

3

Persons and Popper's World 3: Do Humans Dream of Abstract Sheep?

RAY SCOTT PERCIVAL

Introduction

In the film classic *Blade Runner*, the story explores the notion of personal identity through that of carefully crafted androids. Can an android have a personality; can androids be persons? The title of the original story by Philip K. Dick is *Do Androids Dream of Electric Sheep?* The story suggests that our sense of being a person depends on our having memories that connect us with our childhood. In the movie, the androids are only a couple of years old, but have adult bodies. To complete them as persons they are given simulated memories of childhood. Some psychiatrists have decided that even humans dream only of electric sheep.

Modern psychiatry is premised on the reduction of the human person to a complex set of chemical states or processes of the brain. This is implied by its conception of mental problems, or fundamental life problems, as mental diseases. Psychiatry is also committed to a now refuted deterministic view of the physical world, and hence of people's life problems. Roughly speaking, determinism asserts that every event has a cause. More precisely, determinism asserts that any event can be explained in any level of detail given the relevant laws and a sufficiently precise statement of the initial conditions. If you combine this with physicalism, then all causes are physical causes. For psychiatry the world of the person is, therefore, a world closed off from any other type of influence outside the world of the physical.

I need to clarify my claim about psychiatry's philosophical commitments. The ubiquitous use of Freud's talking therapy that attempts to explore and solve the person's problems through a long series of in depth conversations about the person's memories, desires, conflicts, anxieties, and so forth, does not automatically suggest a commitment to a determinist physical reductionism. However, this type of reductionism characterises Freud's original metaphysical programme. Freud confidently expected progress in brain physiology to achieve a complete reduction and for this to then allow both chemical and surgical therapies to take over from what he regarded as a stopgap method.

I would like to illustrate just what this type of reduction would mean in terms of interpreting people's mental life. Suppose a person loves life, but has also adopted a theory (such as a religion or world-view) that seems to him to imply that human life is base, disgusting, or immoral. After many sleepless nights and depression, he decides to take his life. Psychiatry completely ignores the abstract aspect of the case by saying that this person took his life because of some yet undiscovered lesion in his brain. To admit that his theory of human life had any influence would be to open up the deterministic physical world into which psychiatry has placed all humans. From this perspective, the human being is hence no longer a person but a machine that has gone wrong.

The Epistemology of Coping with Life

I want to suggest that Szasz's position is lacking a strong epistemology and is therefore unnecessarily open to attack. Szasz suggests that the term "life problem" more accurately captures the phenomena that the term "mental illness" is meant to denote. I think that Szasz's emphasis on life problems suggests that the most appropriate epistemology for Szasz's perspective can be found in the work of Karl Popper.

As Popper said, "All life is problem solving." He meant this in the most general and abstract way, so that all life is covered by this formula: from the humble bacterium seeking out better conditions of warmth, and so forth, to the highly sophisticated scientist trying to unravel the explanation behind some wonderful phenomenon. In dealing with the problems we encounter in life, we adopt, shape, create, and abandon a host of theories, arguments, plans, and strategies in our attempts to solve or avoid them. This is most powerfully described in terms of a conjecture and refutation model. Popper argued that science should be a matter of different scientists advancing competing bold guesses about the world, guesses which are then subjected to unremitting criticism in the

hope that they may weed out the false theories and be left with those that are at least closer to the truth. Thankfully, for us, science has often managed to achieve this ideal.

An analogous model applies to the way we live. We actively try on different lifestyles, approaches, world-views, habits, and so forth, for size, testing them against criteria and standards (such as truth, beauty, moral goodness) that we have adopted or created or have genetically inherited (the need for warmth and food and human contact). The extent to which this is a deliberate and systematic enterprise varies between individuals and it may be more readily practised systematically only in the more developed countries, but its form can be discerned even in the most conservative or traditional societies and the most inept, slothful individual. The process is analogous to the evolution of organisms and to the development of science in so far as there is a population of variants, some of which meet the pressures of selection and some that do not. It has a greater similarity with science as far as language plays a key role through the formulation and arguing about world-views and the myriad less grandiose theories that the person finds important.

