

# Locke's Doctrine of Substantial Identity & Diversity

by

C. D. BROAD

---

Locke discusses the subject of identity and diversity of substance in Chapter 27 of Book II of his *Essay on Human Understanding*. The question to be considered is this. Under what circumstances would it be correct to say of a certain thing, which occupies a certain place at a certain time, that it is *the same thing as* or *a different thing from* one which occupied the same or a different place at a certain other time?

(1) *General Principles*. LOCKE takes it to be self-evident that one thing cannot occupy two places at the same time. I think that this needs more careful statement. What is the place occupied by a sponge at any moment? Is it the whole of the space which would be enclosed by a flexible sponge-bag which touches the exterior surface at as many points as possible? Or does it exclude the holes in the sponge? I think that we should sometimes be inclined to take the former and sometimes the latter definition. If we take the more rigid view, we could define the statement that body *B* occupies place *P* at time *t* as follows. It means that every part of *B* is located within a sub-region of *P*, and every sub-region of *P* has a part of *B* located within it, at the moment *t*.

Now what is regarded as a single thing includes three different possible cases. (i) It may be strictly *continuous*, as an extended atom was supposed to be. (ii) It may be *connected*, but

have holes in it, i. e., be *porous*. A sponge would be an example of this, if the sponge-material surrounding the holes were strictly continuous. The essential point here is that from any region occupied by a part of the body to any other region occupied by a part of the body there are innumerable paths which do not pass through any region not occupied by a part of the body, although there are such regions, viz., the holes. (iii) It may be *disconnected*. An example would be a heap of stones considered as a single thing. In this case the place strictly occupied by the thing is a set of regions disconnected with each other.

I take it that LOCKE was thinking primarily of bodies which are continuous or porous. I take it that what he wished to assert is this. If a body is continuous or porous, then at any moment there is one and only one place such that every part of the body is located in some sub-region of that place and every sub-region of that place contains some part of the body. If a thing is disconnected, like a heap of stones or a heap of sponges, then the above statement can be applied to each of the continuous or porous bodies of which it is composed.

The next assertion which Locke makes is as follows. He distinguishes three kinds of substance, viz., God, finite spirits, and bodies. He sees no objection to supposing that the same place can be occupied at the same time by two substances, provided that one is of one of these kinds and the other is of another, e. g., by a body and a finite spirit, or by a substance of either of these kinds and by God. But he takes it to be self-evidently impossible that two substances of any *one* of these kinds should occupy the same place at the same time.

Now it must be admitted that there are *prima facie* counter-instances to this last principle. *Prima facie* in a vessel containing water and alcohol, which have been well mixed, the same volume is continuously occupied by water and continuously occupied by alcohol at the same time. LOCKE would have to say that this *prima facie* appearance *must* be deceptive. One possibility is that both the water and the alcohol are composed of large numbers of disconnected particles, and that the water-

particles and the alcohol-particles are interspersed with each other, as black men and white men might be in a crowd. Another possibility is that either the water or the alcohol is a porous body, with enormous numbers of small holes in it, and that each hole in the one is continuously occupied by parts of the other. What LOCKE takes to be self-evidently impossible is that any region, great or small, should be strictly filled by each of two different substances of the same kind at the same time.

LOCKE asserts further that anything which began to exist at a different time or place from that at which a thing *A* began to exist would necessarily be a different thing from *A*. Let us first take the case of different times of origin. I think that the essential premiss is that a thing lasts for a period, finite or infinite. Now suppose, if possible, that a thing *A* were to begin to exist at two different moments  $t_1$  and  $t_2$ . This implies that it first existed from  $t_1$  to some intermediate moment  $t$ , then ceased to exist, and then began to exist again at the end of the interval from  $t$  to  $t_2$ . I take it that what LOCKE regards as self-evident is that a thing which ceased to exist at a certain moment and a thing which began to exist at a later moment must be different things.

We can next deal with the case of different places of origin. If a thing cannot begin to exist at two different moments, a thing which began to exist at two different places would have to be at two different places at the *same moment*, viz., the moment of its origination. But this LOCKE regards as impossible.

Lastly, LOCKE asserts that anything of the same kind as *A* (e. g., a body, if *A* is a body, or a finite spirit, if *A* is a finite spirit), which began to exist at the same time and place as *A*, would necessarily be the same thing as *A*. This does not seem to me at all obvious in the case of finite spirits. Is it not quite plausible to suggest, e. g., that the souls of 'identical' twins begin at the same time and place, and yet are two substances?

