
Plato

Plato (429–347 BCE) is so important to philosophy's inceptions and subsequent development as a discipline that the philosopher and mathematician Alfred North Whitehead once wrote, "The safest general characterization of the European philosophical tradition is that it consists of a series of footnotes to Plato."¹ This excerpt from *Republic*, Bk. 7, the famous allegory of the cave, introduces readers to Plato's theory of the forms and his views on education and the philosopher's proper role in society.

Cave Allegory

From *The Republic*

Characters: Socrates, Glaucon

1. "Next," said I [Socrates], "compare our nature in respect of education and its lack to such an experience as this. Picture men dwelling in a sort of subterranean cavern with a long entrance open to the light on its entire width. Conceive them as having their legs and necks fettered from childhood, so that they remain in the same spot, able to look forward only, and prevented by the fetters from turning their heads. Picture further the light from a fire burning higher up and at a distance behind them, and between the fire and the prisoners and above them a road along which a low wall has been built, as the exhibitors of puppet-shows have partitions before the men themselves, above which they show the puppets." "All that I see," he [Glaucon] said. "See also, then, men carrying past the wall implements of all kinds that rise above the wall, and human images and shapes of animals as well, wrought in stone and wood and every material, some of these bearers presumably speaking and others silent." "A strange image you speak of," he said, "and strange prisoners." "Like to us," I said; "for, to begin with, tell me do you think that these men would have seen anything of themselves or of one another except the shadows cast from the fire on the wall of the cave that fronted them?" "How could they," he said, "if they were compelled to hold their heads unmoved through life?" "And again, would not the same be true of the objects carried past them?" "Surely." "If then they were able to talk to one another, do you

not think that they would suppose that in naming the things that they saw they were naming the passing objects?" "Necessarily." "And if their prison had an echo from the wall opposite them, when one of the passersby uttered a sound, do you think that they would suppose anything else than the passing shadow to be the speaker?" "By Zeus, I do not," said he. "Then in every way such prisoners would deem reality to be nothing else than the shadows of the artificial objects." "Quite inevitably," he said. "Consider, then, what would be the manner of the release and healing from these bonds and this folly if in the course of nature something of this sort should happen to them: When one was freed from his fetters and compelled to stand up suddenly and turn his head around and walk and to lift up his eyes to the light, and in doing all this felt pain and, because of the dazzle and glitter of the light, was unable to discern the objects whose shadows he formerly saw, what do you suppose would be his answer if someone told him that what he had seen before was all a cheat and an illusion, but that now, being nearer to reality and turned toward more real things, he saw more truly? And if also one should point out to him each of the passing objects and constrain him by questions to say what it is, do you not think that he would be at a loss and that he would regard what he formerly saw as more real than the things now pointed out to him?" "Far more real," he said.

2. "And if he were compelled to look at the light itself, would not that pain his eyes, and would he not turn away and flee to those things which he is able to discern and regard them as in very deed more clear and exact than the objects pointed out?" "It is so," he said. "And if," said I, "someone should drag him thence by force up the ascent which is rough and steep, and not let him go before he had drawn him out into the light of the sun, do you not think that he would find it painful to be so haled along, and would chafe at it, and when he came out into the light, that his eyes would be filled with its beams so that he would not be able to see even one of the things that we call real?" "Why, no, not immediately," he said. "Then there would be need of habituation, I take it, to enable him to see the things higher up. And at first he would most easily discern the shadows and, after that, the likenesses or reflections in water of men and other things, and later, the things themselves, and from these he would go on to contemplate the appearances in the heavens and heaven itself, more easily by night, looking at the light of the stars and the moon, than by day the sun and the sun's light." "Of course." "And so, finally, I suppose, he would be able to look upon the sun itself and see its true nature, not by reflections in water or phantasms of it in an alien setting, but in and by itself in its own place." "Necessarily," he said. "And at this point he would infer and conclude that this it is that provides the seasons and the courses of the year and presides over all things in the visible region, and is in some sort the cause of all these things that they had seen." "Obviously," he said, "that would be the next step." "Well then, if he recalled to mind his first habitation and what passed for wisdom there, and his fellow-bondsmen, do you not think that he would count himself happy in the change and pity them?" "He would indeed." "And if there had been honours and commendations among them which they bestowed on one another and prizes for the man who is quickest to make out the shadows as they pass and best able to remember their customary precedences, sequences and co-existences, and so most successful in guessing at what was to come, do you think he would be very keen about such

rewards, and that he would envy and emulate those who were honoured by these prisoners and lorded it among them, or that he would feel with Homer (*Odyss.* 11.489) and 'greatly prefer while living on earth to be serf of another, a landless man', and endure anything rather than opine with them and live that life?" "Yes," he said, "I think that he would choose to endure anything rather than such a life." "And consider this also," said I, "if such a one should go down again and take his old place would he not get his eyes full of darkness, thus suddenly coming out of the sunlight?" "He would indeed." "Now if he should be required to contend with these perpetual prisoners in 'evaluating' these shadows while his vision was still dim and before his eyes were accustomed to the dark—and this time required for habituation would not be very short—would he not provoke laughter, and would it not be said of him that he had returned from his journey aloft with his eyes ruined and that it was not worthwhile even to attempt the ascent? And if it were possible to lay hands on and to kill the man who tried to release them and lead them up, would they not kill him?" "They certainly would," he said.

3. "This image then, dear Glaucon, we must apply as a whole to all that has been said, likening the region revealed through sight to the habitation of the prison, and the light of the fire in it to the power of the sun. And if you assume that the ascent and the contemplation of the things above is the soul's ascension to the intelligible region, you will not miss my surmise, since that is what you desire to hear. But God knows whether it is true. But, at any rate, my dream as it appears to me is that in the region of the known the last thing to be seen and hardly seen is the idea of good, and that when seen it must needs point us to the conclusion that this is indeed the cause for all things of all that is right and beautiful, giving birth in the visible world to light, and the author of light and itself in the intelligible world being the authentic source of truth and reason, and that anyone who is to act wisely in private or public must have caught sight of this." "I concur," he said, "so far as I am able." "Come then," I said, "and join me in this further thought, and do not be surprised

that those who have attained to this height are not willing to occupy themselves with the affairs of men, but their souls ever feel the upward urge and the yearning for that sojourn above. For this, I take it, is likely if in this point too the likeness of our image holds." "Yes, it is likely." "And again, do you think it at all strange," said I, "if a man returning from divine contemplations to the petty miseries of men cuts a sorry figure and appears most ridiculous, if, while still blinking through the gloom, and before he has become sufficiently accustomed to the environing darkness, he is compelled in courtrooms or elsewhere to contend about the shadows of justice or the images that cast the shadows and to wrangle in debate about the notions of these things in the minds of those who have never seen justice itself?" "It would be by no men strange," he said. "But a sensible man," I said, "would remember that there are two distinct disturbances of the eyes arising from two causes, according as the shift is from light to darkness or from darkness to light, and, believing that the same thing happens to the soul too, whenever he saw a soul perturbed and unable to discern something, he would not laugh unthinkingly, but would observe whether coming from a brighter life its vision was obscured by the unfamiliar darkness, or whether the passage from the deeper dark of ignorance into a more luminous world and the greater brightness had dazzled its vision. And so he would deem the one happy in its experience and way of life and pity the other, and if it pleased him to laugh at it, his laughter would be less laughable than that at the expense of the soul that had come down from the light above." "That is a very fair statement," he said.

4. "Then, if this is true, our view of these matters must be this, that education is not in reality what some people proclaim it to be in their professions. What they aver is that they can put true knowledge into a soul that does not possess it, as if they were inserting vision into blind eyes." "They do indeed," he said. "But our present argument indicates," said I, "that the true analogy for this indwelling power in the soul and the instrument whereby each of us apprehends is that of an eye that could not be

converted to the light from the darkness except by turning the whole body. Even so this organ of knowledge must be turned around from the world of becoming together with the entire soul, like the scene-shifting periact³ in the theatre, until the soul is able to endure the contemplation of essence and the brightest region of being. And this, we say, is the good, do we not?" "Yes." "Of this very thing, then," I said, "there might be an art, an art of the speediest and most effective shifting or conversion of the soul, not an art of producing vision in it, but on the assumption that it possesses vision but does not rightly direct it and does not look where it should, an art of bringing this about." "Yes, that seems likely," he said. "Then the other so-called virtues of the soul do seem akin to those of the body. For it is true that where they do not pre-exist, they are afterwards created by habit and practice. But the excellence of thought, it seems, is certainly of a more divine quality, a thing that never loses its potency, but, according to the direction of its conversion, becomes useful and beneficent, or, again, useless and harmful. Have you never observed in those who are popularly spoken of as bad, but smart men, how keen is the vision of the little soul, how quick it is to discern the things that interest it, a proof that it is not a poor vision which it has, but one forcibly enlisted in the service of evil, so that the sharper its sight the more mischief it accomplishes?" "I certainly have," he said. "Observe then," said I, "that this part of such a soul, if it had been hammered from childhood, and had thus been struck free of the leaden weights, so to speak, of our birth and becoming, which attaching themselves to it by food and similar pleasures and gluttonies turn downwards the vision of the soul—if, I say, freed from these, it had suffered a conversion towards the things that are real and true, that same faculty of the same men would have been most keen in its vision of the higher things, just as it is for the things toward which it is now turned." "It is likely," he said. "Well, then," said I, "is not this also likely and a necessary consequence of what has been said, that neither could men who are uneducated and inexperienced in truth ever adequately preside over a state, nor could those who had been permitted to

linger on to the end in the pursuit of culture—the one because they have no single aim and purpose in life to which all their actions, public and private, must be directed, and the others, because they will not voluntarily engage in action, believing that while still living they have been transported to the Islands of the Blest." "True," he said. "It is the duty of us, the founders, then," said I, "to compel the best natures to attain the knowledge which we pronounced the greatest, and to win to the vision of the good, to scale that ascent, and when they have reached the heights and taken an adequate view, we must not allow what is now permitted." "What is that?" "That they should linger there," I said, "and refuse to go down again among those bondsmen and share their labours and honours, whether they are of less or of greater worth." "Do you mean to say that we must do them this wrong, and compel them to live an inferior life when the better is in their power?"

