Chapter 14
Persons

§14.1 Minds and Bodies
If reductionism is false, we can hope to give an un-mysterious account of the interaction between minds and bodies that preserves the autonomy of the mental. At present we have very little idea of how a pattern of retinal stimulation evokes a mental image of, say, a snake in the jungle, or how a decision to attack it, or to run away, results in the corresponding action. But we may progress. I see the snake, and am inclined to have a go at killing it, thereby impressing my mates; but then I think of my responsibilities as the sole breadwinner, and deem discretion the better part of valour. We can imagine how an innate response produces the adrenalin that makes the first response possible, and how established neural patterns bring to mind affection and concern for the family in the future. In each case our understanding will be of the “How Possibly?” kind, not the “Why, necessarily”; if the physiology were to dictate the sequence of mental processes, they would be wooden and inflexible, quite unlike the variegated responses of our actual life. But once we no longer have a metaphysical argument for overall reducibility which would justify a procrustean strait-jacketing of mental activity into rigid patterns, the claim that it can be done because it must be, loses all plausibility.

At the physiological level no intelligible pattern emerges; a number of physiological states succeed one another. The way they develop is perfectly compatible with the principles of physiology, but other courses would have been equally compatible. Granted the first-personal concepts of consciousness, we can see the physiological states as instantiating the conflicting considerations present to my mind as I wrestle with them, and finally decide what to do.
I can make sense of the broad, first-personal picture, but lose my bearings as I descend to more limited processes at a lower level.

Descartes’ dualism required a causal interaction between fundamental substances of a radically different kind: we have not found any fundamental substances in our search for the Ultimate Reality; the one reality we have found is amorphous and lacks the thisness that is the characteristic mark of substance. The interactionism outlined here is an explanatory interactionism, suggesting how different explanations can co-exist and interpenetrate with each other, by virtue of their asking different questions at different levels of complexity of the same sort of stuff. The me whose sensations I am conscious of, and which carries out the decisions I make, is a material me, composed of the same matter as in the cells of my body and the oxygen molecules I breathe in: but it is not just that matter that happens to constitute my body now; my body has a history, which involves many material particles now dispersed far from where my body is, and would be the same body if largely composed of different matter from that of which my body is currently composed. To focus too much on my current components would be to stop focusing on me. We cannot replace asking about me by asking about my components, though we can see how answers to questions about my current components can be fitted into answers to wider-ranging questions about me generally. Although sometimes answers to questions about my physiology, biochemistry, chemistry, or physics, will have an important bearing on the decisions I take, and what I manage to achieve, for the most part they are irrelevant, since if one process was not available, another would do just as well. Minds and the persons they embody are exceptionally independent of circumstances, being able to achieve their purposes by some means or other in spite of adventitious obstacles and difficulties.

Aseity

Just as men are more independent of their surroundings than other animals, so minds are more independent still, have greater aseity, and are correspondingly more real.
It is the principle of Aseity. Independence of circumstance confers ontological status. Minds, embodied in persons, can be called substances, not because they are ultimate non-composite entities, but because they are referred to by key terms we use in explanations of a sort extremely in our discourse. Although there is an indefinite plurality of possible explanations, explanations of why someone did something predominate and differ from those put forward to account for natural phenomena. Mental deliberation, as we have seen, has a “dialogue” logic of prima facie cases, objections, rebuttals, and conclusions subject to ceteris paribus clauses: the natural science generally, and physiology in particular, aims to have a “monologous” logic of necessary and sufficient conditions which ill accords with the process of ratiocination, and precludes there being any simple match between the physiological stimulus of seeing the snake and ensuing response, and the thoughts that actually went through the man’s mind. It is reasonable, therefore, to draw this major distinction among the many different sorts of explanation, and re-introduce dualism, a dualism of topics, based on a dualism of questions, rather than a dualism of substance; that is to say, among the many different types of question we ask, the biggest distinction is between impersonal questions from an external point of view, and those about persons from a personal point of view. Answers to the latter typically involve a vicarious first-personal stance, in which possible reasons for actions are canvassed and assessed, in a situation seen, as it were, first-personally by the person concerned. To do this, it is not enough to stand in another’s shoes; we have to try to get inside his skin, and feel as he does, and see with his eyes. We deem him to be a conscious being with a viewpoint peculiarly his own.

### Dualism

We reject Cartesian dualism which holds that bodies and minds are just two different kinds of fundamental substance, but Descartes was not wrong in picking out minds as what a large part of our discourse was about.

Dualism is not, fundamentally, ontological, but a dualism of stand-point, the (first-)personal and the impersonal, based on a fundamental difference in types of explanation.

---

1 See above, §1.5.
§14.2  Consciousness

Consciousness has long been a problem for philosophers. It is a difficult concept, or, rather, bundle of concepts, but has been made more difficult by philosophers approaching it from the wrong direction. If we start with the impersonal, tenseless view from Nowhere, we find it difficult to construct a plausible concept of consciousness, just as nineteenth-century scientists found it difficult to reinsert the arrow of time into a physics from which it had been filtered out. Often the correct response to a philosopher who demands a convincing account of consciousness, is like that of the yokel who, asked for directions by a lost motorist, replies “If I wanted to get to Dunchideock, I wouldn’t start from here”. Philosophers already have some concept of consciousness. Rather than trying to explicate consciousness in purely impersonal or third-personal terms, we should remind them of their first-personal knowledge of themselves, and encourage them to make use of resources already available to them if they want to gain understanding of other men and some other animals.

Theories of meaning are largely to blame for the philosophers’ self-imposed blindness, who tell themselves that the behavioural criteria for consciousness are all that can be at issue, because, when we say something, we can only mean what the grounds are for our saying it. That this is false in the case of consciousness can be seen if we consider the gruesome case of patients who were administered an ineffective anaesthetic together with an all-too-effective dose of curare, which completely paralysed them, so that they were able to feel the surgeon’s knife without being able to make it known to him that they were still conscious. They were conscious, though there were no signs of consciousness. Hence, consciousness is not constituted by the overt evidence of being conscious.

The conclusion has been resisted by some philosophers in the grip of a theory of meaning which leads them to argue that it is only by reason of the patients’ subsequent testimony that we know of their terrible experiences. If the patients had died, or if they had been given an amnesic drug that made them forget the horrible things they had undergone, then, so the argument runs, there would have been no fact of the matter, beyond the evident absence of overt indications of consciousness, and we should have no warrant

---

2 See below, this section p.420.
3 See above, ch.1, §1.2.
for saying that they ever were conscious of the operative procedures being performed on them. But this is absurd, as well as callously inhumane. I should view entirely differently the prospect of an operation in which I was going to be genuinely unconscious from one in which I was going to suffer all the agonies of unanaesthetized surgery, even if later I was going to die or be made to forget it. As with verificationist arguments generally,\(^4\) so with consciousness; the criteria for ascribing consciousness are one thing, but what we mean when we ascribe consciousness is something more.

Our concept of consciousness is not clear-cut, but is, rather, a cluster of concepts, with a corresponding multiplicity of criteria—those used by an anaesthetist dealing with a human patient, those used by a biologist deciding to anaesthetize an organism before vivisection, those used by an ecofreak, asking a tree's permission before cutting it down for fuel. With human beings we distinguish the questions whether someone is aware of something, and whether he is self-conscious. It is easy to confuse these different senses. Often the best safeguard is to ask what the opposite is, what consciousness is being contrasted with. Often we start by ascribing on the basis of limited criteria the full consciousness that is characteristic of personhood, and subsequently have to modify it, when it becomes evident that other criteria are not satisfied. Our ancestors were animists, ascribing personality to trees and hills and springs, but we back-track, allowing only that trees are alive. Homeostatic behaviour is a necessary, but not a sufficient, condition of something's being conscious. We find it extremely implausible to ascribe consciousness to totally inert minerals—even the ecofreak does not ask the coal's permission before burning it. But the homeostatic behaviour of plants is not enough to make us think they are conscious; and the doubts extend to the behaviour of animals moving away from noxious environments and towards food supplies, or engaging in reproductive activity, which may seem at first sight to be instances of conscious behaviour. But we withdraw the ascription of consciousness, if subsequent investigation shows a machine-like insensitivity to the real needs of the actual situation. The apparently well-thought-out behaviour of a sphex wasp making nests wherein to lay her eggs might seem at first to be evidence of the wasp's consciously carrying out a plan of action. But it turns out to be an automatism, triggered by certain stimuli, even when

\(^4\) See above, §1.2, §4.4.
wildly inappropriate.\textsuperscript{5} Dennett makes much of sphexishness, as showing how wrong we are to ascribe consciousness on the basis of apparently purposive behaviour; but his argument is two-edged. In seeking to rebut the ascription of consciousness to the sphex wasp, he allows that the apparently purposive behaviour constitutes some \textit{prima facie} evidence in favour of the ascription, and in arguing that the ascription ought none the less be withheld in this case, he indicates another criterion for consciousness, which, if present, would strengthen the case for regarding an organism as, indeed, being conscious. The absence of what we might call holisitic assessment in cases of sphexishness not only counts against the sphex wasp’s being conscious, but suggests that where this criterion is satisfied, an ascription of consciousness may be justified.