(2) *Identity of a Particle*. Let us now ignore finite spirits and those bodies which are either porous or disconnected. Let us confine our attention to bodies which are continuous and homogene-

ous. Any such body may be called a 'particle'. (There is no reason in theory why a 'particle', thus defined, should be very small.) I am fairly sure that LOCKE held that no macroscopic body is really continuous or homogeneous, though some appear *prima facie* to be so, and that every macroscopic body is a disconnected collection of ultra-microscopic particles of various kinds.

Now consider the description 'particle occupying place  $s$  at instant  $t$ '. From one of LOCKE's principles it follows that this description applies either to no entity or to only one entity. So, if it applies to something, we can talk of 'the particle which occupies  $s$  at  $t$ '.

Consider next the two definite descriptions 'the particle which occupies  $s_1$  at  $t$ ' and 'the particle which occupies  $s_2$  at  $t$ '. From another of LOCKE's principles it follows that anything answering to the first of these descriptions must be diverse from anything answering to the second of them. So the only two questions that remain are these. (i) Do the two descriptions 'the particle which occupies  $s$  at  $t_1$ ' and 'the particle which occupies  $s$  at  $t_2$ ' apply to the same entity or to different entities, assuming that they both apply to something? (ii) Do the two descriptions 'the particle which occupies  $s_1$  at  $t_1$ ' and 'the particle which occupies  $s_2$  at  $t_2$ ' apply to the same entity or to different entities, assuming that they both apply to something? Now LOCKE's answer seems to be the following. In each case the two descriptions apply to the *same* entity, if the entity which answers to one of them *originated at the same time and place* as the entity which answers to the other of them. They apply to *different* entities, if the entity which answers to one of them *originated at a different time or a different place* from that at which the entity which answers to the other of them originated.

This criterion of identity and diversity seems to me to be completely futile in practice and by no means satisfactory in theory. It is futile in practice for the following reasons. We do not know whether particles begin to exist in the course of nature at all. And, even if we did know this, we do not know where or

when any particular particle originated. It seems to me to be unsatisfactory in theory for the following reason. In order to show, e. g., that the two descriptions 'the particle which occupied  $s_1$  at  $t_1$ ' and 'the particle which occupied  $s_2$  at  $t_2$ ' apply to the same entity, you would have to show with regard to each of them that it applies to the same entity as a certain description of the form 'the particle which originated at  $s_0$  at  $t_0$ '. Now I cannot imagine any way in which you could show this which could not equally be used *directly* to show that the two descriptions 'the particle which occupied  $s_1$  at  $t_1$ ' and 'the particle which occupied  $s_2$  at  $t_2$ ' apply to one and the same entity. It seems to me that the criterion must in any case be the presence or absence of a certain kind of spatial and qualitative continuity, bridging the gap between the two places in question during the period between the two moments in question. I cannot see why this criterion should not be applied directly to the gap between  $s_1$  and  $s_2$  and the period between  $t_1$  and  $t_2$  without dragging in a reference to a date and a place of origin.

There is one other radical criticism to be made. LOCKE assumes in all this discussion that there is no difficulty in deciding whether one is concerned with a single *place* or *date* or with different *places* or *dates*. But, if, as he holds, spatial and temporal position are purely relative and can be determined only by relations to *bodies* and to recurrent processes in *bodies*, respectively, there is surely a grave risk of the whole discussion moving in a circle.

(3) *Aggregates, Organisms, and Machines*. Having dealt with identity and diversity as applied to individual particles, LOCKE proceeds to consider their application to certain other entities.

(i) *Aggregates of Particles*. We say that the same aggregate of particles continues to exist provided that the same particles remain conjoined together. It does not matter if the relative positions of the particles should be altered, e. g., by stirring the aggregate. But if a single one of them has been removed or a single one added, we say that there is a different aggregate. This seems to me to have hardly any application in real life. One

would certainly talk, e. g., of the same heap of stones, even if a few were added or taken away. On the other hand, if the heap were spread out flat, we should be inclined to say that it had ceased to exist, even though not a single stone had been removed or added and they had all remained in contact with each other on the ground.

(ii) *Vegetable and animal Organisms*. A living organism is not a mere aggregate of adjoined particles. It consists of particles of various kinds arranged in certain characteristic ways, which together constitute what LOCKE calls 'one coherent body partaking of one common life'. For its continued existence and self-identity it is unnecessary that the same particles should continue to be present. All that is needed is that new particles taken in should replace old ones that have been lost, that they should be organized on the same characteristic plan, and that they should play their parts in maintaining the characteristic vital functions of the whole.