Popper proposed the following schema for the most abstract account of problem solving:

Problem → Tentative Theory → Error Elimination → New Problem

I want to say that persons are partly and actively constituted by the theories they have about themselves and the world. Persons actively give themselves unity, individuality, and continuity partly by a web of theories, conceptions, problems, arguments, plans and other abstract non-physical things that they have created, adopted, shaped, and adapted for themselves through life. This web of abstract entities makes a difference to what people do and hence opens up their world to the non-physical. Popper's arguments for the existence of three different types of classes of things, World 1, World 2, and World 3, and how they interact with one another help to bolster the rich conception of the person and defend it against the chemical control imposed by the state.

False Theory versus Category Mistake

I agree with Szasz that humans have life problems, but not mental diseases. The medical establishment has overlooked the fact that all life is problem solving, and it is by no means obvious that all people would

produce the same solution, or solutions that deviate only slightly from the norm, or that non of these solutions might be undesirable from a moral point of view. "What should I do?" is a question we face anew every day. Should I marry? Why should I be good? Why should I conform to what others do or say? What is more important: individual achievement in science through "obsessive" devotion or raising a family? Should I grow up? If not, how do I avoid doing so?

The word "disease" is defined within medicine as tissue damage or a condition conducive to damage. However, the mind is not tissue, nor a purely physical state of bodily tissue. Szasz argues that to apply the adjective "disease" to the mind is to incur a category mistake, an expression popularised by the philosopher Gilbert Ryle. It is like saying that numbers are red, or that pain is hexagonal. This approach has the strength of clarity, but it is vitiated by the fact that our language and its categories are a reflection of our theories about the world, and thus change with the advance of our understanding. Hundreds of years ago it would have seemed to be a category mistake to say that whales are mammals, but now we know that whales are in fact mammals, not fish. Once we classed mushrooms with plants; but now we regard mushrooms as belonging to the class of fungi. One could even envisage a sensible use of the expression "pain is hexagonal". We could interpret this phrase as describing the shape of the area of skin affected, for example.

Different theories carve up the world in different ways.

Popper's Worlds 1, 2, and 3

A stronger argument for the myth of mental illnesses attacks the theory of reductionism that lies behind the confusion of these different categories.

Popper's argument for dualism is the strongest case against the reductionist view. Popper argues that there are at least three radically different classes of thing. More concisely, there are three worlds. World 1 is the world of physics. It includes rocks, stars, protons, computers and biological bodies. It also includes the worlds of chemistry and biology. World 2 is the world of our conscious selves, dreams, hopes, pleasures, pains—the world of psychology. World 3 is the world of abstract products of the human mind. It is the world of numbers, theories, arguments, and problems. It also includes works of art and music.

Popper's World 3 is like Plato's world of forms, but has important differences. Plato's world of forms is a collection of eternal, perfect abstract concepts, like beauty, the circle, the good, and so on. In contrast, Popper's World 3 is the creation of the human mind. It contains every-

thing that Plato's world contains but also contains theories, arguments, problems, and works of art and music. It also contains erroneous theories, invalid arguments and other imperfect abstract productions. It retains the autonomy of Plato's world of forms, in that once a World 3 object, like the natural numbers, has been created, it develops a life of its own with its own laws and relationships that are independent of our psychology. For example, once the natural numbers had been created, it could then be discovered that prime numbers existed and this then brought up new unforeseeable problems, such as "is there a highest prime number?" and "do the prime numbers continue to get more scarce as we look further along the sequence of prime numbers?"

Many philosophers are upset by the use of the plural word "worlds," so let us be clear that they are all simply domains within the one world, that which we call the universe. Popper's three worlds could have been called Domain 1, Domain 2, and Domain 3. Some things such as books belong to both World 1 (on account of the fact that books are physical objects) and World 3 (on account of the fact that they contain abstract objects like theories and arguments).

The Reality of World 2

I would like to briefly state my assumption regarding the reality of a non-physical mental domain.

There are active self-conscious minds. The existence of a mind or self is dependent on the brain, but the mind is the pilot of an important range of brain processes. Its evolutionary function was the integration and co-ordination of activities of the brain and the body for the benefit of survival. However, the mind has developed a life of its own in some respects and some of its goals are independent of survival (for example, searching for a solution to an abstruse mathematical problem.) This is not to undermine the theory of evolution, since organs originally used for one function are often used for new functions later. The philosopher A.J. Ayer once said that the problem with radical physicalism is that it requires one to feign anesthesia. The hypothesis of minds explains a whole range of phenomena that cannot be satisfactorily explained simply by brain processes. In this respect, the hypothesis is on a par with the postulation of unobservable atoms to explain the structure of macroscopic objects, so the fact that the minds of other people cannot be directly observed is irrelevant. Moreover, even though it is not quite as open to falsifying tests as the atomic hypothesis or other physical theories, it can be tested.