5. "You have again forgotten, my friend," said I, "that the law is not concerned with the special happiness of any class in the state, but is trying to produce this condition in the city as a whole, harmonizing and adapting the citizens to one another by persuasion and compulsion, and requiring them to impart to one another any benefit which they are severally able to bestow upon the community, and that it itself creates such men in the state, not that it may allow each to take what course pleases him, but with a view to using them for the binding together of the commonwealth." "True," he said, "I did forget it." "Observe, then, Glaucon," said I, "that we shall not be wronging, either, the philosophers who arise among us, but that we can justify our action when we constrain them to take charge of the other citizens and be their guardians. For we will say to them that it is natural that men of similar quality who spring up in other cities should not share in the labours there. For they grow up spontaneously from no volition of the government in the several states, and it is justice that the self-grown, indebted to none for its breeding, should not be zealous either to pay to anyone the price of its nurture. But you we have engendered for yourselves and the rest of the city to be, as it were, king-bees and leaders in the hive. You

have received a better and more complete education than the others, and you are more capable of sharing both ways of life. Down you must go then, each in his turn, to the habitation of the others and accustom yourselves to the observation of the obscure things there. For once habituated you will discern them infinitely better than the dwellers there, and you will know what each of the 'idols' is and whereof it is a semblance, because you have seen the reality of the beautiful, the just, and the good. So our city will be governed by us and you with waking minds, and not, as most cities now which are inhabited and ruled darkly as in a dream by men who fight one another for shadows and wrangle for office as if that were a great good, when the truth is that the city in which those who are to rule are least eager to hold office must needs be best administered and most free from dissension, and the state that gets the contrary type of ruler will be the opposite of this." "By all means," he said. "Will our alumni, then, disobey us when we tell them this, and will they refuse to share in the labours of state each in his turn while permitted to dwell the most of the time with one another in that purer world?" "Impossible," he said: "for we shall be imposing just commands on men who are just. Yet they will assuredly approach office as an unavoidable necessity, and in the opposite temper from that of the present rulers in our cities." "For the fact is, dear friend," said I, "if you can discover a better way of life than office-holding for your future rulers, a well-governed city becomes a possibility. For only in such a state will those rule who are really rich, not in gold, but in the wealth that makes happiness—a good and wise life. But if, being beggars and starvelings from lack of goods of their own, they turn to affairs of state thinking that it is thence that they should grasp their own good, then it is impossible. For when office and rule become the prizes of contention, such a civil and internecine strife destroys the office-seekers themselves and the city as well." "Most true," he said. "Can you name any other type or ideal of life that looks with scorn on political office except the life of true philosophers?" I asked. "No, by Zeus," he said. "But what we require," I said, "is that those who take office should not be lovers of rule.

Otherwise there will be a contest with rival lovers." "Surely." "What others, then, will you compel to undertake the guardianship of the city than those who have most intelligence of the principles that

are the means of good government and who possess distinctions of another kind and a life that is preferable to the political life?" "No others," he said.

Notes

1. Alfred North Whitehead, *Process and Reality: An Essay in Cosmology* (New York: Humanities Press, 1929), 63.—LG
2. Tr.: An obvious allusion to the fate of Socrates.
3. Tr.: Probably a reference to the triangular prisms

on each side of the stage. They revolved on an axis and had different scenes painted on their three faces.

4. Tr.: The world of ideas, the upper world as opposed to that of the cave.

Study Questions

1. Draw a picture that illustrates Plato's cave allegory. Explain how the goings-on of the cave serve as an allegory for Plato's theory of the forms.
2. How do the concrete and particular objects of our sense-experience relate to the abstract and universal

objects that inhabit Plato's realm of the forms?

3. What does Plato's cave allegory suggest about how students should be educated? Once educated, what is the proper role of the philosopher in society, according to Plato?

Plato

Plato (429–347 BCE) was a citizen of the ancient Greek city-state of Athens. Plato's works are not treatises but take the form of dialogues, in which Socrates is usually the main character. Plato's family knew Socrates, and Plato probably began discussions with Socrates at an early age. Plato is thought to have travelled widely after Socrates' death in 399 BCE, returning to Athens around 387 BCE to found a school (a *Mouseion*, dedicated to the Muses) in the grove of the hero Academe. In *Meno*, Plato investigates the question of what excellence (or virtue) is and introduces a theory of knowledge as remembering (or recollection).—JS

From *Meno*¹

Characters: Meno, Socrates, Meno's slave

1. What Is Excellence?

MENO. Can you tell me, Socrates, is excellence teachable? Or is it not teachable, but acquired by practice? Or is it neither acquired by practice nor learned, but something which comes to men by nature or in some other way?

SOCRATES. Meno, the Thessalians used to be famous and much admired among the Greeks for their horsemanship and their wealth; nowadays, I think, they have a reputation for wisdom as well, especially the citizens of **Larissa**, where your companion Aristippus comes from. For this change in you, Gorgias is responsible because when he arrived at your city, he made lovers of his wisdom out of the highest of the Aleuadai, among them your own lover Aristippus, and the other Thessalians too. What is more, he made you adopt the habit of answering confidently and with great style whatever anyone asked you, as befits those who have knowledge, since he is himself ready to answer any Greek who wishes to ask him anything whatever, and never lacks an answer. But here at Athens, my dear Meno, the opposite has happened. It's as if the supply of wisdom had dried up, and perhaps it has left these parts to be with you. At any rate, if you propose to put that question to people here, there isn't a single person who won't laugh and say "Stranger, I must seem to you to be a lucky man—at least if you think that I know whether excellence is teachable or how

it comes. But I'm so far from knowing whether excellence is teachable or not, that I don't even know what excellence itself is at all." So that's my position too, Meno: I'm just as poor as my fellow-citizens in this respect, and blame myself for not knowing about excellence at all. But where I don't know what something is, how could I know what sort of a thing it is? Or do you think that it is possible for someone, who doesn't know at all who Meno is, to know whether he is beautiful or rich or nobly-born, or the opposite of these? Do you really think that's possible?

M. I don't. But do you, Socrates, really not even know what excellence is? Are we to take home this news about you?

S. Not only that, but you can add that I've never yet met anyone else who did, as far as I can tell.

M. What! Didn't you meet Gorgias when he was here?

S. I did.

M. Then you didn't think that he knew?

S. I don't have a very good memory, Meno, so I can't say now what I thought then. But perhaps he does know, and you too know what he said: so do remind me of how he spoke. Or, if you like, tell me yourself. I suppose that you think as he does.

M. I do.

S. Then let's forget about him, since anyway he isn't here: but, for heaven's sake, Meno, what do you say excellence is? Do tell me and don't refuse to give me the chance of discovering that I am fortunate to be caught out in a falsehood, if it proves that you and Gorgias do know, when I said that I'd never yet met anyone who knew.

M. It's not difficult to explain, Socrates. First of all, let's take a man's excellence, that's easy. A man's excellence is this: to be capable in handling the affairs of the city, and, while so doing, to help his friends and harm his enemies, taking care not to suffer anything of the sort himself. Now take a woman's excellence, that's not difficult to describe: she must manage the household well, watch over all household affairs and be obedient to her husband. Different again is the excellence of a child, male or female, so too that of an older man, whether free or slave. And there are a great many other different excellences, so that one is at no loss to say about excellence what it is: corresponding to each action and time of life, and to every function, there is, for each of us, the appropriate excellence and likewise, I'd say, Socrates, the appropriate defect.

S. I seem to have had very good luck, Meno: I was looking for a single excellence and I discover a swarm of them all around you. But, Meno, to pursue this image of swarms, suppose I asked you about the essential nature of bees, what it is, and you said that there were many different kinds of bees, how would you answer me if I asked "Do you hold that they are of many kinds and different from one another in respect of being bees? Isn't it rather that they don't differ at all in this respect, but only in others, as in beauty or size or the like?" Tell me, what answer would you give to that question?

M. What I'd say is that one is no different from another insofar as they are bees.

S. Suppose I said next: "Then tell me this, Meno: what do you say this is, in respect of which they are no different but all the same? You'd have a reply for me, I think?"

M. I would.

S. Then do the same for the excellences: even if they are of many different kinds, at any rate they have this one common character, because of which they are all excellences, and which it's as well to keep your eye on if you're giving an answer to the question: "What is excellence?" Or don't you understand what I mean?

M. I think I do: but I don't yet grasp the question as well as I'd like to.

S. In your opinion, Meno, is it only in the case of excellence that there's a different one for a man and for a woman and so on? Or do you think that the same holds for health and size and strength? Do you think that there is one kind of health for a man and another for a woman? Or is it the same character everywhere, provided that it's health, whether it's in a man or in anything else whatever?

M. Health seems to me to be the same whether in man or woman.

S. Surely this applies to size and strength as well? If a woman is strong, will she be strong because of the same kind of thing, the same strength? What I mean by "the same" here is this: strength is no different in respect of being strength whether it is in a man or in a woman. Or do you think that there is a difference?

M. I don't.

S. Will excellence be any different in respect of being excellence whether it is in a child or an elder, in a woman or in a man?

M. I'm inclined to think, Socrates, that this case is not the same as those other ones.

S. Well, didn't you say that a man's excellence was to run a city well, while a woman's was to run a house well?

M. I did.

S. Is it possible to run a city or a house or anything else well, if one doesn't run it sensibly and justly?

M. No.

S. And surely, if they run anything justly and sensibly, they will run it with justice and good sense?

M. Necessarily.

S. So both a man and a woman will need the same things, justice and good sense, if they are going to be good?

M. So it appears.

S. What about a child or an elder? Could they ever possibly be good if they lacked good sense or were unjust?

M. Indeed not.

S. But if they were sensible and just, they could be?

M. Yes.

S. So everyone is good in the same way; for it is by acquiring the same things that they become good.

M. It seems so.

S. Now surely, they wouldn't all be good in the same way, unless it was the same excellence they possessed.

M. Indeed not.

S. Then since they all possess the same excellence, try to tell me or to remember what Gorgias, and you too, say it is.

M. What else can it be but to be capable of ruling people? If, that is, you are looking for one thing that applies to every case.

S. Indeed I am. But will a child possess this same excellence, Meno, and will a slave be capable of ruling his master? Do you think such a ruler would be a slave any longer?

M. No, I don't think that at all, Socrates.

S. No, it's not likely, my friend: and consider this next point. You say "to be capable of ruling." Shouldn't we add "justly and not unjustly"?

M. I agree because justice, Socrates, is excellence.

S. Excellence, Meno, or *an* excellence?

M. What do you mean?

S. Like any other case. As, for example, I'd say of roundness that it is *a* shape, not simply that it is shape. I'd say this because there are other shapes as well.

M. You'd be right to say that, and I'd say that justice is not alone; there are other excellences as well.

S. What are they? Do tell me. Just as I would tell you of other shapes, if you asked me to, so you tell me of other excellences.

M. I think that courage is an excellence and so are good sense and wisdom and nobility of style, and lots of others.

S. The same thing has happened to us again, Meno: when we were looking for one excellence, we've found lots, though in a different way from just now. But we can't discover that single one which runs through all of them.