“Holistic assessment” has proved a problem for workers in artificial intelligence, who know it as the “Frame Problem”. They are now able to program machines to behave in fairly complicated ways, but not to adjust themselves to the wide range of circumstances which they are likely to encounter. Organisms have evolved to be able—within reason—to adjust. Although they occupy some particular ecological niche, there is a much greater range of variation within that niche than any machine thus far constructed can accommodate itself to. It is a mark of consciousness not to be so severely limited in its responses. Although often we need to concentrate on the matter in hand, disregarding everything we take to be irrelevant to our present purpose, tunnel vision is a defect: we need to be alert to the emergence of unforeseen factors, which may disrupt our carefully contrived plans. Wide-ranging responsiveness to the environment is taken as evidence of awareness, itself a central facet of consciousness.

With human beings (and perhaps some other “higher animals”), a further facet of consciousness is the ability to stand back, and reconsider the whole situation, oneself included. It emerges from the ability to adjust to unexpected variation of circumstance; for this involves an ability to stand back and consider not only the circumstances but also the programs, what they can accomplish, and hence in what circumstances they are appropriate. If I am alert to my environment, I may become aware of factors adverse to my

Consciousness
Different Contrasts:
(1) Medical: as opposed to being asleep, anaesthetized, con-
cussed, in a coma.
(2) Responsive: as opposed to being inert, like a mineral.
(3) Sensitive: as opposed to being merely homeostatic, like a
vegetable.
(4) Alert: as opposed to having tunnel-vision, ?like a sphex
wasp?.
(5) Self-conscious: as opposed to not being able to stand back,
and reconsider.

present plans, and I may then need to weigh up the pros and cons,
and decide whether to continue with the plan I had adopted, or to
modify or abandon it. I am led to detach myself from myself and my
situation, and re-assess the situation from a new standpoint, which
can be seen as a higher standpoint in as much as it is more detached,
and less in thrall to immediate circumstance. Standing-backness
and self-criticism are peculiarly characteristic of consciousness. Its
first appearance among animals—the discovery on the part of tits
how to peck through the tops of milk bottles, of rabbits how to
gnaw through the plastic guards around growing trees, of household
pets how to find their way home over long distances—attract notice
and are commonly cited as evidence of animals being conscious like
us. The full transition from being conscious to being self-conscious
appears fairly late—only in human beings, and often not even then.
But the first origins of this transition seem to be deep in the concept
of consciousness. The conscious organism is something separate
from its environment, acting homeostatically so as to preserve itself,
and reproduce its kind, capable, to some extent, of adjusting its
strategies to fit the situation in which it finds itself, and hence with
some power of reflection.\textsuperscript{6}

Thus far we have distinguished three inter-related marks of con-
sciousness: 1. homeostatic behaviour that suggests the agent has
a goal which is pursued in spite of adventitious alteration of cir-
cumstance; 2. some plasticity of behaviour, revealing a sensitivity

\textsuperscript{6} See Bernd-Olaf Kippers, \textit{Information and the Origin of Life}, (tr. Manu
to a wide range of circumstance and an ability to size up the situation as a whole, and adjust behaviour to the whole of it, and not just a few salient features; and 3. some degree of independence from physical environment. What leads us to ascribe, on the basis of these criteria, actual consciousness is its explanatory power. As we saw in chapter six, if we can integrate a wide variety of different pieces of behaviour as manifestations of a single state of mind, as we can with Ryle’s depiction of the different manifestations of vanity, it is rational to posit that there is, in fact, a mind at work. We can understand diverse pieces of behaviour, provided we posit the existence of a person, conscious of himself and how he appears to others; therefore it is reasonable to make this posit, just as it is reasonable to believe that electrons exist, because by so doing we can explain many diverse phenomena. The same line of argument that justifies our rejecting phenomenalism and believing in material objects as the reality that gave rise to phenomenal appearances, justifies also the ontological assumption that certain patterns of behaviour are to be understood as the actions of conscious beings.

The arguments for ascribing consciousness are presumptive, not conclusive. They can be defeated, as Dennett seeks to defeat the ascription of consciousness to the sphex wasp. There is room for rational disagreement over whether, and to what extent, animals are conscious. Ants show considerable flexibility of response to their environment: do they feel pain? Is it cruel to pull of the wings of house flies? In contemporary Britain there are many who believe that foxes feel fear when pursued by foxhounds, but that fish do not feel pain when pulled by a fish hook in their mouth, and that therefore hunting should be banned, but angling allowed. Often it is the consequences of imputing consciousness that seem to tell most against making any such ascription, and it is useful to consider what follows from regarding an organism or machine as conscious. We are most keenly aware of the moral consequences: we think it wrong to cause animals pain, unless for a sufficiently good reason, whereas we have no qualms in cutting down a tree or boiling unanaesthetized carrots. If computers were conscious and sufficiently self-aware, we might feel it incumbent to consult them about their own future, perhaps even give them the vote. But these moral consequences are not fundamental: rather, they flow from

\[7\] See above, ch.6, §6.13.
a metaphysical view of conscious beings as centres, each having a 
view of its own upon the world. We should not cause people pain, 
because they are sentient beings, for whom pain is, from their point 
of view, bad. In attributing to them a point of view, we attribute 
to them a privileged authority on what that view is. I am the au-
thority on me. What I say about me goes; not absolutely—I may 
be lying, dissimulating or malingering, or I may misunderstand the 
English language, or even misreport my own feelings; but these are 
only subsidiary derogations from my being generally the person pe-
culiarity entitled to speak about me, and to have my words believed.
I may have a terrible disease or a horrible wound, and the doctors 
may expect me to be in great pain: but if I assure them that I feel 
no pain, and my actions confirm that I am being truthful, then 
the doctors have to accept my word for it, and perhaps look for 
other explanations—perhaps I have had the operation under hyp-
nosis, perhaps I have had acupuncture, perhaps I am one of those 
unfortunates who can feel no pain. Conversely, if I feel pain, I am 
to be believed, even though there is no discernible cause for it. It 
may be "only psychological", it may be some malfunctioning of the 
nervous system, but it cannot be denied without imputing to me 
dishonesty. Similarly in the case of animals, except that there we 
have no distinction between behaviour generally and the special 
sub-class of linguistic behaviour: if we read their behaviour as that 
of a sentient being, we believe it in the absence, or even against 
the evidence, of physiological explanation—if a rabbit squeals, I 
think it is hurt or frightened even if I can see no cause of pain or 
ground for fear, whereas if a tyre squeals, I give it no probative 
force and disregard it as soon as I can account for it as the result 
of friction on the road's surface. If we came to regard computers 
as conscious, we should give more credence to their output than 
to their hardware, and be prepared to agree that something was 
wrong even though the most exhaustive examination failed to re-
veal it, provided the output, either by means of direct symbolism 
or through suitably modulated aversive behaviour, indicated that 
this was so. Once this point is reached, moral consequences follow: 
the artefact has a right to consideration, because it has a view of its 
own which cannot be subsumed under views available to us apart 
from its explicit avowals or significant actions.

Consciousness has different facets and different aspects. Our 
ascriptions of it are always fallible, but can be justified. They 
explain. IF the reductionist programme could be carried through,
§14.3 Persons

their explanatoriness would be merely an illumination for us, and ontological parsimony would still counsel not regarding sentient beings, and in particular, persons, as anything very special. But since no reduction can be carried through, ontological parsimony counsels in favour of regarding persons as fundamentally real, since only by doing so, can we make sense of diverse phenomena. If we can attribute feelings and motives, and more generally empathize with other men, and to a more limited extent with other animals, we can achieve an understanding of their reactions and behaviour in the characteristically complicated, and often unique circumstances in which they are situated. We try to see things from their point of view, because there is no other way in which we can see things coherently. But once we accept they have a point of view, which is in the same ontological class as ours, we are impelled to take into consideration not only our own hopes and fear but theirs also. And then we are in deep waters.

§14.3 A Mind of One’s Own

Locke explicated personal identity in terms of memory, and common sense has since accepted a Lockean position which makes memory-claims, along with character traits and first-personal avowals, the definitive criteria of personal identity. Such an account seems to fit our actual practices, acknowledges the prime importance of the first-personal standpoint, allows for the intelligibility of Dr Jekyll and Mr Hyde, and the various Miss Beauchamps, and for the widespread human hopes and fears of a life after death, and all this without being very strongly committed to any particular metaphysical scheme. But it has been criticized, because it fails to guarantee uniqueness. The memory claims of any one person could be put forward by another person, sincerely believing that he was the same person that the first claimant genuinely believed himself to be. Just as physical appearance and behaviour can be duplicated, so memory claims, character traits, and avowals of identity can, in principle, be duplicated too. Whatever evidence we might have for identifying one claimant as the person he claimed to be, it is logically possible that there could be another candidate with exactly the same features, and therefore equally worthy to be identified as the same person. Having unreservedly accepted the first as the genuine claimant, we should be nonplussed on encountering the other; we might, like Buridan’s ass, conclude that since they cannot both be, they neither of them can be, the person they
both claim to be; or, perhaps more reasonably, we might seek some further surety for uniqueness.\textsuperscript{8}

One feature that cannot in our world be duplicated is bodily continuity, and many modern philosophers have concluded that bodily continuity is the constitutive condition of personal identity, as it was of corpuscles, and still is of atoms. But there are difficulties, both in practice, and in principle.

