of the first; for the first could disappear only for another and in the presence of another. I see myself annihilated only if I have already resuscitated myself by an act which is positive, however involuntary and unconscious...If I abolish this inner self, its very abolition becomes an object for an imaginary self which now perceives as an external object the

self that is dying away.” — Bergson, Henri, *Creative Evolution*, The Modern Library, N.Y., 1944, p. 303.

“It may...be possible to construct a computer memory that operates according to the parallel worlds view and not the Copenhagen view [of quantum mechanics]. Accordingly, a parallel world computer memory would be capable of observing itself even while the memory contained two parallel and quantum interfering streams of data — each stream in a parallel universe. Such a computer would probably use magnetic flux quanta instead of today’s chip memory. It would be macroscopic and not atom-sized, however. Consequently, although it would be a large object, it would follow the rules of quantum physics and not Newtonian mechanics.” — Wolf, Fred Allan, *Parallel Universes*, Simon and Schuster, N.Y., 1988, p. 60.

“The thing that I have the greatest difficulty in believing in, is my own reality. I am constantly getting outside myself, and as I watch myself act I cannot understand how a person who acts is the same as the person who is watching him act, and who wonders in astonishment and doubt how he can be actor and watcher at the same moment.” Edouard, in his journal, in Gide, Andre, *The Counterfeiters*, Vintage Books, N.Y., 1973, p. 71.

*Exercise 10:* Consider the sentence, “If you can think it, it isn’t True.” Recalling from the Preface that “True” means “true” in any sense that the word is used, are there ever occasions in which the sentence is appropriate? The sentence might be considered a vague metaphysical analogy to Goedel’s famous sentence in formal logic, “This sentence cannot be proved true.”

This exercise brings up an important idea concerning philosophical idealism.

“Once the idealist argument is accepted, I understand that it is possible — perhaps inevitable — to go even further. For Hume, it is not licit to speak of the form of the moon or its color: its form and color *are* the moon. Neither can one speak of the mind’s perceptions, inasmuch as the mind is nothing but a series of perceptions. The Cartesian ‘I think, therefore I am’ is thus invalidated: to say *I think* is to postulate the I, and is a *petitio principii*. In the eighteenth century, Leibniz proposed that in place of *I think*, we should say, impersonally, *it thinks*, just as one could say *it thunders* or *it flashes* (lightning). I repeat: there is not, behind the visages, a secret *I* governing our acts and receiving our impressions.” — Borges, Jorge Luis, “A New Refutation of Time”, in *A Personal Anthology*, Grove Weidenfeld, N.Y., 1967, p. 49.

Whenever someone argues that an intellectual construct is unnecessary, among the questions we should ask is, “How would the world be different if we were able to get rid of this unnecessary construct?” In this case, we ask, “How would the world be different if we were able to live in the belief that the mind is nothing but a series of perceptions, without an underlying, separate *I* which experiences these?”

Let us begin by considering certain cases where the abandonment of unnecessary intellectual constructs seem likely *not* to produce any changes in the world. Suppose, for example, that scientists on another planet had discovered all the physical laws that we have discovered up to the present, but that in addition, for cultural reasons, they postulated demons that put the laws into effect. Thus, instead of saying that the force of gravity is such-and-such, they would say, The gravity demon makes the force of gravity be such-and-such. And similarly for the electricity

demon, and the chemistry demon, etc. Now if some scientist on that planet were to propose that these demons are unnecessary, that nothing in science would be different if these demons were banished from science (although the scientific world might become a little lonelier), then I think most scientifically trained persons on the planet Earth would agree with him. Certainly scientific laws would not change.

Similarly, if medicine had, in addition to all known diseases, the symptomless disease, which medical students were told about and taught to watch for, I think that most doctors and patients would agree that little would change if this disease were no longer lectured upon in medical schools, or written about in medical books, or discussed with patients.

However, the elimination of the *I could* be an entirely different matter, depending on whether living in the belief that the mind is nothing but a series of perceptions, is like living in the belief that, e.g., our feelings are nothing but the result of electrochemical processes, or if it is closer to the Buddhist goal of experiencing the world selflessly, which I take to mean, without any sense of self, i.e., without any sense of one's being an individual, separate *I*, or if it is like neither of these.

Now, I argue that, if it is closer to the Buddhist goal, then for one thing, it would lead to the end of science, mathematics, and technology, not to mention the law, because it would remove self-consciousness from the world, which in turn would remove the possibility of posing, and solving, problems in these disciplines. It would remove the concept of *the world*.

“Jarring chord. The door flies open and Cardinal Ximenez of Spain enters flanked by two junior cardinals...

“Ximenez: Nobody expects the Spanish Inquisition. Our chief weapon is surprise ... surprise and fear...fear and surprise ... our *two* weapons are fear and surprise...and ruthless efficiency. Our *three* weapons are fear and surprise and ruthless efficiency and an almost fanatical devotion to the Pope...Our *four* ...” — *The Complete Monty Python's Flying Circus: All the Words*, Vol. 1, , Pantheon Books, N.Y., 1989, pp. 192-193.

A: “All there is, is sense-impressions!”

B: “...plus your awareness that that is all there is.”

A: “All right, all right: all there is, is sense-impressions plus my awareness of sense-impressions.”

B: “...plus your awareness that these two things are all there is.”

A: “Yes, yes, all right, OK: all there is, is...”

*Exercise II*: Discuss: (1) a brick falling a great distance; (2) an unconscious person falling a great distance; (3) a self-conscious person falling a great distance.

“Why was I born into this life?” is a question like “Have you stopped beating your wife?” What we are really asking is something like, “What biological and sociological factors produced the personality that I call ‘I’?” Now, at least, we can see what we are up against in trying to answer the question.

## *Self-Consciousness*

We must overcome our naive belief that because “I” for each of us represents something that seems to be infinite and unique and most certainly something that is not an object, it therefore must really *be* such a thing. This is as naive as believing that the images on the movie screen really “move”, or that the depth in a perspective drawing really exists. From the appropriate point of view, each personality — in particular, each self-consciousness — is “flat”.

Philosophers are able to withstand isolation better than most people because, unlike most people, they know that they always have for companionship the study of an inexhaustible subject, namely, their own self-consciousness.

Two friends has the philosopher,  
His lonely days to fill:  
The one is his self-consciousness,  
The other is his Will.

A friend who read this essay said that he felt that all such questions will sooner or later be answered through computer science research. If that should turn out to be so, how will it change our own behavior as self-conscious beings? Is it possible that one result of such research will be that we will understand for the first time that certain questions about self-consciousness can never be answered?

A vision: Whatever something is made out of, that in turn is made out of the first something. In the last analysis, there is no *outside*. To be outside of something is to be inside of something else.

Death is not the end of life, it exists now, is going on now, for each of us, on the other side of the wall between death and life. People there have picnics on sunny days on bright green grass near the banks of sparkling rivers. The only difference is that these people do not say, “At least we are alive!”

A man stands in a chamber holding a sponge ball in his hand. He sticks a pencil slowly into the ball. At that moment, behind him, the point of a giant pencil begins to emerge from the ceiling.

*Self-Consciousness*