(iii) *Machines*. LOCKE thinks that the self-identity of a machine, e. g., a watch, does not differ profoundly from that of a living vegetable or animal body. He asserts that the main difference is the following. A watch may be in perfectly good order, and yet not perform its characteristic functions simply because it has not been wound up, i. e., not provided with energy from outside. In a living organism structure and functions are much more closely bound up with each other. It is true that a plant or an animal will not perform its characteristic movements if it is not supplied with food, water, air, etc., from outside. But, if these are withheld, it does not merely stop, like an unwound watch or a locomotive which lacks fuel or water. It begins to degenerate and soon gets into a state in which it can no longer exist or perform its characteristic functions, even if food, water, and air should again be supplied.

(4) *Identity of a Man*. In § 21 of Chapter 27 of Book II LOCKE discusses the different meanings which may be attached to the phrase 'the same individual man'. It might be taken to mean either (i) the same individual immaterial thinking sub-

stance, taken in complete abstraction from the body; or (ii) the same animal body, taken in complete abstraction from the immaterial soul; or (iii) the whole composed of the same immaterial soul and the same animal organism, with the former animating the latter. LOCKE definitely takes the second of these alternatives. In § 6 he says that the criterion for being the same man is the same *mutatis mutandis* as for any other species of animal, viz., 'a participation of the same continued life by constantly fleeting particles in succession, vitally united to the same organized body'. And in § 8 he says that 'the idea in our minds, of which the word *man* in our mouths is a sign, is nothing else but an animal of a certain form'.

The arguments which LOCKE produces in favour of this view are the following. (i) Identity of the immaterial soul is not *enough* to constitute the same man. In the first place, it is logically possible that the same immaterial soul might successively animate a number of different human bodies, e. g., those of Socrates, Judas Iscariot, and Napoleon. But, even if we knew for certain that these bodies were animated by the same soul, we should say that Socrates, Judas, and Napoleon are three men and not one and the same man. (ii) It is logically possible that a human soul should at times animate a non-human body. But LOCKE feels sure that, if a pig-keeper knew that the body of one of his pigs is animated by the soul which formerly animated the body of the emperor Heliogabalus, he would not call that pig a *man*, and therefore would not say that it is *the same man* as Heliogabalus. (iii) LOCKE tells a long story about a creature which looked like a parrot and had parrots for parents and yet carried on conversations and performed actions which showed every sign of rationality. He says that one would certainly call such a creature a *rational parrot* and not a man. On the other hand one would not hesitate to call any living being in human shape a *man*, even if it never showed any signs of rationality in speech or action.

I will now make some comments on this. (i) LOCKE considers that these remarks show that the property of being a living ani-

mal of human form is *necessary and sufficient* to entitle a thing to be called a *man*. I doubt whether they show this property to be *sufficient*. If one knew for certain that a living organism in human form was *not* animated by a human soul, would one call it a man? Would one not be more inclined to say: 'It is not really a man, but only an animal in human shape'? We generally take for granted that any living organism in human form *is* animated by a human soul, though that soul may be very defective or the body may have defects which prevent the soul from expressing itself properly in speech or action. I suspect that it is only for this reason that we are inclined to agree that to be a living animal of human form is *sufficient* to entitle a thing to be called a man.

(ii) It is more plausible to hold that the arguments prove that this property is *necessary* to entitle a thing to be called a man. I have very little doubt, e. g., that we should call such a creature as LOCKE describes a 'rational *parrot*' and not a '*man* in a parrot's body'.

(iii) I do not attach much weight to LOCKE's other remarks. I think that it is generally futile to take imaginary cases, which, so far as we know, never happen, and then ask whether people would apply or refuse to apply a certain familiar word or phrase to such a case if it did happen. Unless the circumstances supposed are described in considerable detail, the question is too vague to admit of any reasonable answer. If, on the other hand, the circumstances supposed are described in adequate detail, the only sensible answer generally is: 'Some would be inclined to apply it, some to withhold it, and some would just not know what to say; and none of them could be fairly described as using language incorrectly'.