With few exceptions, bodily continuity is not a criterion we actually use. Only with patients intensively watched over in hospital and prisoners securely locked in their cells, can we be sure that they are bodily continuous with their previous selves. Nearly always we rely on faces and self-avowals. That point, by itself, is not decisive. Criteria are features which serve as handy indicators; they are subject to constraints of practicality: they need to be accessible, easy to apply, and reasonably reliable under commonly prevailing conditions, but do not need to exhaust the whole content of the concept, nor to be infallible, nor even reliable under conditions which seldom, if ever, obtain. In order to delineate the concept, we are concerned not just with criteria, but with constitutive conditions. But once we relax the certainties of ordinary life, and take seriously the possibility of two claimants with qualitatively identical characteristics and memories, we cannot rule out in principle the counter-factual possibility of someone’s splitting bodily, amoeba-like, into two.\textsuperscript{9} In the corpuscularian philosophy spatiotemporal continuity was a constitutive criterion of identity, because corpuscles possessed no essential qualities, beyond occupying some place or other at any one time, without there being any place that any one corpuscle necessarily occupied at any given time.\textsuperscript{10} But it is only a contingent, not a necessary truth that material objects and human beings do not divide or join up like rain drops. If they did, our concepts would be very different from


\textsuperscript{10} See above, §9.6.
§14.3 Persons

what they are.\textsuperscript{11} Even if we adopt the metaphysics of modern materialism, bodily continuity would not have for human bodies the pre-eminent position it has for atoms, because human bodies are processes rather than things, processes involving a continual flow of energy and matter, with no particular material particles reliably continuing as permanent constituents of any human body.

Although bodily continuity cannot make good the inadequacies of the traditional account, the criticisms made of the traditional account are cogent. Personal identity is necessarily unique, but every characterization of a person in terms of memory claims about the past or other present mental features, habits, or abilities, can be duplicated. We can secure uniqueness only if we look not just to the past, but to the future, not just to the passive, but to the active, and not just to the actual, but to the possible. Once again, we must conjugate. We need to conjugate over tense and mood, considering not only the actual present and unalterable past, but the possible future; we need to conjugate over voice, considering ourselves not only as the subjects of experience but as the initiators of action; and over person, considering not only third-personal ascriptions of identity, but first-personal avowals, and second-personal discourse.

Locke held ‘person’ to be a forensic term, whereby people were held responsible for what they had done, and blamed for what they had done amiss. But though at this present time I can only remember having done some particular deed in the past, at one time I was deliberating, and was making up my mind whether I should do it or not. At that time it was not a matter of my merely having certain experiences, putative memories of having done it in the past, but of my forming an intention for the future and actually carrying it out. I was an agent, not passively experiencing but actively doing. And until I had carried out my intention, it was an open question whether or not I actually would. Although now that I have done it, the deed is done beyond recall, it was at one time only a possibility, with its being still up to me whether or not it would actually be realised.

The importance of future possibility is shown by an example from Alice Through the Looking Glass. Tweedledum and Tweedledee were more than ordinary identical twins: they shared consciousness and experiences. Each knew what the other was doing, felt the other’s pains, and remembered what the other had

\textsuperscript{11} See K.M.Wilkes, Real People, Oxford: Oxford University Press, 1988, ch.1 for reasoned scepticism about the extent to which counter-factual hypothetical experiments should weigh with us.
done. We might be tempted to say that there was only one person, Tweedle, but with two bodies. Bilocation seems perfectly intelligible, and would have advantages for busy modern man, with 2.8 engagements for every available hour. While one of me was dutifully present at a meeting, my happier half could be reading a good book, buying a surprise present for my wife, or having lunch with a useful contact, only occasionally interrupted by the need to know which proposal my dutiful half should vote against. With four arms and four eyes I could make much better use of my time than I currently do. So long as there were unity of control, and my dutiful half could not vote against the better judgement of my liberated half, it would be right to regard me as two bodies with but a single mind, and therefore just one person. It would be quite different if there were, or at least potentially could be, some conflict. Then there would be two of us, perhaps in very close telepathic communication, and often of one mind on matters of importance, but not necessarily agreeing on all things. If, as Alice was assured, Tweedledum can quarrel with Tweedledee, then they are indeed two separate centres of decision-making, and not two arms of the one unified Tweedle. Tweedledum has a mind of his own, different from Tweedledee’s, since he can make it up differently. Even if they always in fact agree, they could differ, and so are different. Having a mind of one’s own is being able to make it up differently.

Tweedledum and Tweedledee Agreed to have a battle; For Tweedledum said Tweedledee Had spoiled his nice new rattle.

Just then flew down a monstrous crow, As black as a tar-barrel; Which frightened both our heroes so, They quite forgot their quarrel.

The future encompasses many possibilities, whereas the present consists of only one actuality, and the past of only one unalterable course of events. In talking about the future, therefore, we have to consider all possibilities: if Tweedledum and Tweedledee can decide to attempt different things, then, whatever the harmony hitherto, there is some future possibility in which one is trying to counter what the other is trying to do; if Tweedledum and Tweedledee cannot decide to attempt different things, then among all the possible
future courses of events there is none that represents the one trying to thwart what the other is endeavouring to achieve. And since we are already in the realm of possibility, there is no room for a further appeal to possible duplication to confuse identity. If we are faced with two people both plausibly claiming to be the same person, let them meet each other. If each denounces the other as an impostor, then we can be at least sure that they are not both the same person in the way that Tweedle, were it not for the possibility of battle, might have been: one of them at least is an impostor. We may not be able to tell which, and we may have another Tichborne case on our hands; but that will be good news for the lawyers rather than bad news for the philosophers. If on the other hand, each treats the other with fraternal affection, taking him for granted as a fully accredited and completely apprised fellow-worker in a common field of endeavour, we have a bilocated person enjoying two bodies but with only one will. When we survey the future, either there is a possible course of events in which the two putative Tweedles fight, or there is not: if the former, they are different persons, Tweedledum and Tweedledee, not both Tweedle; if the latter, there is only one person we are dealing with, Tweedle. There is no third alternative: having considered all possibilities already, there are no further ones to generate further putative persons. Moreover, the all-or-none character of all possibilities secures the all-or-none character also of individual identity. When dealing with individuals, there is no room for a nicely calculated less or more: each someone is definitely a one, as well as necessarily not anyone else. By considering all the possible actions of an agent, we secure that each is one and all alone, and ever more shall be so.

Locke can be fairly charged with having neglected the future in his account of personal identity, but Williams cannot, and in “The Self and the Future” argues from the self’s future concern for itself: but concern presupposes the possibility of action and thus of intention.\textsuperscript{12} I am afraid of being tortured: and therefore I take action to avoid it. If I did not, other things being equal, try to avoid being tortured, I could not be said to be afraid of being tortured. So we can ask at different stages in Williams’ account who would try to take avoiding action. Suppose the torturer were to drop his

guard, and leave the prison door open: it would be the person who was going to be the inhabitant of A's body at the time of torture who would self-interestedly want to escape. The decision-maker is the person who can try to avoid the threatened evil, and is concerned to do so. Often puzzle cases in the philosophy of mind are constructed in which we are invited to imagine ourselves in some dreadful situation paralysed or completely under the control of a mad scientist. But such situations are, not only fortunately but necessarily, untypical. We are, first and foremost, agents: were we not agents, it is doubtful whether we could be patients either—consciousness is very largely a feedback on endeavour—and we certainly could not have the concept of a person that we do have. *Ego, ergo ago*. As I contemplate the dreadful things that might happen to someone, I identify with him whose actions then, should circumstances turn out to be less unpropitious than currently portrayed, I can decide upon now.

Since each rational agent can make up his mind what he should do, and to a large extent carry his decisions into practice, it matters very much to us with whom we are dealing. We need to be able to distinguish different agents because they have different patterns of behaviour, and these differences may matter to us very much. Individuals signify, not simply as idiosyncratic obstacles we need to learn to manipulate if we are to achieve our ends, but as sources of value, differing from us in some important respects, but none the less valid and worthy of our own affirmation and respect.