M. I can't yet, Socrates, grasp the one excellence that you seek, applying to every case, as in the other examples.

S. That's quite understandable: but I will try, if I can, to get us closer to it. I'm sure you understand that this is a general point: if anyone had asked you as I did just now, "What is shape, Meno?" then if you had told him it was roundness and if he had said to you as I did "Is roundness shape or *a* shape?" you'd have said, I think, that it was *a* shape.

M. Of course.

S. Presumably because there are other shapes as well?

M. Yes.

S. And if he'd then asked you what those were, you'd have said?

M. I would.

S. And if he'd asked you about colour in the same way, what it is, and if when you'd said that it was white, he had followed up your reply by asking "Is white colour or *a* colour?" You'd have said it was a colour because there happen to be others?

M. I would.

S. And if he asked you to cite other colours, you'd have mentioned others, which are colours no less than white is?

M. Yes.

S. So if he pursued the argument, as I did, and said "We always come across several, but I don't want that. Since you speak of those many things by one name, and you say that there is not one of them which is not shape, although they are also opposites of one another, what is this thing, which embraces the round no less than the straight, which you call shape, and say that the curved is no more shape than the straight is?" Isn't that what you say?

M. I do.

S. Well, when you say that, do you hold that the round is no more round than straight, or the straight no more straight than round?

M. Of course not, Socrates.

S. But you do hold that the round is no more shape than the straight is, and vice versa.

M. That's true.

S. What, then, is it, that thing whose name is "shape"? Try to explain. If you'd replied to the man asking you these questions about shape or colour by saying "But I don't understand what you want, sir, I don't know what you mean," he would probably have been surprised and said "Don't you understand that I am looking for what is the same in all these cases?" Wouldn't you even have anything to say along these lines, if someone asked you "What is it that applies to the round and the straight and to everything else that you call shapes, which is the same in all of them?" Try to say, to give yourself some practice for the answer about excellence.

M. No, Socrates, you answer.

S. You want me to humour you?

M. Yes.

S. Will you give me your answer about excellence, if I do?

M. I will.

S. Then I must try hard: it will be worth it.

M. It certainly will.

S. Well then, let's try to tell you what shape is. See if you accept it as this: let us take shape to be the only one of the things that are which always accompanies colour. Will that satisfy you or are you looking for some other kind of answer? I'd be delighted if you would give me an answer like that about excellence.

M. But that's a naive answer, Socrates.

S. How so?

M. On your account, shape is what always accompanies colour. Well, suppose that someone said he didn't know what colour was, but was just as much at a loss here as about shape. What answer would you think you had given him?

S. A true one: and if my questioner was one of those clever, contentious and argumentative people, I'd tell him: "I've had my say: if it isn't right, it's your job to take up the argument and refute it." But if friends like you and me want to have a discussion, they must make their replies more considerate and more properly dialectical. To answer more dialectically is perhaps not only to give true answers, but also to reply in terms that the questioner agrees that he knows. I'll try to reply to you like that. Tell me: is there something you call an end? I mean something that is a limit or an extremity—I say that all these are one and the same: perhaps **Prodicus** would distinguish them for us, but I assume that you would say that anything which has been limited has also been ended—that's all I mean, nothing complicated.

M. I do say that, and I think I understand what you mean.

S. Well, is there something that you call a plane and something else you call a solid, for example those mentioned in geometry?

M. I do.

S. Then you should already understand from those what I mean by shape. Because I say of every shape that it is that in which a solid ends—which I could put shortly as "shape is the limit of a solid."

M. And what do you say colour is, Socrates.

S. You're quite shameless, Meno: you keep demanding answers from an old man, when you won't use your memory and explain what Gorgias says excellence is.

M. But I will answer you, when you've explained this, Socrates.

S. Anyone could tell, Meno, even if he were blindfolded, that you are beautiful and still have lovers from the way you talk.

M. Why?

S. Because in your conversation, you do nothing but lay down the law, as spoilt boys do, who are tyrants as long as they're attractive. Besides you've realized that I can't resist anything beautiful: so I'll satisfy you and give you an answer.

M. Yes, do satisfy me.

S. Do you want me to give you an answer in Gorgias' style, which you would find easiest to follow?

M. Of course I'd like that.

S. You both hold that things that have effluences, as Empedocles' theory states?

M. Yes indeed.

S. And that they have passages into which and through which these effluences travel?

M. Yes.

S. And that some of the effluences fit certain of the passages, but others are too small or too large?

M. That's so.

S. Is there something you call sight?

M. There is.

S. Given this, then, "See what it is I mean" as Pindar says. Because colour is an effluence shapes have, which is commensurate with sight and perceived by it.

M. I think you've given an excellent answer there, Socrates.

S. Probably because it's the kind you're used to hearing. And no doubt you see that you could go on from this to say what sound is, and smell and lots of other things of that kind.

M. Yes, indeed.

S. It sounds a grand answer, Meno, so you prefer it to the answer about shape.

M. I do.

S. But I've convinced myself that it isn't the best one—the other one is. And I suspect that you would think so, if you didn't, as you said yesterday, have to go away before the mysteries, but could stay here and be initiated.

M. Oh but I would stay, Socrates, if you kept on talking to me like that.

S. You can be sure that I shan't be lacking in enthusiasm for talking like that, for my sake as well as yours: but I'm afraid that I won't be able to say much more in that vein. Come on now, you have a try, and do as you promised, explain excellence to me as a whole, saying what it is, and stop making a multitude out of one, as jokers say when people break things; leave it whole and tell me what excellence is when intact. You've had some examples from me. [. . .]

2. Learning as Remembering

MENO. Socrates, I was told even before I met you that you're always puzzled yourself and that you make other people puzzled too. And now, I'd say, you are using spells and drugs and positively bewitching me, so that I'm just a mass of perplexity. And, to be frivolous for a moment, you seem to me to be exactly like one of those flat sea-going electric rays; both in your appearance and in other ways. Because it numbs anyone who approaches and touches it, and that's just what you seem to have done to me now. My mind and my lips have really gone numb and I don't know what reply to give you. Yet thousands of times I've made long speeches about excellence to

large audiences, and did it very well, or so I fancied: but now there isn't a single thing that I can say. It seems to me that you would be well advised not to leave Athens or go to live elsewhere: if you behaved like this as a stranger in some foreign city, you'd probably be arrested as a sorcerer.

SOCRATES. You wretch, Meno, you nearly fooled me.

M. How so, Socrates?

S. I know why you used a simile about me.

M. Why, do you think?

S. So that I would return the compliment. I've discovered that all those who are beautiful enjoy having comparisons made about them—because they gain by it: for likenesses of beautiful things are themselves beautiful—but I'm not going to give you one in exchange for yours. I am like the electric ray, if it numbs others by being numb itself: but if it doesn't then I'm not. It's not that I am well off myself when I puzzle others, but rather I'm so puzzled that I make others puzzled too. And now, as regards excellence, I don't know what it is, and though you perhaps used to know before you came into contact with me, still, now it's just as if you didn't. However, I am ready to join you in investigating and trying to look for what it is.

M. And how will you look for this, Socrates, when you don't know at all what it is? For, which of the things that you don't know are you going to take this to be when you go looking for it? Or if you should happen to come across it, how will you know that this is what you didn't know?

S. I understand what you mean, Meno. Do you see that you are introducing that eristic [controversial] argument that a man cannot try to find out either what he knows or what he doesn't know? He wouldn't seek what he knows—for as he knows it, he has no need to make the inquiry—nor for what he doesn't know—for then he doesn't even know what he is to look for.

M. Don't you think that argument is a sound one, Socrates?

S. I do not.

M. Can you say why?

S. I can: because I have heard from men and women who are wise in matters of religion. . .

M. What account did they give?

S. One that was true, I thought, and beautiful.

M. What was this account and who were the people who told it to you?

S. They are holy men and women who are interested in being able to give an account of the activities in which they engage: Pindar talks of it too, and so do other divinely-inspired poets. What they say is this—just see if they seem to you to be speaking the truth. Because they say that man's soul is immortal, and at one time comes to an end—which they call dying—but at another is born, and is never totally annihilated. Because of this, one must live out one's life in as righteous a manner as one can. [...] Thus, since the soul is immortal and has come to be many times, and has seen both what is on earth and what is in Hades, all things, in fact, there is nothing which it has not learned: so it is not surprising that it is able to remember what it formerly knew, both about excellence and other things. Since all nature is related, and the soul is in a state of having learned everything, once someone has remembered one thing—this is what people call learning—nothing prevents him from finding out everything, provided he is stouthearted and does not tire of his inquiry: for all seeking and learning is remembering. So we mustn't accept that eristic argument, because it would make us lazy, and is one that pleases weak men. The other makes people hardworking and eager inquirers: I trust in its truth and am willing to join you in trying to find out what excellence is.

M. Yes, Socrates. But what do you mean when you say that we don't learn, but that what we call learning is remembering? Can you teach me how this is so?

S. Didn't I call you a wretch just now, Meno, and here you are, when I say that there is no teaching, only remembering, asking me if I can teach you,

so that I shall immediately be caught in an obvious contradiction.

M. No, no, Socrates, I didn't have that in mind, I was using the word out of habit. But if you have any way of showing me that things are as you say, please do show me.

S. Well, it's not easy, but for your sake, I'm willing to make the effort. Call over for me one of this crowd of servants you have here, any you like, so that I can use him to give you a demonstration.

M. Of course. [TO BOY] Come here.

S. He is a Greek and Greek-speaking?

M. Certainly, born in our house.

S. Pay attention then, and see whether he seems to you to be remembering or learning from me.

M. I'll do that.

S. Tell me, boy, do you recognize a square as a figure like this? [SOCRATES DRAWS A SQUARE]

BOY. I do.

S. So is a square a figure having all these lines equal, four of them? [INDICATING THE SIDES]

B. Yes.

S. Isn't it one having also these lines going through the middle equal? [IT IS LEFT TO THE READER TO DECIDE WHICH LINES THESE ARE]

B. Yes.

S. Such a figure could be larger or smaller?

B. Yes.

S. So if this side were two feet long, and this side two feet, how many feet would the whole be? Think of it like this: if it was two feet on this side, but only one foot on this one, the figure would be two feet by one, wouldn't it?

B. Yes.

S. But since it is two feet on this side, can it come to anything but twice two?

B. It comes to that.

S. So it comes to twice two feet?

B. Yes.

S. How many feet are twice two? Work it out and tell me.

B. Four, Socrates.

S. There could be, couldn't there, another figure twice the size of this one, and of the same kind, having all these lines equal like this one?

B. Yes.

S. How many feet will it be?

B. Eight.

S. Well now, try and tell me how long each of its lines will be. This one has a line of two feet: What is the line of the double square?