Sophisticated and Ordinary Cases of World 3 Influence

The designer of a bridge may become deeply depressed if a fault in his calculations for the design leads to a fatal collapse. (The calculations were wrong only relative to the facts of mathematics, which are clearly not chemical or physical. Thus, the builder's behaviour in following the faulty design is not caused simply by his chemistry.) A mathematician may experience life-long frustration at not being able to derive a whole section of maths from a consistent set of axioms. (A logical inconsistency is not a physical or chemical state, process or relation. Hence, the mathematician's frustrating life-long problem is not a product of his chemistry.)

Someone may dismiss the case of the mathematician in search of the properties of prime numbers as irrelevant to the day-to-day thinking of people, but there are innumerable examples from everyday life. Five people out on the town each have 20 dollars, a total of 100 dollars. They all want to go to a Cantonese restaurant for a meal. When they get there, they find that the minimum charge for the five of them would be 150 dollars. Therefore, they decide not to order the meal there. It is a property of the natural numbers that 150 is greater than 100. Moreover, this is clearly not a physical fact; it is a mathematical fact. The reductionist is asking us to believe that the decision of the group not to eat at the restaurant could have nothing whatsoever to do with the fact that 100 is smaller than 150. A little thought will make it obvious that our life is full of instances of our interacting with and making use of abstract things, laws and relationships.

It is astounding that nearly the whole of psychiatry and even psychology implicitly denies any influence in peoples' lives to the existence of plans, designs, theories, numbers and logical arguments and the various non-physical relations that exist within and between these entities. In many cases abstract structures are simply neglected (a recent example would be the work of Antonio Damasio).

Objections

I intend to confine my defence of the person to an attack on one prominent assumption of cognitive science, the idea that the mind can be reduced to a computer program.

It has been argued that the autonomy of World 3 can be fully accounted for by a reduction of World 3 to technology (Levinson 1993). The

most popular version of this is that computer hardware and programs can do all the explanatory work that World 3 is meant to do.

A Technological Version of World 3

A number of attempts have been made to reduce World 3 to psychological or physical states, all of which founder on the infinite richness of at least some World 3 objects. One bold attempt was made by Paul Levinson in his *Mind at Large: Knowing in the Technological Age*.

Levinson argues that technological products, for example a humble nail, consist of a union between World 3 and World 1, since it is a physical object that embodies certain theories (presumably to do with how and for what it can be used). So far Popper would agree. But Levinson says that the autonomy of technology itself gives us all the autonomy that Popper sought in World 3 without our having to concede the existence of unembodied ideas. We have computers and other machines that function quite independently of us once they have been created. Even more fancifully, machines may supplant humans and become the next vehicles for the replication of what Dawkins calls memes.

However, the autonomy of World 3 goes far beyond the autonomy of that part of it that is embodied in technology. The idea that World 3 could be reduced to technology is similar to the idea that World 3 is simply the total library of objective knowledge. This is a suggestive metaphor, but it is also a very misleading error. Think of a theory that gets written down in a book. Some of its implications may be worked out and also written down. Now think of the total class of all the implications of this theory that will ever be worked out and embodied in writing. This perhaps vast amount of written material will still not exhaust the theory's logical content.

The Unfathomable Logical and Information Content of our Objective Theories

One of the strongest arguments for the independence of World 3 from psychology is based on the analysis of a theory's logical and information content. It can be shown that a scientific theory—a typical World 3 object—has an infinite information content. Information here is identified with what a theory denies or rules out. Expressed roughly, if I say that it will rain on at least one day next week, I convey less information than if I say it will rain only on Wednesday, because the second sentence

rules out more possibilities. Now a scientific theory such as Newton's rules out not only Einstein's theory, but also an infinite number of other possible theories. Newton's mind obviously did not contain a representation of Einstein's theory, let alone most of the other theories that his theory rules out.