We can look at this in a Leibnizian way. There is a tension between the necessary uniqueness of the self and the possibility of multiple instantiation of any set of qualities. So long as we are dealing with only a finite set of qualities, they specify some *infima species*, which only contingently and not necessarily has only a sole specimen. If, however, we move into the infinite realm of possible courses of events, we can specify an individual uniquely by an infinite specification characterizing every response he might make under every set of conditions. If two putative persons would respond to every situation in the same way, then they are one and the same person, numerically identical: if there could be some discordant response, they are to that extent qualitatively as well as numerically distinct. Instead of needing the bodily continuity of the corpuscularians to guarantee the separate individuality of each entity necessarily located in space, each monad differs from every other one by virtue of the way it actualises the infinite potentiality
open to it as an agent. The Leibnizian picture of individuals as infinitely complicated monads, each qualitatively different from all others, is important as a prophylactic against a scepticism that has prevented many philosophers from conjugating over persons. They have been confined to the third person by qualms about the first-personal approach. But while we should sympathize with those who are in the grip of a theory of meaning, we need not share their affliction. We have experience, most notably of other people but also of concepts, which, because they go beyond anything that could be programmed or specified from outside, seem to us to be indubitably real. People have style. Not only men of letters in their literary output, but all of us in our ordinary activities, do things in our own characteristic and inimitable way. Although a well-trained classicist can produce a passable parody of ancient authors, extended imitation is soon distinguished from the real thing, because the real is original, with unanticipated features which ex post facto are recognised as being absolutely right, but which did not conform to any antecedently specified rule. At a much lower level it is often the mark of someone’s having understood a concept that he is able to apply it in new cases that lie altogether outside the existing rules, and, at a slightly higher level, of his having got inside a poem or play that he can make critical comments that go beyond the most that the most assiduous student could have mugged up. In the non-denumerably infinite welter of possible performances of possible monads we can distinguish the rational from the irrational even though we do not have any explicitly formulated rule to tell us how to do it. In mathematical terms people are capable of non-algorithmic reasoning which is right without being rule-bound, and individual without being random. Hence it is that we are, most of us, robustly immune to doubts whether we can meaningfully ascribe to another person a first-personal standpoint and private experiences all of his own: we have on occasion felt the force of a individual personality, different from anything we could have thought up for ourselves, yet rational none the less; and that gives us confidence to believe that we are dealing with something real, an entity that is the same on different occasions of encountering it, an entity such as we hold ourselves to be. If I recognise something as a person, I recognise that he has his own first-personal standpoint, and can use the word ‘I’ to express it in the same way as I do to express mine: and if I re-identify him as a particular person, I ascribe to him the same coherent identity
as I assume for myself in my view of my own affairs. Were it not for this, I should not be interested in the identity of individuals, but should be content to identify species, as I am with flowers. Although in each case I may turn out to be wrong in my identification of a person, the point of the claim is to establish him as an *alter ego*, a particular *alter ego* with whom I have had dealings in time gone by, or may have dealings in time to come.

If I regard someone as an *alter ego*, as someone who can use the word ‘I’ to express it in the same way as I do to express mine: I shall want to address him in the second person: I ask *you* to give your own account, in the first person, of what you did and why you did it, the view you now take of your actions, your opinions about current affairs, and your plans for the future. Personal identification is important because it opens up the possibility of dialogue in which the person concerned can give a first-hand account, and when he uses the first-person singular, we shall know to which person the word ‘I’ in his mouth refers. The fact that we often have to ask people their names has not been sufficiently attended to. It proves the prime importance of the first person even where mere criteria are concerned, and is becoming increasingly important as other ways cease to be available in consequence of the meagre information provided by electronic means of communication.

Potential disagreement is constitutive of personal identity. If two people can disagree, they are different persons: if they cannot, they are the same person. But it does not provide a convenient criterion. Faces and avowals are what we commonly use, and are nearly always reliable, sometimes supplemented by other bodily features, character traits, and memory claims. These work well enough in practice, though if the world were very different, we should be unable to rely on them as much as we do. Underlying all personal discourse is our seeing persons as agents, each an ἀρχή τοῦ κινήν (*arche tou kinein*), an initiator of action, with a mind of his own, which he makes up for himself, and can make up differently if he so decides.
§14.4 Fusion and Fission

The prime importance of the first-personal approach should guide us when we come to consider Derek Parfit’s difficult cases. It is the point of ascribing an identity, not the criteria which have served us well in more mundane circumstances, that should carry the most weight with us. In so far as we can make sense of fusion and fission thought experiments, we need to ask how the story is to be told from the inside before attempting to say how we, on the outside, would respond.

Let us deal with fusion first. There is an obvious difficulty, analogous to that in spare-part surgery. Transplants are commonly rejected, just because, as we have seen, human bodies are not material objects but processes, more like flames or eddies than pieces of furniture or wooden ships, and what counts is not the material particles composing a body at any one time, but the metabolic flow of matter and energy, and the information implicit in the organization of the organism. In the same way, any hypothetical experiment involving the fusion of two personalities would have comparable difficulties in integrating disparate memories into one coherent and plausible history of the putative person. Of course, a story can be told of a mad scientist getting hold of me, and implanting some of the memories of Fiona, and of my waking up with apparent memories of my first pony, Queen Charlotte’s Ball, a London season, triumph at Badminton, and my marriage at St George’s, Hanover Square, to a sprig of aristocracy. But apparent memories are not, contrary to the opinion of many philosophers, accepted always without question. In the quiet watches of the night I vividly remember many exploits, sometimes as Napoleon, sometimes as Pericles, sometimes, less properly, as the Sublime Porte, surrounded by eunuchs and concubines, all eager to gratify my merest whim; but these impressions, however vivid, are rejected because they do not fit in with the rest of my remembered experience; and, though less spectacularly, some waking memory-impressions are likewise subject to critical scrutiny, and in spite of the fact that I could have sworn that I posted the cheque, I ruefully concede that I must have misremembered it as I find the cheque in its envelope still nestling among the other papers on my desk. Memory, though generally reliable, is not of its nature infallible, and we only accept those memories which do not run counter to the rest of what we know and understand. One look in the mirror will convince me that I was never the deb of the year, half a
minute on a horse that I could never have gained even a rosette at a pony-club rally. Just as the body rejects intruded cells, so each man’s autobiography rejects intrusive matter which cannot be accommodated within what is already known and understood; and although the immune system of the body can, at great cost and risk, be sufficiently suppressed to allow transplants to be made, the suppression of the critical faculties required before I could be content to accept Fiona’s memories as my own would be so great as to leave me a completely irrational, and barely self-conscious, being. No first-personal account of fusion is coherent: although I may by my own free decision share the rest of my life with another, merging my individual will in our joint choices, and coming to know all her memories as she recounts them, neither of us can know the other’s early memories first-hand, and our sharing could never conceivably become a complete coalescence.

Much the same is true of fission, though not so absolutely. Each one of us needs to be able to make sense of his future as well as of his past. A person, as Lockwood maintains, arguing from a very different starting point, is essentially a diachronic concept: we should take it as definitive of mind that it encompasses its own past actions and reaches out towards what is to happen. As Glover emphasizes, our identities are very largely what we ourselves create. If I could not form a coherent view of my own future, I should not have a coherent concept of myself. I find no difficulty in envisaging a future in which I am bilocated, but formidable difficulties in making sense of the prospect of my being two mes, not of one mind with each other. I should not want my Abel-self to be slain by my Cain-self, undeterred by Swinburne and anxious to secure his sole claim to be me. So I decide now not in either of my successor selves to slay my other self. But will this decision stick? After all, it is always possible to change one’s mind. I might, I suppose, make sure by deciding now to separate, like Abraham and Lot, but our experience with actual identical twins makes this seem unnatural, and in the face of such a prospect I am much

---

more likely to resolve very firmly not to change my mind. If that possibility is excluded, or if intentions do not survive the process of fission, then I am already no longer a person, and nobody can be identified with me. In order for there to be any question of my continuing to exist, I must be able to form intentions and stick to them. In that case, I shall resolve not only that we shall, neither of us, hurt or thwart the other, but that we shall stick together through thick and thin. Tweedle will rule out the possibility of a battle, and will remain, so much as in them lies, still just Tweedle, though bilocated Dumly and Deely. So, too, the tribe of Lucas clones will stick together, like a pack of wolves, each abiding by the decision made in the original position, each identifying himself as part of the pack, quite possibly using only the first person plural; and should any have the misfortune to be separated from the pack, likely to fade away and die through inanition and loss of identity.

Difficulties remain. The two most notable are the division of consciousness and the unforeseen fission. It could be the case that Dum and Dee did not know, until told, what the other was doing. But this often happens in ordinary life, where I do not know what I am doing until I tell myself and consciously recognise it. More pertinently, in the real-life split brain cases there is some division of consciousness between, as it were, two centres of consciousness. But we should note our response: a few philosophers excepted, we continue to regard the patient as a single person, because for the most part he is able to make over-all decisions and integrate his behaviour into a coherent pattern. This no accident. Consciousness is not entirely separable from decision-making, but is, as we have seen, largely a feed-back on action.16 If we are of one mind what to do, we shall be mutually aware of how we fare in consequence.