B. Obviously, Socrates, it is double.

S. Do you see, Meno, that I'm not teaching him anything, but always asking questions? He now thinks that he knows what kind of line it is from which the eight-foot figure will come to be; don't you agree?

M. I do.

S. So does he know?

M. By no means.

S. He thinks it is from the double line?

M. Yes.

S. Just watch him remembering in order, as one should remember. [SOCRATES TURNS BACK TO THE BOY] Tell me: you say the double figure comes to be from the double line? What I mean is not a figure that is long in one direction and short in another, but one that is to be equal in every direction like this one, but an eight-foot one, double its size. Do you still think that it will come to be from the double line?

B. I do.

S. Won't a line double that one come to be if we add another similar line to it?

B. Of course.

S. It is from this line, you say, that the eight-foot figure will be constructed, if there were four similar lines?

B. Yes.

S. Let's draw in the four equal lines. This will be what you say the eight-foot figure is, won't it?

B. Of course.

S. Surely there are in this figure these four figures, each of which is equal to the original four-foot square?

B. Yes.

S. So how big is it? Isn't it four times the size?

B. It must be.

S. So is what is four times the size double?

B. Certainly not.

S. But how many times as big?

B. Four times.

S. Then from the double line, boy, it's not a double figure that is generated, but one four times as big?

B. That's true.

S. Because four by four is sixteen, isn't it?

B. Yes.

S. What line gives the eight-foot one? Doesn't this line give a figure four times the size?

B. I agree.

S. And from the one half its length, the four-foot figure?

B. Yes.

S. Well; isn't the eight-foot figure double this one and half that?

B. Yes.

S. Won't it be generated from a line greater than one of this length, but less than one of that?

B. I think so.

S. Good. Do answer what you think. Tell me, wasn't this line two feet long, and this one four?

B. Yes.

S. So the line for the eight-foot figure must be longer than this two-foot one, but shorter than the four-foot one.

B. It must.

S. Now try and tell me how long you think it is.

B. Three feet.

S. If it's to be three feet, shall we increase this line by half and will that be three feet? Because this length is two feet and this is one: and from here again, this is two feet and this is one: and here is the figure which you mention.

B. Yes.

S. Surely if it's three feet this way and three that, the whole figure will come to be three by three feet?

B. It looks like it.

S. How many is three by three feet?

B. Nine.

S. How many is the double figure to be?

B. Eight.

S. So the eight-foot figure doesn't come to be from the three-foot line either?

B. Obviously not.

S. Then from what line? Try to tell us exactly: or if you don't want to put a number to it, then indicate its length.

B. My goodness, Socrates, I just don't know.

S. Do you see, Meno, where he has now got to in his journey of remembering? To begin with he didn't know what the line of the eight-foot figure is, just

as now he still doesn't know, but of course then he thought he knew it and answered confidently, like one who knows, and he didn't suppose he was at a loss. Now, though, he does suppose he's at a loss and he still doesn't know—but he doesn't think he knows either.

M. That's true.

S. Surely he's now better off as regards the thing which he didn't know?

M. I'd agree with that.

S. So by making him puzzled and numbing him like an electric ray, have we really done him any harm?

M. I don't think so.

S. In fact we've done him a service, as far as finding out how things are concerned: because now he will be glad to investigate, since he doesn't know, but previously he would have thought he could easily often have made fine speeches before large audiences about the double figure, saying that it must have a line double the length.

M. Presumably.

S. So do you think that he would then have tried to investigate or to learn what he didn't know but thought he knew, before he came to suppose he didn't know, fell into perplexity, and started wanting the knowledge?

M. I don't think so, Socrates.

S. So being numbed has helped him?

M. I think so.

S. Now see what, starting from this puzzlement, he'll discover, investigating with my help, when I do nothing but ask questions and don't do any teaching: watch out in case you catch me teaching or explaining to him, and not just asking for his opinions. [TURNING BACK TO THE BOY, SOCRATES BEGINS TO DRAW A NEW FIGURE.] Tell me, isn't this our four-foot figure? Do you understand?

B. I do.

S. We could add another one equal to it here?

B. Yes.

S. And a third equal to each of these?

B. Yes.

S. Then we could complete this space in the corner?

B. Certainly.

S. So these will be four equal figures?

B. Yes.

S. Well: how many times the size of this original figure is this whole one?

B. Four times.

S. But what we needed was one double the size, if you remember.

B. Of course.

S. Isn't this line from corner to corner one which cuts each of these figures in half?

B. Yes.

S. Don't these four lines turn out to be equal lines, enclosing this figure here?

B. They do.

S. Think then; what size is this figure?

B. I don't understand.

S. Doesn't each line cut off the inner half of each of these four figures?

B. Yes.

S. How many halves are there in this figure?

B. Four.

S. How many in this one?

B. Two.

S. What multiple is four of two?

B. Double.

S. So how many feet does this figure come to?

B. Eight feet.

S. From which line?

B. From this one.

S. From the one stretching from angle to angle of the four-foot figure?

B. Yes.

S. **Sophists** call this the “diagonal”: so if its name is the diagonal, it would be from the diagonal, in your opinion, Meno’s boy, that the double square would come to be.

B. Of course, Socrates.

S. What do you think, Meno? Has he expressed any opinion not his own when he answered?

M. No, all his own.

S. And yet he didn’t know, as we said a little while ago.

M. That’s true.

S. Still, these opinions were somehow in him, weren’t they?

M. Yes.

S. So in one who doesn’t know whatever it is he doesn’t know, there are true opinions about those things which he doesn’t know.

M. So it appears.

S. And just for now these opinions have been freshly aroused in him as if in a dream: but if anyone asks him these same things on many occasions and in many ways, you know very well that he will end up knowing about these things as accurately as anybody.

M. It seems so.

S. Surely he will know not from someone teaching him but from questioning, himself taking up the knowledge from himself?

M. Yes.

S. For oneself to take up knowledge that’s in oneself—isn’t that to remember?

M. Of course. [...]

Note

- Trans. Janet D. Sisson (2008). Sisson has an MA in Literae Humaniores and a BPhil in Philosophy from the University of Oxford and a PhD from the University of Calgary. A specialist in ancient philosophy, Sisson was a tenured lecturer at the

University of Glasgow until 1987. She has also taught at the Universities of Leicester, Oxford, Kansas, Alberta, Calgary, and most recently at Mount Royal University, Calgary.—LG

Study Questions

- Meno repeatedly goes wrong in searching for a definition of “excellence.” Explain how Socrates uses his examples of bees and shapes to set Meno straight.
- What definitions of “shape” and “colour” does Socrates propose in order to help Meno arrive at a definition of “excellence”? What might explain

Socrates’ preference for his definition of “shape” over his definition of “colour”?

- What is the pedagogical significance of accepting, with Socrates, that what we call learning is actually remembering?
- Why does Socrates reject the “eristic argument” that seems so convincing to Meno?

~ KNOWLEDGE OF THE EXTERNAL WORLD ~

René Descartes

3/30/15

René Descartes (1596–1650), who is considered to be the first modern philosopher, also made important contributions to mathematics and physics. Descartes was born in La Haye, France, and received a Jesuit education before attending the University of Poitiers, where he graduated in law. In *Meditations on First Philosophy*, Descartes methodically casts doubt on his most deeply held beliefs in order to provide new, secure foundations for scientific knowledge.—JF

From *Meditations on First Philosophy*

Meditation I: Of the Things on Which We May Doubt

Several years have now elapsed since I first became aware that I had accepted, even from my youth, many false opinions for true and that, consequently, what I afterward based on such principles was highly doubtful; and from that time I was convinced of the necessity of undertaking once in my life to rid myself of all the opinions I had adopted, and of commencing anew the work of building from the foundation, if I desired to establish a firm and abiding superstructure in the sciences. But as this enterprise appeared to me to be one of great magnitude, I waited until I had attained an age so mature as to leave me no hope that at any stage of life more advanced I should be better able to execute my design. On this account, I have delayed so long that I should henceforth consider I was doing wrong were I still to consume in deliberation any of the time that now remains for action. Today, then, since I have opportunely freed my mind from all cares and am happily disturbed by no passions, and since I am in the secure possession of leisure in a peaceable retirement, I will at length apply myself earnestly and freely to the general overthrow of all my former opinions. But, to this end, it will not be necessary for me to show that the whole of these are false—a point, perhaps, which I shall never reach; but as even now my reason convinces me that I ought not the less carefully to withhold belief from what is not entirely certain

and indubitable than from what is manifestly false, it will be sufficient to justify the rejection of the whole if I shall find in each some ground for doubt. Nor for this purpose will it be necessary even to deal with each belief individually, which would be truly an endless labour; but, as the removal from below of the foundation necessarily involves the downfall of the whole edifice, I will at once approach the criticism of the principles on which all my former beliefs rested.

All that I have, up to this moment, accepted as possessed of the highest truth and certainty, I received either from or through the senses. I observed, however, that these sometimes misled us; and it is the part of prudence not to place absolute confidence in that by which we have even once been deceived.

But it may be said, perhaps, that although the senses occasionally mislead us respecting minute objects, and such as are so far removed from us as to be beyond the reach of close observation, there are yet many other of their informations (presentations), of the truth of which it is manifestly impossible to doubt: as for example, that I am in this place, seated by the fire, clothed in a winter dressing gown, that I hold in my hands this piece of paper, with other intimations of the same nature. But how could I deny that I possess these hands and this body, and withal escape being classed with persons in a state of insanity, whose brains are so disordered and clouded by dark bilious vapours as to cause them pertinaciously to assert that they are monarchs when they are in the greatest poverty; or clothed in gold and purple when destitute of

any covering; or that their head is made of clay; their body of glass; or that they are gourds? I should certainly be not less insane than they, were I to regulate my procedure according to examples so extravagant.

Though this be true, I must nevertheless here consider that I am a man, and that, consequently, I am in the habit of sleeping, and representing to myself in dreams those same things, or even sometimes others less probable, which the insane think are presented to them in their waking moments. How often have I dreamt that I was in these familiar circumstances, that I was dressed and occupied this place by the fire, when I was lying undressed in bed? At the present moment, however, I certainly look upon this paper with eyes wide awake; the head which I now move is not asleep; I extend this hand consciously and with express purpose, and I perceive it; the occurrences in sleep are not so distinct as all this. But I cannot forget that, at other times, I have been deceived in sleep by similar illusions; and, attentively considering those cases, I perceive so clearly that there exist no certain marks by which the state of waking can ever be distinguished from sleep, that I feel greatly astonished; and in amazement I almost persuade myself that I am now dreaming.