In the *Logic of Scientific Discovery* (1934), Popper put forward the idea that a statement says more the more it forbids. Carnap, accepting Popper's suggestion, defined the assertive power of a sentence as the class of possible cases it excludes (Carnap 1942, p. 151). Carnap attributes it to Wittgenstein, an attribution he later explained as an error of memory. Later Popper (1974) reformulated the intuitive idea in terms of theories rather than possible cases, of both high and low universality. The information content is then defined as the class of all those statements that are logically incompatible with the given theory. Thus since Einstein's theory contradicts Newton's theory, Einstein's theory is part of the information content of Newton's. Newton could hardly have known this, and so it could not have been part of his psychology. Furthermore, there are an infinite number of unknown theories that form part of the information content of Newton's theory, and indeed of any empirical theory.

The argument for the infinite logical content of a theory t can be put thus. Suppose an infinite list of statements that are pair-wise contradictory and which individually do not entail t : a, b, c, \dots . Then the statement " t or a or both" follows from t . The same holds for each and every one of the statements in the infinite list. Since the statements in the list are pair-wise contradictory one can infer that none of the statements " t or a or both," " t or b or both," etc., is interderivable. Thus the logical content of t must be infinite.

The proof of the assumption that no pair of the statements " a or t or both", " b or t or both," etc., are interderivable is as follows. " b or t or both" follows from " a or t or both" if and only if the theory t follows from " a and non- b ." But because a and b contradict each other, " a and non- b " says the same as a . Thus " b or t or both" follows from " a or t or both" if and only if t follows from a , which by assumption it does not.¹

This in itself is not so important, but when combined with the idea of information content, the two notions produce some very interesting ramifications. As Popper shows, when we combine this result with the idea of logical content we obtain a parallel result, for if Einstein's theo-

¹ This proof is due to David W. Miller. See footnote 18 in *Unended Quest*.

ry E is part of the information content of Newton's theory N then Non- E is part of N 's logical content. Thus both the logical and information content of theories consist of an infinite number of non-trivial consequences. As Popper says, it follows that the task of understanding a theory is infinite.

As Popper used to say, we never fully know what we are talking about. Expressed more generally, when someone creates a theory he creates an object whose properties transcend his psychological make-up.

Barrow and Tipler estimate that the information storage capacity of the human brain is between 10 to the power 10 and 10 to the power 15 bits, with the lower figure assuming that each brain cell stores on average 1 bit. While a colossal figure, this is clearly smaller than the infinite content

The Causal Potential of Logical Standards

Cognitive science, which tries to model the way humans think simply in terms of brain states or computer programs, has yet to come to terms with the causal effectiveness of logical standards. A physical brain state cannot logically contradict a theory, but the logical contradictions between Einstein's and Newton's theories obviously made a difference to the thought of scientists. We know this independently of being able to supply an adequate theory as to how contradictions do make a psychological difference. The same point can be made in connection with technology. In explaining why an engineer rejects a proposed building project (that if adopted would have created a dangerous building) because he noticed an error in the reasoning that it was based on, we have to take into account two things:

(a) the engineer's knowledge of logic and mathematics (perhaps describable in terms of dispositions to carry out certain algorithms), and

(b) the objective fact that there was an error in the reasoning to notice. But this latter fact is neither a physical nor a psychological fact.

I think that one of the most challenging problems is to explain how standards can influence our thought. It cannot be a logical relationship between the standard and the psychological state, but there must be some patterned relationship between the logical relationship and the psychological states. This problem is connected to what has come to be called the problem of the empirical basis of science.

But the point I want to make is that current cognitive science is forced to say that the discovery of a logical contradiction never has anything to do with its actually being a contradiction. It cannot explain the

psychological impact that the discovery of an error in reasoning can have on us because a contradiction qua contradiction is impotent. For the cognitive scientist, performing an inference validly or discovering a logical error is either an accident of following certain conventional rules that one has been taught or a mysterious pre-established harmony.

As far as computer models are concerned, we must appeal to logical standards in order to make computers perform logical operations properly; we do not appeal to computers to judge logic. After all, computers break down. In the face of a global computer breakdown caused by a computer virus, we would still have recourse to the notions of validity and invalidity.

There is some truth in the idea that we can use computers to judge logic that must not be confused with the idea that computer programs can constitute validity. We can program a computer to perform according to a given set of axioms and inference rules. We can instruct it to draw out implications to see if any contradictions appear. If they do we can say that the putative logic is in fact invalid. But we would be appealing to an independent standard of validity. In an important sense, the computer is just a glorified pencil that helps us perform and check our inferences and calculations.