Fission is not altogether symmetrical with fusion. With fusion, whatever memories may have been implanted in me, I shall have to decide whether to own them as memories of my own actions, or disown them as fantasies that can have played no part in my real life. In the case of fission that I have considered, I was able to decide antecedently what I should subsequently do, and impose on

---

my future an orderly theme. But what if an individual is divided unawares? If I did not know I was fissile, and it just happened to me? I meet someone who looks like me, behaves like me, has all my memories, aspirations and ambitions, and who says he is me? Do I regard him as an impostor to be killed, or an alter ego, to be cherished and cooperated with? Answers give out: but perhaps the questions have run out too. I share Dr Wilkes’ scepticism both as regards the feasibility of surgical Parfitry and its relevance to our understanding of the concept of a person. The very fact that our actual concepts have been developed against a particular background of fact means that our criteria are likely to pick on contingent rather than necessary features as reliably characteristic of what it is to be a person, or to be the same person as one previously identified. Habeas corpus makes sense in humdrum England, but would have been of no avail to Ariel seeking his liberty from Prospero’s spell-enforced servitude. If the transmigration of souls were an established fact, if we lived in one of Shorter’s imaginary societies, if most of our pupils behaved, and avowed having experiences, like Miss Sally Beauchamp, if our colleague Dr Jekyll alternated in behaviour with the altogether un-electable Mr Hyde, our concepts and forensic practices would be very different. So too with other, non-personal concepts. If the world were very different, our concepts would be different too. We should not have the concept of a material object if we lived in an aqueous medium, in which beneficial and noxious chemicals were diffused, but with no definite sources or boundaries: we might then have a purely olfactory experience, in which we could distinguish various sorts of “good—here—now” and “bad—here—now”, but no numerically distinct objects we could re-identify and manipulate. Hence it is reasonable to meet difficult fission possibilities by pointing out that we should not have the concept of a person, if agents did not continue over extended periods of time, pursuing long-term plans, remembering what they had done, and what they had decided to do. I should not be conscious of myself as a self if I could not decide what I was going to do, and always forgot what I had done and what I had been going to do. Not that I have to have complete plans for all my future life or complete memory of all my past doings: I can tolerate a large measure of present indecision and forgetfulness. But once I am not totally undecided and amnesiac, I can take steps to transcend my limited abilities. I can write memos to myself, to remind myself of what decisions I either have
made or need to make, and I can keep a diary. Thus, granted only the minimal capacities required of our being agents at all, we can extend our effective forward planning and backward memories to connect our present selves with our future concerns and previous deeds. It is only if we are hopelessly restricted to the present that we cannot get a grip on the future and past, and then we should no longer be agents at all, and should not know ourselves as such.

So, if we cannot rule out puzzle cases absolutely, we can at least marginalise them, either by reckoning them so bizarre that they altogether undercut the notion of one’s having a first-personal view of oneself as a person, or, if any reasonable power of decision-taking is left, by extending it to cover the difficult cases, and taking them into account as we envisage possibilities and form life projects, and by discounting awkward possibilities by definite decisions. If we are to have a concept of an agent, he must be able, at least a little, to extend himself towards the future and to own his deeds done in time past, and any such extension is itself extensible. I can make up my own mind what I shall do, and it is the mark of its being my mind that I can make it up for myself, differently from you.

§14.5 In Praise of Bodies

If in our minds we can range freely over space and time, we might suppose that we could do without our bodies altogether. Plato thought so. Many religions teach that our mortal flesh is an encumbrance to purity of spirit, and popular ideas of life after death envisage some sort of disembodied existence.

But there are difficulties. Minds have been shown to be largely independent of bodily circumstance, but not absolutely so. Over a wide range it does not matter which particular atoms, molecules, physiological processes are involved, but it does not follow that it does not matter whether any atoms, molecules, physiological processes are involved. As far as explanations are concerned, minds are often the ultimate entities: but as far as ontology is concerned, they are not ultimate simples, but rather patterned composites. And although it is the pattern we should be primarily interested in, it is relevant question whether the pattern is actually instantiated or not.

But how relevant? A strong Platonist holds that universals exist ante res, and could in principle work out theoretic chemistry from quantum mechanics, not much caring whether any particular element or compound was actually to be found somewhere in
the universe or not. Disembodied minds, it might seem, are in the same case. Might we not work out what their personal qualities must be, and assess their moral worth? I can imagine myself without a body—perhaps looking down at my own funeral, and seeing my body committed to the flames or the earth, but how could people recognise me thereafter, if I did not have a face? Indeed, could I have visual experience if I was not located anywhere? The argument that led to the perspectival view posited windowed monads who could not only communicate with one another—the radio rule—but interact causally with one another to the extent that each could observe where others were located—the radar rule—for which they needed sense organs to observe, and bodies to be observed.

Some philosophers find these difficulties insuperable; Sir Peter Strawson wittily observes, alluding to the Apostles’ Creed, that it is just as well that the orthodox believe in the resurrection of the body.\(^7\) Some of the difficulties, however, though real, are not insuperable. After all, in the E-mail community we may never have met or heard those we have converse with, but know them only by the different addresses at the head of messages that flash up on our VDUs. Cyberspace affords a model of a universe of windowed monads, who do not essentially have visual experience, and are known by their E-mail addresses rather than their faces. Although we, who are sighted, use screens, blind people can use E-mail, with auditory and tactile substitutes for visual display units. Some of the adolescent young live primarily in a world of text-messaging, only occasionally emerging to eat, and make sure their parents are still around. Although we need some sensory input and some motor output to do E-mail or surf the web, we are not dependent on any particular sense or any particular mode of operation. We can imagine a world of nerds, each being able to register a unique E-mail address at some designated site.

McTaggart would have found such a world intelligible, and answering, though perhaps a trifle austerely, to his ideal. In spite of difficulties over initiation into such a community of souls, learning their language and, more particularly, establishing intersubjective referents for common nouns (we could posit that they all have a natural—“innate”, one might almost say—ability to surf the web), they could have meaningful dialogues about mathematics, not only

elementary number theory, but analysis, the theory of groups, and perhaps even analytic geometry. Such a picture presupposes that each communicating entity is an individual, and reliably the same individual, who has exclusive and secure control over the use of his own identification number in authenticating messages, so that there is no danger of messages to or from john.lucas@merton.ox.ac.uk going to, or being attributed to, the wrong person. Such a cyber-society, in which we had no eyes or faces, seems conceivable. Admittedly, if its members never had had any embodied existence, they might have rather little to talk about, and could become boring, unless they developed advanced mathematical interests. (Minds that had been embodied in some previous existence would fare better; they could remember and recount scenes and sounds from their yesteryears, and reminisce about the hardships they had endured, and the successes they had achieved.) But possible boredom apart, would disembodied minds be at disadvantage compared with their embodied colleagues?

They might be. It depends on how their existence is structured with regard to what they can do and what may be done to them. Minds differ in being able to decide differently. In the world we live in, we are not always able to carry out our decisions. There is a gap between aspiration and achievement. Our aspirations are often thwarted by the untowardness of circumstance or the interventions of others. In a disembodied existence, we should not be able to do things by means of our bodies, but we can conceive of our being able to do things by some form of telekinesis. I will that the mountain be removed to another place, and lo! it is removed. But what if you would not have it so, and will that it stay put? We cannot both be telekinetically omnipotent.

If I am to operate in a non-solipsistic universe, there must be a gap between aspiration and achievement. I am free to aspire only on condition that wanting is not the same as getting. In the public world I am necessarily not the only pebble on the beach, but if I am to be not merely a pebble, but to have a mind of my own, there must be some realm where my mind can roam unfettered by the vetos of others or the recalcitrance of reality. I need some privacy of intention, if I am to have a mind of my own: and I must operate in a public realm, if I am to be an effective agent. The distinction between public and private is a deep conceptual one, deeper than the traditional distinction between the bodies
and minds we are acquainted with, linked only contingently by discoveries in neurophysiology and our everyday experience.

In cyber-space I can send E-mails to all and sundry, but am often the victim of spam—unwanted E-mails inviting me to download porn, or make $10,000. I can protect myself by setting up filters that siphon off E-mails containing certain words or the $ sign. Adolescent nerds might gang up on an unpopular one, and agree that they each would block his E-mails to them. It could happen on a world-wide scale. The fashion-setters might decide that my views were insufficiently PC to deserve a hearing, and declare that it was uncool for anyone not to block E-mails emanating from john.lucas@merton.ox.ac.uk. I could go on composing E-mails to my heart’s content, expressing ever more politically incorrect views, but it would be the freedom of impotence. Only if I am the only effective person in the world, can I be both omnipotent and invulnerable. If there are others like me, then I shall be able to do some things, but unable to prevent them doing others, and unable sometimes to prevent them preventing me from achieving what I want.

Although I am quite fond of my body—a poor thing, but my own—I can imagine myself dispensing with it; but I cannot imagine there being absolutely no distinction between the public and the private, between an omnipersonally accessible world and my own first-personal perspective, with the concomitant possibility of my aspirations being frustrated, and my being vulnerable to unwelcome concatenations of public events.

<table>
<thead>
<tr>
<th>I and Me</th>
</tr>
</thead>
<tbody>
<tr>
<td>I can in my thoughts imagine that I am free of all material constraints, and can occupy any location in space and time</td>
</tr>
<tr>
<td>But</td>
</tr>
<tr>
<td>in point of fact it is not possible for Me to be without a material body, or to be located at any other place or time.</td>
</tr>
</tbody>
</table>
§14.6 The Subjective View

In the austere world of windowed monads there is a tight correlation between the perspective of any frame of reference—the coordinates assigned to a particular point—and its position in the external world. With human observers there is a correlation, but not a tight one: the letter on the optician’s screen may be an A, but the patient truthfully reports his seeing an R. If our prime concern is the nature of the external world, it does not matter much. One observer’s errors can be corrected by the observations made by others. But once we seek to include the subjective as part of reality, we have to consider the status of genuine but mistaken reports, together with other purely subjective experiences.