Let us suppose, then, that we are dreaming, and that all these particulars—namely, the opening of the eyes, the motion of the head, the forth-putting of the hands—are merely illusions, and even that we really possess neither an entire body nor hands such as we see. Nevertheless it must be admitted at least that the objects which appear to us in sleep are, as it were, painted representations, which could not have been formed unless in the likeness of realities, and, therefore, that those general objects, at all events, namely, eyes, a head, hands, and an entire body are not simply imaginary, but really existent. For, in truth, painters themselves, even when they study to represent sirens and satyrs by forms the most fantastic and extraordinary, cannot bestow upon them natures absolutely new, but can only make a certain medley of the members of different animals; or if they chance to imagine something so novel that nothing at all similar has ever been seen before, and such as is, therefore, purely fictitious and absolutely

false, it is at least certain that the colours of which this is composed are real.

And on the same principle, although these general objects, viz., a body, eyes, a head, hands, and the like, be imaginary, we are nevertheless absolutely necessitated to admit the reality at least of some other objects still more simple and universal than these, of which, just as of certain real colours, all those images of things, whether true and real, or false and fantastic, that are found in our consciousness (*cogitatio*) are formed.

To this class of objects seem to belong corporeal nature in general and its extension: the figure of extended things, their quantity or magnitude, and their number, as also the place in, and the time during, which they exist, and other things of the same sort. We will not, therefore, perhaps reason illegitimately if we conclude from this that Physics, Astronomy, Medicine, and all the other sciences that have for their end the consideration of composite objects are indeed of a doubtful character, but that Arithmetic, Geometry, and the other sciences of the same class, which regard merely the simplest and most general objects, and scarcely inquire whether or not these are really existent, contain somewhat that is certain and indubitable: for whether I am awake or dreaming, it remains true that two and three make five, and that a square has four sides; nor does it seem possible that truths so apparent can ever fall under a suspicion of falsity or incertitude.

Nevertheless, the belief that there is a God who is all powerful, and who created me, such as I am, has, for a long time, obtained steady possession of my mind. How, then, do I know that he has not arranged that there should be neither earth, nor sky, nor any extended thing, nor figure, nor magnitude, nor place, providing at the same time, however, for the rise in me of the perceptions of all these objects, and the persuasion that these do not exist otherwise than as I perceive them? And further, as I sometimes think that others are in error respecting matters of which they believe themselves to possess a perfect knowledge, how do I know that I am not also deceived each time I add together two and

three, or number the sides of a square, or form some judgment still more simple, if more simple indeed can be imagined? But perhaps Deity has not been willing that I should be thus deceived, for he is said to be supremely good. If, however, it were repugnant to the goodness of Deity to have created me subject to constant deception, it would seem likewise to be contrary to his goodness to allow me to be occasionally deceived; and yet it is clear that this is permitted. Some, indeed, might perhaps be found who would be disposed rather to deny the existence of a Being so powerful than to believe that there is nothing certain. But let us for the present refrain from opposing this opinion, and grant that all which is here said of a Deity is fabulous; nevertheless, in whatever way it be supposed that I reached the state in which I exist, whether by fate, or chance, or by an endless series of antecedents and consequents, or by any other means, it is clear (since to be deceived and to err is a certain defect) that the probability of my being so imperfect as to be the constant victim of deception will be increased exactly in proportion as the power possessed by the cause to which they assign my origin is lessened. To these reasonings I have assuredly nothing to reply, but am constrained at last to avow that there is nothing of all that I formerly believed to be true of which it is impossible to doubt, and that not through thoughtlessness or levity, but from cogent and maturely considered reasons; so that henceforward, if I desire to discover anything certain, I ought not the less carefully to refrain from assenting to those same opinions than to what might be shown to be manifestly false.

But it is not sufficient to have made these observations; care must be taken likewise to keep them in remembrance. For those old and customary opinions perpetually recur—long and familiar usage giving them the right of occupying my mind, even almost against my will, and subduing my belief; nor will I lose the habit of deferring to them and confiding in them so long as I shall consider them to be whatin truth they are, viz., opinions to some extent doubtful, as I have already shown, but still highly probable, and such as it is much more reasonable to believe than deny. It is for this reason

I am persuaded that I shall not be doing wrong, if, taking an opposite judgment of deliberate design, I become my own deceiver, by supposing, for a time, that all those opinions are entirely false, and imaginary, until at length, having thus balanced my old by my new prejudices, my judgment shall no longer be turned aside by perversed usage from the path that may conduct to the perception of truth. For I am assured that, meanwhile, there will arise neither peril nor error from this course, and that I cannot for the present yield too much to distrust, since the end I now seek is not action but knowledge.

I will suppose, then, not that Deity, who is sovereignly good and the fountain of truth, but that some malignant demon, who is at once exceedingly potent and deceitful, has employed all his artifice to deceive me; I will suppose that the sky, the air, the earth, colours, figures, sounds, and all external things are nothing better than the illusions of dreams, by means of which this being has laid snares for my credulity; I will consider myself as without hands, eyes, flesh, blood, or any of the senses, and as falsely believing that I am possessed of these; I will continue resolutely fixed in this belief, and if indeed by this means it be not in my power to arrive at the knowledge of truth, I shall at least do what is in my power, viz., suspend my judgment, and guard with settled purpose against giving my assent to what is false and being imposed upon by this deceiver, whatever be his power and artifice.

But this undertaking is arduous, and a certain indolence insensibly leads me back to my ordinary course of life; and just as the captive, who, per chance, was enjoying in his dreams an imaginary liberty when he begins to suspect that it is but a vision, dreads awakening and conspires with the agreeable illusions that the deception may be prolonged, so I, of my own accord, fall back into the train of my former beliefs and fear to arouse myself from my slumber, lest the time of laborious wakefulness that would succeed this quiet rest, in place of bringing any light of day, should prove inadequate to dispel the darkness that will arise from the difficulties that have now been raised.

Meditation II: Of the Nature of the Human Mind; and That It Is More Easily Known Than the Body

The Meditation of yesterday has filled my mind with so many doubts that it is no longer in my power to forget them. Nor do I see, meanwhile, any principle on which they can be resolved; and, just as if I had fallen all of a sudden into very deep water, I am so greatly disconcerted as to be unable either to plant my feet firmly on the bottom or sustain myself by swimming on the surface. I will, nevertheless, make an effort, and try anew the same path on which I had entered yesterday, that is, proceed by casting aside all that admits of the slightest doubt, not less than if I had discovered it to be absolutely false; and I will continue always in this track until I shall find something that is certain, or at least, if I can do nothing more, until I shall know with certainty that there is nothing certain. **Archimedes**, that he might transport the entire globe from the place it occupied to another, demanded only a point that was firm and immovable; so, also, I shall be entitled to entertain the highest expectations, if I am fortunate enough to discover only one thing that is certain and indubitable.

I suppose, accordingly, that all the things which I see are false (fictitious); I believe that none of those objects which my fallacious memory represents ever existed; I suppose that I possess no senses; I believe that body, figure, extension, motion, and place are merely fictions of my mind. What is there, then, that can be esteemed true? Perhaps this only, that there is absolutely nothing certain.

But how do I know that there is not something different altogether from the objects I have now enumerated, of which it is impossible to entertain the slightest doubt? Is there not a God, or some being, by whatever name I may designate him, who causes these thoughts to arise in my mind? But why suppose such a being, for it may be I myself am capable of producing them? Am I, then, at least not something? But I before denied that I possessed senses or a body; I hesitate, however, for what follows from that? Am I so dependent on the body and the senses

that without these I cannot exist? But I had the persuasion that there was absolutely nothing in the world, that there was no sky and no earth, neither minds nor bodies; was I not, therefore, at the same time, persuaded that I did not exist? Far from it; I assuredly existed, since I was persuaded. But there is I know not what being, who is possessed at once of the highest power and the deepest cunning, who is constantly employing all his ingenuity in deceiving me. Doubtless, then, I exist, since I am deceived; and, let him deceive me as he may, he can never bring it about that I am nothing, so long as I shall be conscious that I am something. So that it must, in fine, be maintained, all things being maturely and carefully considered, that this proposition (*pronunciatum*) I am, I exist, is necessarily true each time it is expressed by me, or conceived in my mind.

But I do not yet know with sufficient clearness what I am, though assured that I am; and hence, in the next place, I must take care, lest perchance I inconsiderately substitute some other object in room of what is properly myself, and thus wander from truth, even in that knowledge (cognition) which I hold to be of all others the most certain and evident. For this reason, I will now consider anew what I formerly believed myself to be, before I entered on the present train of thought; and of my previous opinion I will retrench all that can in the least be invalidated by the grounds of doubt I have adduced, in order that there may at length remain nothing but what is certain and indubitable. What then did I formerly think I was? Undoubtedly I judged that I was a man. But what is a man? Shall I say a rational animal? Assuredly not; for it would be necessary forthwith to inquire into what is meant by animal, and what by rational, and thus, from a single question, I should insensibly glide into others, and these more difficult than the first; nor do I now possess enough of leisure to warrant me in wasting my time amid subtleties of this sort. I prefer here to attend to the thoughts that sprung up of themselves in my mind, and were inspired by my own nature alone, when I applied myself to the consideration of what I was. In the first place, then, I thought that I possessed a countenance, hands, arms, and

all the fabric of members that appears in a corpse, and which I called by the name of body. It further occurred to me that I was nourished, that I walked, perceived, and thought, and all those actions I referred to the soul; but what the soul itself was I either did not stay to consider, or, if I did, I imagined that it was something extremely rare and subtle, like wind, or flame, or ether, spread through my grosser parts. As regarded the body, I did not even doubt of its nature, but thought I distinctly knew it, and if I had wished to describe it according to the notions I then entertained, I should have explained myself in this manner: By body I understand all that can be terminated by a certain figure; that can be comprised in a certain place, and so fill a certain space as therefrom to exclude every other body; that can be perceived either by touch, sight, hearing, taste, or smell; that can be moved in different ways, not indeed of itself, but by something foreign to it by which it is touched and from which it receives the impression; for the power of self-motion, as likewise that of perceiving and thinking, I held as by no means pertaining to the nature of body; on the contrary, I was somewhat astonished to find such faculties existing in some bodies.