You can set up so-called "logic gates" in a computer in order for it to perform "logical inferences". But these structures and operations are only called logical because we interpret them so. The action of electrical impulses in a computer is an all or nothing phenomenon. Because pulses are precisely timed, even the absence of a pulse can be interpreted as a signal. When we want a set of possible combinations of signals to make a logic gate, the signals are interpreted as true or false (true = presence of a pulse; false = absence of a pulse). You can then make a logic gate for each of the logical operations: conjunction, disjunction, implication, and so forth. Each logic gate will be defined by what may be called a pulse-analogue of a propositional truth table.

It is clear that the action of a computer has to be suitably interpreted before we can use it for logic. Indeed, a great deal of logic and mathematics is used in interpreting the action of computers to make them useful tools of our reasoning. The more general point that any structure supporting a repeatable process involving the right conditionalities can be interpreted by us as a "logic gate" and as performing a "logical inference". Whether these interpreted processes can be put to any use is another matter.

World 3 as Linguistic Conventions

O'Hear claims that Popper's World 3 is not needed because we can account for the objectivity of World 3 by referring to linguistic conventions. We are simply drawing out the consequences of a set of rules. Of course, some of them may be unintended and unforeseen, but there is nothing more to what we are doing.

However, the first person to discriminate between a valid and invalid argument was not simply applying a set of conventional rules (or manifesting a set of dispositions) that he had been taught. By what convention was the first valid argument a valid argument?

There is another fundamental objection to O'Hear's view that derived from Gödel. Kurt Gödel showed that we cannot set down once and for all a set of rules that will tell us all the valid rules of inference. There will always be some valid rules of inference that remain undiscovered and not even a consequence of our current set.

Conclusion

My intention in this paper has been to argue that Szasz has left his position unnecessarily open to attack. Szasz has failed to supply an epistemology and a sufficiently elaborate philosophical case to defend his thesis about the myth of mental illness. A great deal has been written on the relation between mind and body, and it is not possible for me to cover even a significant amount of the debate. I have only been able to expound the relevant parts of Popper's epistemology and ontology and offer some introductory defence of this perspective on the mind-brain problem.²

² I am grateful for criticism and moral support from my wife Tamara Lynn Schreiber, my stepson Jacob Schreiber, David Barker, Patrick F. Murphy, and David Ramsay Steele.

REFERENCES

- Bartley, W.W. 1985 [1973]. *Wittgenstein*. La Salle: Open Court.
- . 1990. *Unfathomed Knowledge, Unmeasured Wealth: On Universities and the Wealth of Nations*. Chicago: Open Court.
- Carnap, R. 1942. *Introduction to Semantics*. Cambridge Massachusetts: Harvard University Press.
- Dick, Philip K. 1968. *Do Androids Dream of Electric Sheep?* Garden City: Doubleday.
- Levinson, P. 1988. *Mind at Large: Knowing in the Technological Age*. London: JAI Press.
- Penrose, R. 1989. *The Emperor's New Mind: Concerning Computers, Mind, and the Laws of Physics*. New York: Oxford University Press.
- Popper, Karl R. 1959 [1934]. *The Logic of Scientific Discovery*. London: Routledge.
- . 1963. *Conjectures and Refutations*. London: Routledge.
- . 1972. *Objective Knowledge*. New York: Oxford University Press.
- . 1976. *Unended Quest*. London: Fontana. Originally published as the "Intellectual Autobiography" in Schilpp 1974.
- . 1982. *The Open Universe*. Volume II of the *Postscript to The Logic of Scientific Discovery*. London: Routledge.
- Popper, Karl R., and John Eccles. 1977. *The Self and Its Brain*. London: Routledge.
- Ryle, G. 1949. *The Concept of Mind*. London: Hutchinson.
- Schilpp, Paul A., ed. 1974. *The Philosophy of Karl Popper*. La Salle: Open Court.
- Szasz, Thomas S. 1974 [1961]. *The Myth of Mental Illness: Foundations of a Theory of Personal Conduct*. Revised edition New York: Harper Row.
- . 1996. *The Meaning of Mind: Language, Morality, and Neuroscience*. Westport: Praeger.
- . 1974. *Ceremonial Chemistry: The Ritual Persecution of Drugs, Addicts, and Pushers*. Garden City: Doubleday.