Many thinkers shy away from acknowledging purely subjective experiences on account of Wittgenstein’s argument against private language. But that argument does not prove as much as is commonly supposed. Wittgenstein argues that my use of language must be corrigible to count as a use of language at all, and that if there are no public criteria, I cannot be corrected by anyone else. And from this he concludes that I cannot be corrected at all. But I can be corrected by myself. I often make mistakes, and correct them as soon as I make them; and, less often, come to realise later that I was mistaken, and then silently withdraw my previous usage. Wittgenstein sneers at the man who, dubious about the truth of what he reads in a newspaper, buys another copy of the same newspaper; but if it were a later edition, it would not be futile—occasionally errors in earlier editions have been corrected in later ones. Often, of course, they are not; and similarly, often I don’t realise I was wrong, or won’t admit it. But equally often in my use of public language my mistakes go uncorrected. With my private thoughts there is a crucial difference between those that are in principle incorrigible and those that may on occasion be uncorrectable in practice. In practice my later self may fail to notice and correct mistakes made by my earlier self, but this does not make my usage incorrigible in principle. I can be wrong in my use of my own private language, and that is enough to make it a rule-governed activity, and hence a language, and not mere babbling.

Wittgenstein’s argument works only against a private language in which what I say is absolutely ungainsayable. When what I say is gainsayable, if only by me, my saying it (or even only thinking it) is not an inherently idle exercise. Its relevance, however, is tenuous, if it is open to correction only by my subsequent correction. Often,
however, it is open to correction in other ways too. Often I use a public language to tell you about my subjective experiences: you then can understand me, and can also contradict me: the problem is not one of intelligibility, but of warrant. How can you have grounds for correcting what I say about my own private experience? But the case is not hopeless. A sceptic insinuates that he and I do not see blue and yellow the same way; my experience of blue is like his of yellow, and vice versa. It is clear that we cannot settle the matter by reference to the external world. We both call the sky blue and primroses yellow, but that does not touch the question of how we see them. But I can show him a spectrum, and ask him to point out the lightest colour. If he deems the lightest colour to be the one at the further end of the spectrum away from red, then I shall believe that he is seeing as yellow what I see as blue; but if he deems the lightest colour to be the one in the middle, next to red, then I shall have some warrant for believing that he sees as yellow what I see as blue. Again, the sceptic insinuates that he and I do not see red and green the same way. I show him a spectrum, and draw his attention to the extreme blue end, asking if it shows tinges of red or of green. If beyond indigo he sees violet, his experience of red is the same as mine: if the colour beyond indigo is reported as having a greenish hue, I think he may be onto something.

Musical notes have a much more definite intrinsic structure, and we could go a long way in establishing that a man who was not tone-deaf had auditory experiences much like ours. Most of our experience is not purely sensory, and fits in with other experiences to form a more or less coherent whole, which not only registers the world around us, but reacts to it in various ways, some of which may issue in action. I sympathize less with someone who tells me he is in great pain, when I see him engaged in lively banter with members of the opposite sex. It may be that he is putting on a brave show, but I begin to look out for other signs: does he avoid certain positions, situations, topics? does he show the same gay abandon when not in company? can he concentrate with equal vigour on dull and boring work without being distracted by his excruciating pain? Pain is not just a sensation: it is something which obtrudes itself on our attention, and which we are averse from. Pleasure is what we have when we enjoy what we are doing, and is usually a concomitant of our doing it well, and wanting to do it again. Other feelings are more specific, often responses to events, and often connected with aspirations or intentions, which
could be realised in subsequent action. Although we may misunderstand them or misreport them—psychotherapists spend their time freeing patients from such errors—we usually have sufficient self-knowledge to chart our feelings and describe them reasonably well to others. I share my perspective with you. It would not be possible if we did not both inhabit the same public external world, which provides a common anchorage for many of the things you and I both experience, but it does not require that most of what I tell you obtains its meaning from its being about something in our common external world. You know what I am talking about because I tell you, using a common language which is much used for communicating perspectives as seen from the point of view of the speaker and intended to be understood by the hearer.

I can talk about much more than my sense-experience and basic feelings. I can express my point of view about all manner of subjects, including value-judgements and wide-ranging attitudes to public events and private aspirations. Since all our sense-organs select what they sense, and all our narratives select what they report, it is a mistake to condemn the subjective view as failing to make the grade of objective respectability. Rather, we should recognise that we have different objectives. Sometimes we are down-playing the subjective, and trying to report on a reality as absolutely conceived. But it is equally legitimate to be concerned to convey some thing more personal, seeking to enable others to see things the way I see it. Poets and novelists sometimes succeed in enabling us to alter our own perspectives. So, too, painters and sculptors. A picture is not a photograph. A photograph does not pick out what is significant from what is not, whereas a picture, in showing what the artist thought was significant, may teach us to recognise the significance of what had hitherto eluded us. We conjugate. And in expressing a first-personal subjective view, may communicate to others a second- or third-personal understanding, which in turn may illuminate, enlarge, or deepen, their own first-personal understanding.
§14.7 Secondary Qualities

If first-person subjective views can convey truth, secondary qualities need no longer be discountenanced. The colours, sounds, smells and tastes that appear to me are not for that reason unreal. Although appearances can be contrasted with reality, they can also convey perspectival information. In the table in §8.3, it is not the first sense of appearance that is in issue, but the fourth. Secondary qualities are not just what I, in the first-person singular, say they are, but what we, in the first-person plural, agree about. They may be subjective, but they are inter-subjectively monitored. There is nothing of arbitrariness or subjective whim about them. They are candidates in good standing for telling the truth about the world as it really is.

Nevertheless, they are discountenanced. The standard opinion, first formulated by Democritus, and of great importance to Descartes, and popularised in the English-speaking world by Locke, excludes them from any fundamental account of the world, in favour of primary qualities, which alone are accounted real. In the real world things really do have shapes and sizes, and can be counted and measured, but though we speak of them possessing colours, and emitting sounds and scents, we are, it is alleged, only painting our sensations on to them, attributing to them qualities they do not really possess, on the strength of how they appear to us.

Primary and Secondary Qualities

Primary: bulk, figure, number, situation, motion or rest, ?solidity/impenetrability?, mass, energy, momentum, parity, charge, ?spin?.

Secondary: colours, sounds, smells, tastes.

---

18 p.230.
The metaphysics of projectivism, discussed in Chapter Five, supports this view. Projectivism has the great virtue of ontological economy, as well as fitting in with the scientific world-view. Since, however, we cannot economize on objective probabilities, if we are to make sense of quantum mechanics, the argument from economy is broken-backed. But still, the argument from the scientific world-view carries weight. Whereas primary qualities were ones that entered into scientific explanations, secondary qualities did not, and did not play any part in explaining our perception of them. It was a telling criticism in the Seventeenth Century which had had hopes of explaining all secondary qualities in terms of primary qualities; but now we are seeking deeper explanations still, by means of theories that do not take primary qualities to be fundamental qualities, but rather answers to questions elicited by a certain sort of operator. Secondary qualities may, indeed, be explicable: but so, it is hoped, are primary qualities, which in any case no longer have the primacy in our scientific thinking that they once had. In any case, to explain is not to explain away. The Special Theory explains in terms of its position and relative velocity why each light-cone has the shape it has, but that does not make it any the less real. Even if secondary qualities can be explained, it does not follow that they are unreal.

We may be concerned merely to describe the world, but even then, it is claimed, secondary qualities are unsatisfactory because they are inherently first-personal, depending on the observer’s autobiography, rather than omni-personal, independent of the particular idiosyncrasies of the observer. Primary qualities, by contrast, are said to be observer-independent. If I see three cows in the field, each with four legs, there really are three cows in the field, each with four legs. I can go to a shop, and order twelve yards of curtain material, confident that when it arrives, it will appear to be twelve yards long, whereas the paint I chose at the same time may seem quite different when I put it on the woodwork. Similarly with sound, the whistle sounds different to me as the train speeds by. When I come in from a snowball fight, even the water from the cold tap feels warm. If I have a fever, sweet delicacies seem

20 See above, §§11.4 and 11.5.
21 See above, §11.10.
nauseous. Locke asked if porphyry was black at night.\textsuperscript{22} and we wonder whether chlorophyll was green in the Jurassic age.

These attacks on the descriptive integrity of secondary qualities can be met. Shape and size are in equally bad case. They, too, vary with the position of the observer.\textsuperscript{23} And although we may be able to assign numbers which do not vary from one observer to another, measuring is a highly theory-laden procedure, not the simple reading off of some intrinsic quality.

In observing shape and size we can, as Descartes’ examples of the distant tower and statues show, draw the distinction between how something appears to us, and how it really is.\textsuperscript{24} But so we can with colours and sounds. I realise now that I was looking at the paint chart under artificial illumination in the shop, whereas it is being seen by daylight in the house. Similarly with sound, I attribute to the Doppler effect the change in how the whistle sounds to me as the train speeds by, and believe that the note was really the same all the time. In the same way we differentiate between describing sensations of warmth and coolness and ascribing a temperature, which is intended to be a real quality of an object. I can say that the cold water feels warm to my chilly hand, though really it is 40°F. I can also say that the wine was really quite sweet, though it tasted sharp after the baked Alaska.