But as to myself, what can I now say that I am, since I suppose there exists an extremely powerful, and, if I may so speak, malignant being, whose whole endeavours are directed toward deceiving me? Can I affirm that I possess any one of all those attributes of which I have lately spoken as belonging to the nature of body? After attentively considering them in my own mind, I find none of them that can properly be said to belong to myself. To recount them were idle and tedious. Let us pass, then, to the attributes of the soul. The first mentioned were the powers of nutrition and walking; but, if it be true that I have no body, it is true likewise that I am capable neither of walking nor of being nourished. Perception is another attribute of the soul; but perception too is impossible without the body; besides, I have frequently, during sleep, believed that I perceived objects which I afterward observed I did not in reality perceive. Thinking is another attribute of the soul; and here I discover what properly

belongs to myself. This alone is inseparable from me. I am—I exist: this is certain; but how often? As often as I think; for perhaps it would even happen, if I should wholly cease to think, that I should at the same time altogether cease to be. I now admit nothing that is not necessarily true. I am therefore, precisely speaking, only a thinking thing, that is, a mind (*mens sive animus*), understanding, or reason, terms whose signification was before unknown to me. I am, however, a real thing, and really existent; but what thing? The answer was, a thinking thing. The question now arises, am I aught besides? I will stimulate my imagination with a view to discover whether I am not still something more than a thinking being. Now it is plain I am not the assemblage of members called the human body; I am not a thin and penetrating air diffused through all these members, or wind, or flame, or vapour, or breath, or any of all the things I can imagine; for I supposed that all these were not, and, without changing the supposition, I find that I still feel assured of my existence.

But it is true, perhaps, that those very things which I suppose to be nonexistent because they are unknown to me are not in truth different from myself whom I know. This is a point I cannot determine, and do not now enter into any dispute regarding it. I can only judge of things that are known to me: I am conscious that I exist, and I who know that I exist inquire into what I am. It is, however, perfectly certain that the knowledge of my existence, thus precisely taken, is not dependent on things, the existence of which is as yet unknown to me: and consequently it is not dependent on any of the things I can feign in imagination. Moreover, the phrase itself, I frame an image (*effingo*), reminds me of my error; for I should in truth frame one if I were to imagine myself to be anything, since to imagine is nothing more than to contemplate the figure or image of a corporeal thing; but I already know that I exist, and that it is possible at the same time that all those images, and in general all that relates to the nature of body, are merely dreams or chimeras. From this I discover that it is not more reasonable to say, I will excite my imagination that I may know more distinctly what I am, than to express myself as

follows: I am now awake, and perceive something real, but because my perception is not sufficiently clear, I will of express purpose go to sleep that my dreams may represent to me the object of my perception with more truth and clearness. And, therefore, I know that nothing of all that I can embrace in imagination belongs to the knowledge which I have of myself, and that there is need to recall with the utmost care the mind from this mode of thinking, that it may be able to know its own nature with perfect distinctness.

But what, then, am I? A thinking thing, it has been said. But what is a thinking thing? It is a thing that doubts, understands, conceives, affirms, denies, wills, refuses; that imagines also, and perceives. Assuredly it is not little, if all these properties belong to my nature. But why should they not belong to it? Am I not that very being who now doubts of almost everything; who, for all that, understands and conceives certain things; who affirms one alone as true, and denies the others; who desires to know more of them, and does not wish to be deceived; who imagines many things, sometimes even despite his will; and is likewise peripient of many, as if through the medium of the senses. Is there nothing of all this as true as that I am, even although I should be always dreaming, and although he who gave me being employed all his ingenuity to deceive me? Is there also any one of these attributes that can be properly distinguished from my thought, or that can be said to be separate from myself? For it is of itself so evident that it is I who doubt, I who understand, and I who desire, that it is here unnecessary to add anything by way of rendering it more clear. And I am as certainly the same being who imagines; for although it may be (as I before supposed) that nothing I imagine is true, still the power of imagination does not cease really to exist in me and to form part of my thought. In fine, I am the same being who perceives, that is, who apprehends certain objects as by the organs of sense, since, in truth, I see light, hear a noise, and feel heat. But it will be said that these presentations are false, and that I am dreaming. Let it be so. At all events it is certain that I seem to see light, hear a noise, and feel heat; this cannot be false, and this is what in me is

properly called perceiving (*sentire*), which is nothing else than thinking. From this I begin to know what I am with somewhat greater clearness and distinctness than heretofore.

But, nevertheless, it still seems to me, and I cannot help believing, that corporeal things, whose images are formed by thought which fall under the senses, and are examined by the same, are known with much greater distinctness than that I know not what part of myself which is not imaginable; although, in truth, it may seem strange to say that I know and comprehend with greater distinctness things whose existence appears to me doubtful, than that are unknown, and do not belong to me, than others of whose reality I am persuaded, that are known to me, and appertain to my proper nature; in a word, than myself. But I see clearly what is in the state of the case. My mind is apt to wander, and will not yet submit to be restrained within the limits of truth. Let us therefore leave the mind to itself once more, and, according to it every kind of liberty permit it to consider the objects that appear to it from without, in order that, having afterward withdrawn it from these gently and opportunely and fixed it on the consideration of its being and the properties it finds in itself, it may then be the more easily controlled.

Let us now accordingly consider the objects that are commonly thought to be the most easily, and likewise the most distinctly known, viz., the bodies we touch and see—not, indeed, bodies in general, for these general notions are usually somewhat more confused, but one body in particular. Take, for example, this piece of wax: it is quite fresh, having been but recently taken from the beehive; it has not yet lost the sweetness of the honey it contained, it still retains somewhat of the odour of the flowers from which it was gathered; its colour, figure, size are apparent (to the sight); it is hard, cold, easily handled; and sounds when struck upon with the finger. In fine, all that contributes to make a body as distinctly known as possible is found in the one before us. But, while I am speaking, let it be placed near the fire—what remained of the taste exhales, the smell evaporates, the colour changes, its figure

is destroyed, its size increases, it becomes liquid, it struck upon, it emits no sound, and, although still remain after this change? Does the same wax that it does remain; no one doubts it? It must be admitted tinctness in the piece of wax? Assuredly, it could be nothing of all that I observed by means of the senses, since all the things that fell under taste, smell, sight, touch, and hearing are changed, and yet the same wax remains. It was perhaps what I now think, viz., pleasant odour of flowers, the whiteness of honey, the nor the sound, but only a body that a little before appeared to me conspicuous under these forms, and which is now perceived under others. But, to speak precisely, what is it that I imagine when I think of it in this way? Let it be attentively considered, and, retrenching all that does not belong to the wax, let us see what remains. There certainly remains nothing, except something extended, flexible, and movable. But what is meant by flexible and movable? Is it not that I imagine that the piece of wax, being round, is capable of becoming square, or of passing from a square into a triangular figure? Assuredly such is not the case, because I conceive that it admits of an infinity of similar changes; and I am, moreover, unable to compass this infinity by imagination; and consequently this conception which I have of the wax is not the product of the faculty of imagination. But what now is this extension? Is it not also unknown? For it becomes greater when the wax is melted, greater when it is boiled, and greater still when the heat increases; and I should not conceive clearly and according to truth the wax as it is, if I did not suppose that the piece we are considering admitted even of a wider variety of extension than I ever imagined. I must, therefore, admit that I cannot even comprehend by imagination what the piece of wax is, and that it is the mind alone (*mens*, Lat., *entendement*, F.) which perceives it. I speak of one piece in particular; for as to wax in general, this is still more evident. But what is the piece of wax that can be perceived only by the understanding or mind? It is certainly the same which I see, touch,

composed.

But, meanwhile, I feel greatly astonished when I observe the weakness of my mind, and its proneness to error. For although, without at all giving expression to what I think, I consider all this in my own mind, words yet occasionally impede my progress, and I am almost led into error by the terms of ordinary language. We say, for example, that we see the same wax when it is before us, and not that we judge it to be the same from its retaining the same colour and figure; whence I should forthwith be disposed to conclude that the wax is known by the act of sight, and not by the intuition of the mind alone, were it not for the analogous instance of human beings passing on in the street below, as observed from a window. In this case I do not fail to say that I see the men themselves, just as I say that I see the wax; and yet what do I see from the window beyond hats and cloaks that might cover artificial machines, whose motions might be determined by springs? But I judge that there are human beings from these appearances, and thus I comprehend, by the faculty of judgment alone which is in the mind, what I believed I saw with my eyes.

The man who makes it his aim to rise to knowledge superior to the common ought to be ashamed to seek occasions of doubting from the vulgar forms of speech; instead, therefore, of doing this, I shall proceed with the matter in hand, and inquire whether I had a clearer and more perfect perception of the piece of wax when I first saw it, and when I thought I knew it by means of the external sense itself, or, at all events, by the common sense (*sensus communis*), as it is called, that is, by the imaginative faculty; or whether I rather apprehend it more clearly

at present, after having examined with greater care, both what it is, and in what way it can be known. It would certainly be ridiculous to entertain any doubt on this point. For what in that first perception, was there distinct? What did I perceive which any animal might not have perceived? But when I distinguish the wax from its exterior forms, and when, as if I had stripped it of its vestments, I consider it quite naked, it is certain, although some error may still be found in my judgment, that I cannot, nevertheless, thus apprehend it without possessing a human mind.

But, finally, what shall I say of the mind itself, that is, of myself? For as yet I do not admit that I am anything but mind. What, then! who seem to possess so distinct an apprehension of the piece of wax, do I not know myself, both with greater truth and certitude, and also much more distinctly and clearly? For if I judge that the wax exists because I see it, it assuredly follows, much more evidently, that I myself am or exist, for the same reason: for it is possible that what I see may not in truth be wax, and that I do not even possess eyes with which to see anything; but it cannot be that when I see, or, which comes to the same thing, when I think I see, I myself who think am nothing. So likewise, if I judge that the wax exists because I touch it, it will still also follow that I am; and if I determine that my imagination, or any other cause, whatever it be, persuades me of the existence of the wax, I will still draw the

same conclusion. And what is here remarked of the piece of wax is applicable to all the other things that are external to me. And further, if the notion or perception of wax appeared to me more precise and distinct, after not only sight and touch but many other causes besides rendered it manifest to my apprehension, with how much greater distinctness must I know myself, since all the reasons that contribute to the knowledge of the nature of wax, or of any body whatever, manifest still better the nature of my mind? And there are, besides, so many other things in the mind itself that contribute to the illustration of its nature that those dependent on the body, to which I have here referred, scarcely merit to be taken into account.

But, in conclusion, I find I have insensibly reverted to the point I desired: for, since it is now manifest to me that bodies themselves are not properly perceived by the senses nor by the faculty of imagination, but by the intellect alone; and since they are not perceived because they are seen and touched, but only because they are understood or rightly comprehended by thought, I readily discover that there is nothing more easily or clearly apprehended than my own mind. But because it is difficult to rid one's self so promptly of an opinion to which one has been long accustomed, it will be desirable to tarry for some time at this stage; that, by long continued meditation, I may more deeply impress upon my memory this new knowledge.

Study Questions

1. Detail Descartes' employment of methodological scepticism in the *Meditations*; be sure to consider each of the devices he uses and the specific role it plays for him.
2. What is the subject matter of the sciences of arithmetic and geometry?
3. In the second Meditation, Descartes determines the certainty of these two propositions: (i) "I am, I exist"; and (ii) "I am . . . in the strict sense

only a thing that thinks, that is, I am a mind, or intelligence, or intellect, or reason." How do these two propositions differ in their philosophical significance?