Take, e. g., LOCKE's supposition that a certain pig-keeper *knows* that the soul of the late emperor Heliogabalus now animates the body of a certain one of his pigs. *How* is he supposed to know this fact? Are we to suppose that this pig looks and behaves just like all the others? I should have thought that any evidence that one had for believing that Heliogabalus's soul is



now animating a certain pig's body would be evidence inclining one to say that that pig is a certain man 'in disguise' or 'transformed'.

Suppose that one had a friend who was a were-wolf, and that one had repeatedly seen his body change from the human to the lupine form at sunset and back again at sunrise. One might describe the situation by saying: 'Mr. Jones is a man by day and turns into a wolf at night; he is the *same were-wolf* all the time, but he is sometimes in human and sometimes in lupine form'. But it would also be quite sensible to say of him when he is in his lupine form: 'He is not just an ordinary wolf; he is really a man transformed into a wolf'. And it would be about equally sensible to say of him when he is in his human form: 'He is not just an ordinary man; he is really a wolf transformed into a man'.

It is fairly easy for us to deal with the above situation linguistically, because it is one which has been believed to occur quite often, so that our ancestors have coined the name *were-wolf* to cover it. But suppose that I had owned an ordinary parrot; that shortly after McTAGGART's death it had laid an egg; and that, when the egg was hatched, it had given rise to a creature which looked exactly like a parrot but had talked in McTAGGART's characteristic style, enunciated and discussed philosophical propositions characteristic of McTAGGART's system, and so on. I suppose that I should hardly call this creature a *man*, and therefore should hardly say that it is *the same man* as McTAGGART. But I certainly should not be content to say that it is a rational parrot. I suppose that I should be inclined to say: 'It looks as if McTAGGART had survived and been born again as this parrot'.

(5) *Personal Identity*. LOCKE distinguishes between being the same *man* and being the same *person*. He defines a 'person' in § 9 as 'a thinking intelligent being, which has reason and reflexion and can consider itself . . . as the same thinking thing in different times and places'. He says in § 26 that 'person' is primarily a legal term, concerned with the question of what ac-

tions can properly be rewarded or punished. It therefore applies only to intelligent beings, capable of understanding laws and deliberately obeying or disobeying them, and capable of feeling pleasure and pain and recognizing that certain pleasures are rewards and certain pains are punishments for their acts of obedience or disobedience respectively.

Now LOCKE is concerned to maintain that *memory* is the sole and sufficient test for personal identity. Anything which a person remembers at a given moment as having been done or experienced was done or experienced by that person. Anything which he does not remember at a given moment as having been done or experienced was not done or experienced by him.

It seems to me that this is terribly ambiguous, and that the first thing is to try to clear up the ambiguities in it. (i) Is LOCKE referring to *actual* remembering only, or to remembering in the *dispositional* sense also?

(a) On the first alternative nothing which a person is not actually remembering at  $t$  was done or experienced by him, and anything that he is actually remembering at  $t$  was done or experienced by him. It is plain that this criterion would be hopelessly at variance with common usage. Everyone admits that what a person is actually remembering at any moment is only a tiny fraction of what he has done and experienced before then. Another absurdity which would follow is this. It would be impossible to say that the same person is at some times remembering and at other times not remembering a certain act or experience  $x$ . For the person who at  $t_1$  *is* remembering  $x$  *is* the person who did or experienced  $x$ ; and the person who at  $t_2$  *is not* remembering  $x$  *is not* the person who did or experienced  $x$ . Therefore the person who at  $t_1$  is remembering  $x$  must be other than the person who at  $t_2$  is not remembering  $x$ . But this is absurd.

(b) Let us then try the dispositional sense of 'remember'. LOCKE's doctrine then comes to the following. Anything which a person actually does remember or could remember at a given moment as having been done or experienced *was* done or experienced by that person. Anything which he neither does nor could

remember at a given moment as having been done or experienced was *not* done or experienced by him.

This is unsatisfactory also. How much is implied by 'could' here? Does it mean just 'could, if he had tried at the time'? On this interpretation we shall be back at our old difficulties. Everyone would admit that each person has done many acts and had many experiences which he has forgotten, and which henceforth he could not remember however hard he might try. Again, it is quite common to find that at a certain moment one cannot remember having done or experienced *x*, however hard one may try, whilst at a later moment one can and does remember it. On the present criterion it would be contradictory to say that one and the same person sometimes can and sometimes cannot remember the same act or experience.