Doubts about the reliability of our sense organs are most persistent in the case of colour, for good reason. It has been hitherto the case that we are not so much aware of the variability of apparent shape as we are of the variability of apparent colour; the reason is that when we are conversing, we are necessarily conversing at the same time, and therefore occupying different positions; so we learn to talk not about the apparent shape and size of things, which would be different for each of us, but about a “real” size, which is the same for us all, wherever we happen to be. With colour, however, the pressure to talk about real colours rather than apparent colours has, for most of man’s history, been much less insistent: when we converse we converse at the same time, and since illumination has varied with time, but not significantly with space, and

\textsuperscript{22} Essay Concerning Human Understanding, II:8:19.

\textsuperscript{23} See phenomenal regression to real object in §8.3.

has varied only from daylight to dark, there has been no need to
distinguish a “real” colour of an object from its apparent colour
as seen under daylight. With the advent of artificial illumination,
the situation is changing. We are learning to differentiate between
how an object appears under various different sorts of artificial light
from its “true” colour, which is an invariant relation between inci-
dent and reflected light.25 In so far as we can make this distinction,
the secondary quality of colour is in the same case as the primary
quality of shape. If we are describing the world as it was long ago,
then it really was green, even if there was nobody around to see it.
The contrast is between the autobiographical statement of experi-
ence, as I, you, he, she, or anyone, might have stated it, and the
publicly accredited account of what there is in reality, according to
the rules of the English language. Green then is what things really
are, and if I report something’s being green, when I am in good
health, good sight, and am observing under normal conditions, it
does not just appear green, but really is green. If I want to alter
the colour of a material object, I dye it; I do not put on coloured
spectacles, or alter the illumination under which I view it. Green is
a real objective property of material objects, not as it may appear
to me, you, him, her, or anyone, but as we all characterize it.

From the descriptive point of view secondary qualities are in
the same case as primary qualities. We factor out each individual
standpoint, and adopt a person-invariant mode of discourse, just
as we do for primary qualities. In each case we attribute them
to the external world, being able to make the distinction between
what they really are, and how they appear to us, and being ready
to allow correction to our reports of what is really the case, either
on the strength of the testimony of others, or on coming to realise
that our observations might have been skewed by some adventitious
circumstance.

Jonathan Bennett argues that secondary qualities are periph-
eral, whereas primary ones are central:26 if I am colour-blind or
tone-deaf, I can manage quite well, whereas if I am shape-blind, or

25 Similarly with coloured filters: in my motorcycling youth, I used to wear
greenish-yellow goggles; when I first put them on, everything seemed green-
ish yellow, but soon I began to see things as they really were. When I took
the goggles off, the effect was reversed, but again, it soon wore off.

26 Jonathan Bennett, Locke, Berkeley, Hume: Central Themes, Oxford, 1971,
ch.IV, pp.89-123.
unable to estimate distances or durations, I shall for ever be trying to put square pegs into round holes, bumping into obstacles, and failing to arrive in time. This is true. Colours, in particular, are peculiar, in being remarkably separate from most other features. (It is a pity that philosophers have tended to take statements about colour as paradigm. The fact that the word ‘good’ is unlike the word ‘yellow’ shows not that the former is non-natural, but that the latter is non-typical.) But there are degrees of non-centrality, and they cut across the primary/secondary distinction. Primitive tribes have weak powers of shape-discrimination, but need to be well able to discriminate sounds.

**Primary and Secondary Qualities**

Four different grounds of distinction:

1. Direction of Scientific Explanation.
2. Observer-independent v. Observer-dependent
3. Central v. Peripheral
4. Dispositional v. Non-dispositional

Locke assimilates the observer-independent v. observer-dependent ground of distinction to a dispositional v. non-dispositional one.\(^{27}\) He compares colour with sensations of warmth, and taste; and these with pain. For something to be painful, it has the power of causing pain in a sentient being. The definition is entirely in terms of the pain which it would cause if a person were to be affected. It seems reasonable, he argues, to construe the term as a dispositional one, ascribing to the object said to be painful simply a power which in turn is characterized by the sensations of a sentient being, so that if there were no sentient beings, there would be no pain, and hence no painful objects or processes. In the same way, the word ‘poisonous’ refers to a propensity of a substance, no doubt due to its chemical structure, to damage organisms which come in contact with it: but there is no further special quality of being poisonous. And again, if there were no organisms, there would be no poisonous substances. He then suggests that just as a substance’s being poisonous is simply its having a disposition to cause damage on someone who takes it, so its being green is simply

---

but the argument does not survive scrutiny: sharp objects can cause pain, and bullets wound or kill, but even if there were no sentient beings, flakes of flint would still be sharp, and moving objects would still possess kinetic energy and momentum. Almost every quality has some relational aspect, and can therefore be construed as being at least covertly dispositional. Although secondary qualities can, indeed, be construed as being “constitutionally iffy”, so that to say that chlorophyll is green is to ascribe a disposition which manifests itself in suitable perceivers calling it ‘green’; its being dispositional no more downgrades it ontologically, than the fact that a magnetic field manifests itself in suitable magnetic needles aligning themselves along it, shows that magnetic fields do not “really exist”. Locke’s argument seems to be an attempt to shift the boundary between pukka qualities and suspect dispositions, so as to reclassify all sorts of erstwhile respectable qualities as being dispositional in the last resort. But then shapes and sizes can be construed as dispositions to remain unaltered under transformations of the Euclidean group.

Still, there is some distinction to be drawn. Being poisonous is not merely a power, but a power to produce effects in us. Bernard Williams cites the example of ‘amusing’:28 To call something amusing is on the face of it to ascribe a property, in much the same way as to call it green: but it is a property relative to human interests and tastes. It would be awkward to describe incidents in the Jurassic age as being amusing at the time they took place; and similarly, it is argued, we should not describe anything as being green then, in the absence of any observers who could see it as green. All we should do is to say that the leaves would have looked green to us, had we been there to see them, since they contained chlorophyll, whose chemical structure makes it reflect light with wavelength around 600 nanometres. But there is an important difference between something’s being amusing and its being green. Tastes and interests vary: Queen Victoria was not amused by many of the things her contemporaries found amusing, but her water colours

---

showed the same trees as green as were so observed by her sub-
jects. Against this it can be argued that ascriptions of colour are
also variable, and do, in fact, vary with language and culture. The
Greeks had no word for yellow, using the word χλωρός (chloros)
for greenish yellow, and ξανθός (xanthos) for orangeish yellow.29
But it is a different variability. There are differences over where to
draw boundaries—whether turquoise should be counted as green
or blue—but, apart from colour-blindness, no radical differences
over central cases. Were that not so, we should not identify some
people as colour-blind.

The colour-blindness argument tells both ways. It is effective
in showing that our perception of colour is not like our sense of
humour, nor radically dependent on culture and language. But it
reveals also how much our judgements about colour depend on the
majority of mankind agreeing. Where there is a serious division of
discriminatory ability—whether phenol-thio-urea tastes bitter or
not—we hesitate to attribute either bitterness or tastelessness to
it, and speak only of how it appears to different people.30 It can
then be argued by extrapolation that the ascription of colour, being
similarly observer-dependent should be construed similarly safely.
Although “locally” there is agreement, the agreement is only local,
being found among a certain range of human observers, but not
extending to members of other species, nor even to all members of
homo sapiens.31

It is quite reasonable to consider the senses of non-human an-
imals. Mammals other than primates are colour-blind. Many a
dog-owner must have wondered what a dog’s world would be like,
somewhat shadowy, no doubt, but full of exciting smells. Dogs
and birds, we believe, can hear sounds of a higher pitch than we
can, and it is a profitable topic of contemplation to wonder what
it would be like to be a bat, using sonar to locate its surround-
ings. The world of an electric eel might be even more interesting

29 This seems less strange when we look at a spectrum, and note how small
is the band that looks yellow to us.

30 Jonathan Bennett, “Substance, Reality and Primary Qualities”, American
Philosophical Quarterly, 2, 1965, part II, §1, p.9.; reprinted in C.B.Martin
p.105. A convincing counter is put forward by P.M.S.Hacker, Appearance

philosophically, though perhaps rather dull if actually experienced. But we must be wary of a covert anti-anthropomorphism in these speculations. If in our avoidance of mere “local” perspectives, we rule out all human observers, we have to be ready to say what sort of observer would be allowed. Newton’s universe was observed by God. He knew where everything was in His sensorium, because He had created everything, and sustained its continuing existence by an act of will. But God need not be colour-blind. If God is allowed as a non-local observer, we cannot exclude secondary qualities from our absolute conception of reality. If God’s non-local observing is disallowed, if we seek a conception of the world as it is independently of all observers, we are in danger of evacuating it of all empirical content. The final test is observation. An account that does not save the appearances is not relevant to us, and is not about the world we live in. We cannot purge our account of the world of secondary qualities on the ground that they are observer-dependent and therefore local, because the resulting absolute conception of reality without observation is idle.