4. According to Descartes, what is the essence of material things? On what bases does Descartes argue that we arrive at this knowledge through reason, not through our senses or imagination?

George Berkeley

4/6/15

George Berkeley (1685–1753) was born near Kilkenny, Ireland, and completed his education at Trinity College, Dublin. Berkeley's best-known works, including *Three Dialogues*, arose in critical response to the writings of Descartes, Locke, and other modern philosophers while he was a fellow at Trinity College between 1707 and 1724. An ordained Anglican priest, Berkeley was appointed dean of Derry in 1724 and bishop of Cloyne in 1734. Between 1728 and 1731, Berkeley lived in Rhode Island, waiting in vain for a promised parliamentary grant in support of his planned seminary in Bermuda.—JF

From *Three Dialogues between Hylas and Philonous, in Opposition to Skeptics and Atheists*

The First Dialogue

PHILONOUS. Good morrow, Hylas: I did not expect to find you abroad so early.

HYLAS. It is indeed something unusual: but my thoughts were so taken up with a subject I was discoursing of last night that finding I could not sleep, I resolved to rise and take a turn in the garden.

P. It happened well, to let you see what innocent and agreeable pleasures you lose every morning. Can there be a pleasanter time of the day, or a more delightful season of the year? That purple sky, those wild but sweet notes of birds, the fragrant bloom upon the trees and flowers, the gentle influence of the rising sun, these and a thousand nameless beauties of nature inspire the soul with secret transports; its faculties too, being at this time fresh and lively, are fit for these meditations, which the solitude of a garden and tranquillity of the morning naturally dispose us to. But I am afraid I interrupt your thoughts; for you seemed very intent on something.

H. It is true, I was, and shall be obliged to you if you will permit me to go on in the same vein; not that I would by any means deprive myself of your company, for my thoughts always flow more easily in conversation with a friend than when I am alone: but my request is, that you would suffer me to impart my reflections to you.

P. With all my heart, it is what I should have requested myself, if you had not prevented me.

H. I was considering the odd fate of those men who have in all ages, through an affectation of being distinguished from the vulgar, or some unaccountable turn of thought, pretended either to believe nothing at all, or believe the most extravagant things in the world. This however might be borne, if their paradoxes and skepticism did not draw after them some consequences of general disadvantage to mankind. But the mischief lieth here: that when men of less leisure see them who are supposed to have spent their whole time in the pursuits of knowledge professing an entire ignorance of all things, or advancing such notions as are repugnant to plain and commonly received principles, they will be tempted to entertain suspicions concerning the most important truths, which they had hitherto held sacred and unquestionable.

P. I entirely agree with you, as to the ill tendency of the affected doubts of some philosophers, and fantastical conceits of others. I am even so far gone of late in this way of thinking that I have quitted several of the sublime notions I had got in their schools for vulgar opinions. And I give it you on my word, since this revolt from metaphysical notions to the plain dictates of nature and common sense, I find my understanding strangely enlightened, so that I can now easily comprehend a great many things which before were all mystery and riddle.

H. I am glad to find there was nothing in the accounts I heard of you.

P. Pray, what were those?

H. You were represented in last night's conversation as one who maintained the most extravagant opinion that ever entered into the mind of man, to wit, that there is no such thing as *material substance* in the world.

P. That there is no such thing as what philosophers call material substance, I am seriously persuaded; but if I were made to see anything absurd or skeptical in this, I should then have the same reason to renounce this that I imagine I have now to reject the contrary opinion.

H. What! Can anything be more fantastical, more repugnant to common sense, or a more manifest piece of skepticism than to believe there is no such thing as matter?

P. Softly, good Hylas. What if it should prove that you, who hold there is, are by virtue of that opinion a greater skeptic, and maintain more paradoxes and repugnancies to common sense, than I who believe no such thing?

H. You may as soon persuade me, the part is greater than the whole as that, in order to avoid absurdity and skepticism, I should ever be obliged to give up my opinion in this point.

P. Well then, are you content to admit that opinion for true, which upon examination shall appear most agreeable to common sense and remote from skepticism?

H. With all my heart. Since you are for raising disputes about the plainest things in nature, I am content for once to hear what you have to say.

P. Pray, Hylas, what do you mean by a skeptic?

H. I mean what all men mean, one that doubts of everything.

P. He then who entertains no doubts concerning some particular point, with regard to that point cannot be thought a *skeptic*.

H. I agree with you.

P. Whether doth doubting consist in embracing the affirmative or negative side of a question?

H. In neither; for whoever understands English cannot but know that *doubting* signifies a suspense between both.

P. He then that denieth any point can no more be said to doubt of it than he who affirmeth it with the same degree of assurance.

H. True.

P. And consequently, for such his denial is no more to be esteemed a *skeptic* than the other.

H. I acknowledge it.

P. How cometh it to pass then, Hylas, that you pronounce me a *skeptic* because I deny what you affirm, to wit, the existence of matter? Since, for aught you can tell, I am as peremptory in my denial, as you in your affirmation.

H. Hold, Philonous, I have been a little out in my definition; but every false step a man makes in discourse is not to be insisted on. I said, indeed, that a *skeptic was one who doubted of everything*; but I should have added, or who denies the reality and truth of things.

P. What things? Do you mean the principles and theorems of sciences? But these you know are universal intellectual notions, and consequently independent of matter; the denial therefore of this doth not imply the denying them.

H. I grant it. But are there no other things? What think you of distrusting the senses, of denying the real existence of sensible things, or pretending to know nothing of them? Is not this sufficient to denominate a man a *skeptic*?

P. Shall we therefore examine which of us it is that denies the reality of sensible things, or professes the greatest ignorance of them, since, if I take you rightly, he is to be esteemed the *greatest skeptic*?

H. That is what I desire.

P. What mean you by sensible things?

H. Those things which are perceived by the senses. Can you imagine that I mean anything else?

P. Pardon me, Hylas, if I am desirous clearly to apprehend your notions, since this may much shorten our inquiry. Suffer me then to ask you this further question. Are those things only perceived by the senses which are perceived immediately? Or may those things properly be said to be sensible which are perceived mediately, or not without the intervention of others?

H. I do not sufficiently understand you.

P. In reading a book, what I immediately perceive are the letters, but mediately, or by means of these, are suggested to my mind the notions of God, virtue, truth, etc. Now that the letters are truly sensible things, or perceived by sense, there is no doubt: but I would know whether you take the things suggested by them to be so too.

H. No, certainly, it were absurd to think *God or virtue* sensible things, though they may be signified and suggested to the mind by sensible marks, with which they have an arbitrary connection.

P. It seems, then, that by *sensible things* you mean those only which can be perceived immediately by sense.

H. Right.

P. Doth it not follow from this that though I see one part of the sky red, and another blue, and that my reason doth thence evidently conclude there must be some cause of that diversity of colours, yet that cause cannot be said to be a sensible thing, or perceived by the sense of seeing?

H. It doth.

P. In like manner, though I hear variety of sounds, yet I cannot be said to hear the causes of those sounds.

H. You cannot.

P. And when by my touch I perceive a thing to be hot and heavy, I cannot say with any truth or propriety that I feel the cause of its heat or weight?

H. To prevent any more questions of this kind, I tell you once for all that by *sensible things* I mean those only which are perceived by sense, and that in truth the senses perceive nothing which they do not

perceive immediately: for they make no inferences. The deducing therefore of causes or occasions from effects and appearances, which alone are perceived by sense, entirely relates to reason.

P. This point then is agreed between us, that sensible things are those only which are immediately perceived by sense. You will further inform me, whether we immediately perceive by sight anything beside light, and colours, and figures: or by hearing anything but sounds: by the palate, anything beside tastes: by the smell, besides odours: or by the touch, more than tangible qualities.

H. We do not.

P. It seems, therefore, that if you take away all sensible qualities, there remains nothing sensible.

H. I grant it.

P. Sensible things therefore are nothing else but so many sensible qualities, or combinations of sensible qualities.

H. Nothing else.

P. Heat then is a sensible thing.

H. Certainly.

P. Doth the reality of sensible things consist in being perceived? Or is it something distinct from their being perceived, and that bears no relation to the mind?

H. To exist is one thing, and to be perceived is another.

P. I speak with regard to sensible things only; and of these I ask, whether by their real existence you mean a subsistence exterior to the mind, and distinct from their being perceived?

H. I mean a real absolute being, distinct from, and without any relation to their being perceived.

P. Heat, therefore, if it be allowed a real being, must exist without the mind.

H. It must. [. . .]

P. Can any doctrine be true that necessarily leads a man into an absurdity?

H. Without doubt it cannot.

P. Is it not an absurdity to think that the same thing should be at the same time both cold and warm?

H. It is.

P. Suppose now one of your hands hot and the other cold, and that they are both at once put into the same vessel of water in an intermediate state; will not the water seem cold to one hand and warm to the other?

H. It will.

P. Ought we not therefore by your principles to conclude, it is really both cold and warm at the same time, that is, according to your own concession, to believe an absurdity?

H. I confess it seems so.

P. Consequently, the principles themselves are false, since you have granted that no true principle leads to an absurdity.

H. But after all, can anything be more absurd than to say, *there is no heat in the fire?*

P. To make the point still clearer; tell me, whether in two cases exactly alike, we ought not to make the same judgment?

H. We ought.

P. When a pin pricks your finger, doth it not rend and divide the fibres of your flesh?

H. It doth.

P. And when a coal burns your finger, doth it any more?

H. It doth not.

P. Since therefore you neither judge the sensation itself occasioned by the pin nor anything like it to be in the pin, you should not, conformably to what you have now granted, judge the sensation occasioned by the fire or anything like it to be in the fire.

H. Well, since it must be so, I am content to yield this point, and acknowledge that heat and cold are

~~only sensations existing in our minds; but there still remain qualities enough to secure the reality of external things.~~

P. But what will you say, Hylas, if it shall appear that the case is the same with regard to all other sensible qualities, and that they can no more be supposed to exist without the mind than heat and cold?

H. Then indeed you will have done something to the purpose; but that is what I despair of seeing proved.

P. Let us examine them in order. What think you of tastes, do they exist without the mind, or no?

H. Can any man in his senses doubt whether sugar is sweet, or wormwood bitter?

P. Inform me, Hylas. Is a sweet taste a particular kind of pleasure or pleasant sensation, or is it not?

H. It is.

P. And is not bitterness some kind of uneasiness or pain?

H. I grant it.

P. If therefore sugar and wormwood are unthinking corporeal substances existing without the mind, how can sweetness and bitterness, that is, pleasure and pain, agree to them?