Suppose, on the other hand, that we widen the sense of 'could' in order to avoid this difficulty. How far is it to be widened? Is it to include 'could, if he had been hypnotized or psycho-analyzed by a sufficiently skilful practitioner'? Or even to include 'could, if God were miraculously to stir his memory'? Or what?

(ii) We can now consider another ambiguity in the word 'remember', which cuts across the one which we have just been discussing. It is common to use the word 'remember' in such a way that it would be contradictory to say '*A* remembers *x*'; but *x* either never happened at all or, if it did, it was done or experienced by someone other than *A*'. Now it is notorious that there are experiences which are introspectively indistinguishable from remembering, but are *delusive*. Let us call all experiences which have the introspective character of rememberings 'ostensible rememberings'. We can then divide these into *veridical* and *delusive*. And we can say that the word 'remembering' is commonly used to mean veridical ostensible remembering.

Now is LOCKE using 'remember', in his criterion of personal identity, in the sense of *veridical* ostensible remembering or in the sense of ostensible remembering *without qualification*? On the former interpretation it is quite trivial to say that anything

which a person remembers at a given moment as having been done or experienced was done or experienced by that person. For that is part of the definition of 'remembering' in that sense of the word. On the second interpretation the statement is not trivial, but it is either false or completely at variance with the common usage of the phrase 'same person'. For a person may ostensibly remember something which never happened at all, or something which was done or experienced by what we should all unhesitatingly describe as another person. George IV in his declining years ostensibly remembered leading a charge at the Battle of Waterloo, and would often describe this alleged incident in his life to the Duke of Wellington. But everyone knows that he was never near the battle and that all charges in it were led by persons other than he.

Let us next consider the question of just reward and punishment. One has a good deal of sympathy with the view that it is unjust, or at any rate what we call 'hard lines', to *punish* a person for an act, done by his body, which he cannot now remember however hard he may try. But suppose that a soldier had done a deed of great heroism, and had been badly wounded and thereafter could not remember doing the act however hard he tried. Should we feel that it was unjust to give him the V. C.? Should we not rather be inclined to say that it would be unjust to refuse to give it to him merely because his memory had become deranged?

Suppose, again, that a person were to find in his diary a statement in his own handwriting that he had committed a certain crime on a certain day. Suppose that he could remember the events recorded immediately before and immediately afterwards, and that there was not the least reason to think that this entry was, or could have been, forged. Suppose that he horror of the crime, or some other psychological cause, makes it impossible for him to remember doing the action, even when he reads the entry which his hand has made in his diary. Should we feel *certain* that it would be unjust to punish him for this crime? Would even *he* feel certain of this?

Again, would anyone say that George IV deserved a decoration for the gallantry which he ostensibly but delusively remembered that he had displayed at the Battle of Waterloo? And, if he had ostensibly but delusively remembered that he had displayed disgraceful cowardice, would anyone say that he ought to have been court-martialled and shot for it?

LOCKE admits that a man is often punished by the law-courts for acts which he says that he cannot remember having done. His answer is that human law can administer only rough justice. If the man is telling the truth as to what he can and cannot remember, it *is* morally unjust to punish him. But the law must take some risks. We can never be sure that such a man is telling the truth about his memories; and we can be quite certain that, if the allegation that one could not remember doing an act were made a sufficient ground for not inflicting punishment, the majority of criminals would get off by telling this particular kind of lie. At the Last Judgment God will reward and punish each man for *all* that he can remember and *only* for what he can remember. God will know exactly what this is, and it will be idle for any man to lie to God about the extent of his memory.

I think it is plain that LOCKE's account of personal identity wholly in terms of actual or possible memory will not do. But I think that this kind of theory could be made less hopelessly unpalatable if it were transformed on the following lines. Suppose we were to say that the person who is now having the experience *x* is the same as the person who formerly had the experience *z*, if and only if *either* (i) he is remembering or could at will now remember *z*; *or* (ii) he is remembering or could at will now remember an experience *y*, such that at the time when he was having *y* he did remember or could have remembered *z*; *or* (iii) he is now remembering or could at will now remember an experience *y*, such that at the time when he was having *y* he was remembering or could have remembered an experience *w*, such that at the time when he was having *w* he was remembering or could have remembered *z*; *or* so on . . .

The point of the amendment is this. There need not be a

single direct memory (actual or possible), spanning the gap between the occurrence of  $x$  and the occurrence of  $z$ . It is enough that the gap shall be filled by any finite number of 'piers', with a different memory (actual or possible) spanning the gaps between the successive 'piers', in the way described above.