§14.8 Conjugating Viewpoints

Although our bodies are located in physical space, and can move only relatively little and relatively slowly, our minds can range over all space and all time, as well as being able in some cases to project themselves into other men’s minds. I can imagine that I am at some other place at some other time than the here and now of my actual bodily location: when I went to catch my plane at Heathrow, I had already bought my ticket; by the time you get this letter, I shall already have left the country. In each case the ‘already’ indicates a different date from that of speaking. Reichenbach gives an illuminating account of tense structure, distinguishing a “reference point” (R) from both the time of speaking (S) and that of the event being spoken (E)about. S and E are not under our control, but we can choose R as seems most illuminating. Reichenbach is able to bring out the difference between the simple past, or aorist, tense as in ‘I went to Cambridge yesterday’ which simply says something about yesterday, namely that it was a Cambridge day, and the perfect tense, as in ‘I have been in Cambridge’, which says something

32 See §8.4 above.
about me now—for example, I could go on to say “so I know where Magdalene is”. Similarly, we can at the cost of some circumlocution distinguish the simple future, ‘I shall be dead by then’ from the analogue of the perfect ‘I am about to die’; and likewise we can say ‘I shall be about to lecture then’ and ‘I was going to write to you’ as analogues to the pluperfect and future perfect. The distinction between the future simple and the analogue to the perfect is difficult to draw, but important philosophically: the Schoolmen distinguished situations described by the latter, as ones where the future was present in its causes; and failure to keep the distinction in mind has led to many thinkers accepting some form of fatalism.

I can conjugate complex tenses, because in my mind I am unfettered in my ability to adopt different points of view; similarly, I can conjugate over person and number, because I can project onto others the emotions, intentions, ambitions and aspirations of which I have first-personal experience myself. I may be wrong, and any ascription I make may be proved wrong by irrefutable third-personal evidence. But I can be right, and just as in my own understanding of myself, I can always stand back, and review the views I had hitherto held, so I may always have fresh insights as I seek to understand other people.

Most important is the second-person singular. I come to know thee better and correct my previous misapprehensions about thee, by entering into dialogue and conversing with thee. In doing so, I not only discover thy mind, but develop my own attitudes and sympathies and insights. I become more myself in relation to thee.

It follows that there is not a simple unitary ego, as assumed by many thinkers, which remains unaltered as the person I really am. Rather, I am an overlapping set of variegated personalities, the different mes realised in different contexts with regard to different people. A portrait painter once exclaimed on seeing his subject “I see seven archbishops: which one am I to paint?”. So it is with all of us. The father, the Captain, the colleague, the pupil, the swain and

Modern English is unfortunate in having almost lost the words ‘thou’, ‘thee’ and ‘thine’. French and German have still kept it, and can distinguish intimate togethernesses, where it is appropriate tutoyer from the more public occasions where I address you more distantly as vous or Sie. At the risk of seeming absurdly old-fashioned, I shall here use ‘thou’, ‘thee’ and ‘thine’ to express the personal second-person singular as distinct from the impersonal ‘you’.
the lover are different, with different, and sometimes incompatible, characteristics, which cannot be massaged into one complete and coherent whole. Personality is not only infinitely complex, but to some extent indeterminate.

And yet there is only one person. A parallel with quantum mechanics suggests itself. There we reckon that there is only quantum-mechanical reality, but different operators elicit different physical properties of it; so with personality, the same self is understood and described in different ways by different persons who have inter-acted with him. We are resonators. When I am on the same frequency as thee, thou pickest up one signal coming over on that frequency, whereas when I am in the ward room, the common room, or the pub, my mates detect a different message coming over amidst the general noise.

Paolo and Francesca were consigned by Dante to purgatory. Though for a season lovers can revel in each other’s existence, sweet nothings pall, and honeymoons go sour, if the happy couple do not move on, and begin to consider what, as an item, they are going to do. There is a necessary progression from the second person singular to the first person plural (or better, in Greek, the first person dual). We are also impelled to do things together by external pressures. Although man is much less dependent on his environment than other animals or other things, he is not self-sufficient. I am mortal, and need the cooperation of another, if my name and genes are to last beyond my own death. In a multitude of other ways we need help if we are to survive and achieve our aims. Man is a social animal who needs must cooperate with his fellows.

If thou and I cooperate with each other we cooperate in a common enterprise. Together we lift the stone. The reason why each of us heaved and strained was that we wanted the stone lifted. Purposes, motives, and reasons for action generally, can be shared. They have a different logic from that of bodies, things and material objects generally. Thou and I can have the same purposes, motives and values, though not the same body: ideals are not privative, as ordinary possessions are—my having an ideal does not preclude thy having it; often the contrary. Values and aspirations can inter-penetrate one another, whereas things are impenetrable, exclusive space-occupiers. (Sometimes, of course, values are opposed, and one cannot be realised without abandoning the other; but usually
they are not inherently exclusive, and often we can find ways of realizing them both.

Cooperation is best explicated by means of the Theory of Games. In many simple cases it is obvious what each must do for both to benefit; but three crucial ones, the Rule of the Road, the Prisoners’ Dilemma, and the Battle of the Sexes, reveal external pressures on us to conjugate. The Rule of the Road arises where there are two or more outcomes with equally good pay-offs for all concerned. If we all drive on the left, we all avoid collisions; but equally if we all drive on the right. There is no way whereby each of us, guided solely by the light of nature, will arrive at the same strategy as everybody else. We need an antecedent agreement, a convention, a rule, which we all recognise and all abide by. I cannot make a go of it in the first person singular alone. I have to regard myself as one of us, and accept as my own the arbitrary convention rule agreed upon by everybody else. Language is the most important example. With few exceptions, it is only by convention that certain sounds have their meanings. There is no way we can reason out why in English ‘Nay’ indicates a negative, while in Greek ναι (nai) means ‘Yes’; and in English ‘OK’ means Yes, while in Greek ουχί (oukhi) indicates a negative. Words mean what they mean by convention, not nature. But once established, conventions constrain us as well as enabling us. I may be an anarchist, and deny the right of the Queen-in-parliament to make laws about how I should drive my car, but if I am wise, I will drive on the left, simply because that is what other, unliberated, drivers do, and I would rather pass them safely than have a head-on collision. I may be an existentialist, and seek to avoid the bad faith of doing the done thing, but unless I take care to keep to the grammatical rules of the French language, I shall not be able to share my thoughts with my fellow intellectuals, or obtain a following in the national press. These and many other ties bind me, and incorporate me into society. We have to have rules.

Rules are important, but inconveniently rigid. If I am a motorist, I need to be sure what the other driver is going to do, and I need to be able to tell on the basis of limited information. I cannot enter into an intimate and wide-ranging I-thou relationship with him, but only a limited, and therefore somewhat impersonal I-you or I-he relationship; and he likewise with me. The first-personal we implies a third-personal they. We avoid the accident, because each of us can know what other drivers in that situation are going
to do. Similarly with language, similarly with dancing, similarly with etiquette.

The Prisoners’ Dilemma proves how self-defeating the me-first strategy can be. For not all activities are mutually beneficial. If the Fierce Bad Rabbit takes the Good Rabbit’s carrot, it may be good from the Fierce Bad Rabbit’s point of view, but it is very much not so from that of the Good Rabbit. Taking whatever you want may seem like a good policy at first, but is less so if other people adopt it too, and help themselves to your possessions whenever they can. We are all losers in a free-for-all, because the benefits of taking whatever one wants, are outweighed by the disadvantage of having one’s own things nicked. We all do better if each forgoes the opportunity to steal, and enjoys the benefit of not having his possessions stolen. Since we are all worse off if we consider only our individual selves, it makes obvious sense for each of us to identify with us all; then we shall all benefit by our each forbearing to maximise his own pay-off at the expense of everyone else. Once again, I cannot make a go of it in the first person singular alone. I have to regard myself as one of us, and be ready to forgo my own individual interest for the sake of our collective good.

The Battle of the Sexes shows that it is self-defeating to consider only future outcomes. If I resolutely let bygones be bygone, and take no account of what has already been done, I lay myself open to manipulation by unscrupulous others. We need to conjugate over the past as well as the future. Just as we make the transition from the immediate present to the future, as we urge agents not to go just for immediate pleasure, but to consider the long-term future as well, so we are impelled to widen the range of our consideration, and to take in the past as well as the future, since it turns out to be imprudent to be guided by prudential considerations alone.

It is tempting to generalise, and identify with all humanity over all time, adopting utilitarianism as the policy that allows personhood to flourish in caring for the whole of mankind. But there are many fellow human beings. If one is to love them all equally, one cannot love any one of them specially. The Utilitarian can have an I-them relationship with humanity in general, perhaps even rising to an I-you one, but not an I-thou relationship with anyone in particular. I can appreciate his rational benevolence, and may benefit from it, but in its impartial administration to all and sundry it will again seem impersonal, and un-self-enhancing. Utilitarianism treats men as units, not as individuals, recognising that different
outcomes have different pay-offs for different people, but not accommodating each person as an agent, with his own interests and ideals he aspires to achieve by his own actions.

We need to conjugate. A full concept of personality needs to consider the possible as well as the actual, in order to delimit separate selves, and to take into account the past and the future as well as the present, and to recognise that the first person singular must be accompanied by the second person singular and the first person plural, if it is to develop into a coherent concept. These developments engender further tensions: Each is a definite individual, ultimately responsible for what he decides to do, while being also an indeterminate shimmering of different personalities, revealed and developed in different personal relationships. Each is unique, of infinite complexity, transcending all stereotypes and neat classification, while needing also to be a safe pair of hands, who can be relied on to do his bit when required.