H. Hold, Philonous; I now see what it was deluded me all this time. You asked whether heat and cold, sweetness and bitterness, were not particular sorts of pleasure and pain; to which I answered simply, that they were. Whereas I should have thus distinguished: those qualities, as perceived by us, are pleasures or pains, but not as existing in the external objects. We must not therefore conclude absolutely that there is no heat in the fire, or sweetness in the sugar, but only that heat or sweetness, as perceived by us, are not in the fire or sugar. What say you to this?

P. I say it is nothing to the purpose. Our discourse proceeded altogether concerning sensible things, which you defined to be the things we immediately perceive by our senses. Whatever other qualities

therefore you speak of, as distinct from these, I know nothing of them, neither do they at all belong to the point in dispute. You may indeed pretend to have discovered certain qualities which you do not perceive, and assert those insensible qualities exist in fire and sugar. But what use can be made of this to your present purpose, I am at a loss to conceive. Tell me then once more, do you acknowledge that heat and cold, sweetness and bitterness (meaning those qualities which are perceived by the senses), do not exist without the mind?

H. I see it is to no purpose to hold out, so I give up the cause as to those mentioned qualities. Though I profess it sounds oddly to say that sugar is not sweet.

P. But for your further satisfaction, take this along with you: that which at other times seems sweet shall to a distempered palate appear bitter. And nothing can be plainer than that diverse persons perceive different tastes in the same food, since that which one man delights in, another abhors. And how could this be, if the taste was something really inherent in the food?

H. I acknowledge I know not how.

P. In the next place, odours are to be considered. And with regard to these, I would fain know, whether what hath been said of tastes doth not exactly agree to them? Are they not so many pleasing or displeasing sensations?

H. They are.

P. Can you then conceive it possible that they should exist in an unperceiving thing?

H. I cannot.

P. Or can you imagine that filth and ordure affect those brute animals that feed on them out of choice with the same smells which we perceive in them?

H. By no means.

P. May we not therefore conclude of smells, as of the other forementioned qualities, that they cannot exist in any but a perceiving substance or mind?

H. I think so.

P. Then as to sounds, what must we think of them: are they accidents really inherent in external bodies, or not?

H. That they inhere not in the sonorous bodies is plain from hence: because a bell struck in the exhausted receiver of an air-pump sends forth no sound. The air therefore must be thought the subject of sound.

P. What reason is there for that, Hylas?

H. Because when any motion is raised in the air, we perceive a sound greater or lesser, according to the air's motion; but without some motion in the air, we never hear any sound at all.

P. And granting that we never hear a sound but when some motion is produced in the air, yet I do not see how you can infer from thence that the sound itself is in the air.

H. It is this very motion in the external air that produces in the mind the sensation of sound. For striking on the drum of the ear, it causeth a vibration, which by the auditory nerves being communicated to the brain, the soul is thereupon affected with the sensation called sound.

P. What! Is sound then a sensation?

H. I tell you, as perceived by us, it is a particular sensation in the mind.

P. And can any sensation exist without the mind?

H. No, certainly.

P. How then can sound, being a sensation, exist in the air, if by the air you mean a senseless substance existing without the mind.

H. You must distinguish, Philonous, between sound, as it is perceived by us, and as it is in itself, or, (which is the same thing) between the sound we immediately perceive, and that which exists without us. The former indeed is a particular kind of sensation, but the latter is merely a vibrative or undulatory motion in the air.

P. I thought I had already obviated that distinction by the answer I gave when you were applying it in a like case before. But to say no more of that: are you sure then that sound is really nothing but motion?

H. I am.

P. Whatever therefore agrees to real sound may with truth be attributed to motion.

H. It may.

P. It is then good sense to speak of *motion*, as of a thing that is *loud, sweet, acute, or grave*.

H. I see you are resolved not to understand me. Is it not evident, those accidents or modes belong only to sensible sound, or sound in the common acceptation of the word, but not to sound in the real and philosophic sense, which, as I just now told you, is nothing but a certain motion of the air?

P. It seems then there are two sorts of sound, the one vulgar, or that which is heard, the other philosophic and real.

H. Even so.

P. And the latter consists in motion.

H. I told you so before.

P. Tell me, Hylas, to which of the senses, think you, the idea of motion belongs: to the hearing?

H. No, certainly, but to the sight and touch.

P. It should follow, then, that according to you, real sounds may possibly be *seen or felt*, but never *heard*.

H. Look you, Philonous, you may if you please make a jest of my opinion, but that will not alter the truth of things. I own, indeed, the inferences you draw me into sound something oddly: but common language, you know, is framed by, and for, the use of the vulgar: we must not therefore wonder if expressions adapted to exact philosophic notions seem uncouth and out of the way.

P. Is it come to that? I assure you, I imagine myself to have gained no small point since you make so light of departing from common phrases and opinions, it

being a main part of our inquiry to examine whose notions are widest of the common road, and most repugnant to the general sense of the world. But can you think it no more than a philosophical paradox to say that real sounds are never heard, and that the idea of them is obtained by some other sense. And is there nothing in this contrary to nature and the truth of things?

H. To deal ingenuously, I do not like it. And after the concessions already made, I had as well grant that sounds too have no real being without the mind.

P. And I hope you will make no difficulty to acknowledge the same of colours.

H. Pardon me; the case of colours is very different. Can anything be plainer than that we see them on the objects? [...]

P. What! Are then the beautiful red and purple we see on yonder clouds really in them? Or do you imagine they have in themselves any other form than that of a dark mist or vapour?

H. I must own, Philonous, those colours are not really in the clouds as they seem to be at this distance. They are only apparent colours.

P. Apparent call you them? How shall we distinguish these apparent colours from real?

H. Very easily. Those are to be thought apparent which, appearing only at a distance, vanish upon a nearer approach.

P. And those I suppose are to be thought real, which are discovered by the most near and exact survey.

H. Right.

P. Is the nearest and exactest survey made by the help of a microscope, or by the naked eye?

H. By a microscope, doubtless.

P. But a microscope often discovers colours in an object different from those perceived by the unassisted sight. And in case we had microscopes magnifying to any assigned degree, it is certain that no object whatsoever viewed through them would

appear in the same colour which it exhibits to the naked eye.

H. And what will you conclude from all this? You cannot argue that there are really and naturally no colours on objects because by artificial managements they may be altered, or made to vanish.

P. I think it may evidently be concluded from your own concessions that all the colours we see with our naked eyes are only apparent as those on the clouds since they vanish upon a more close and accurate inspection which is afforded us by a microscope. Then as to what you say by way of prevention, I ask you, whether the real and natural state of an object is better discovered by a very sharp and piercing sight, or by one which is less sharp.

H. By the former without doubt.

P. Is it not plain from *dioptrics* that microscopes make the sight more penetrating, and represent objects as they would appear to the eye in case it were naturally endowed with a most exquisite sharpness?

H. It is.

P. Consequently the microscopical representation is to be thought that which best sets forth the real nature of the thing, or what it is in itself. The colours therefore by it perceived are more genuine and real than those perceived otherwise. [. . .]

H. I frankly own, Philonous, that it is in vain to stand out any longer. Colours, sounds, tastes, in a word, all those termed *secondary qualities*, have certainly no existence without the mind. But by this acknowledgment I must not be supposed to derogate anything from the reality of matter or external objects, seeing it is no more than several philosophers maintain, who nevertheless are the furthest imaginable from denying matter. For the clearer understanding of this, you must know sensible qualities are by philosophers divided into primary and secondary. The former are extension, figure, solidity, gravity, motion, and rest. And these they hold exist really in bodies. The latter are those above enumerated, or briefly, all sensible qualities.

beside the primary, which they assert are only so many sensations or ideas existing nowhere but in the mind. But all this, I doubt not, you are already apprised of. For my part, I have been a long time sensible there was such an opinion current among philosophers, but was never thoroughly convinced of its truth till now.

P. You are still then of opinion that extension and figures are inherent in external unthinking substances.

H. I am.

P. But what if the same arguments which are brought against secondary qualities will hold proof against these also?

H. Why then I shall be obliged to think, they too exist only in the mind.

P. Is it your opinion, the very figure and extension which you perceive by sense exist in the outward object or material substance?

H. It is.

P. Have all other animals as good grounds to think the same of the figure and extension which they see and feel?

H. Without doubt, if they have any thought at all.

P. Answer me, Hylas. Think you the senses were bestowed upon all animals for their preservation and well-being in life? Or were they given to men alone for this end?

H. I make no question but they have the same use in all other animals.

P. If so, is it not necessary they should be enabled by them to perceive their own limbs, and those bodies which are capable of harming them?

H. Certainly.

P. A mite therefore must be supposed to see his own foot, and things equal or even less than it, as bodies of some considerable dimension; though at the same time they appear to you scarce discernible, or at best as so many visible points.

H. I cannot deny it.

P. And to creatures less than the mite they will seem yet larger.

H. They will.

P. Insomuch that what you can hardly discern will to another extremely minute animal appear as some huge mountain.

H. All this I grant.

P. Can one and the same thing be at the same time in itself of different dimensions?

~~H. That were absurd to imagine.~~

P. But from what you have laid down it follows that both the extension by you perceived and that perceived by the mite itself, as likewise all those perceived by lesser animals, are each of them the true extension of the mite's foot, that is to say, by your own principles you are led into an absurdity.

H. There seems to be some difficulty in the point.

P. Again, have you not acknowledged that no real inherent property of any object can be changed without some change in the thing itself?

H. I have.

P. But as we approach to or recede from an object, the visible extension varies, being at one distance ten or a hundred times greater than at another. Doth it not therefore follow from hence likewise that it is not really inherent in the object?

H. I own I am at a loss what to think.

P. Your judgment will soon be determined, if you will venture to think as freely concerning this quality as you have done concerning the rest. Was it not admitted as a good argument that neither heat nor cold was in the water because it seemed warm to one hand and cold to the other?

H. It was.

P. Is it not the very same reasoning to conclude, there is no extension or figure in an object because to one eye it shall seem little, smooth, and round, when at

the same time it appears to the other, great, uneven, and angular?

H. The very same. But does this latter fact ever happen?

P. You may at any time make the experiment by looking with one eye bare, and with the other through a microscope.

H. I know not how to maintain it, and yet I am loath to give up extension; I see so many odd consequences following upon such a concession.

P. Odd, say you? After the concessions already made, I hope you will stick at nothing for its oddness. But on the other hand should it not seem very odd, if the general reasoning which includes all other sensible qualities did not also include extension? If it be allowed that no idea nor anything like an idea can exist in an unperceiving substance, then surely it follows that no figure or mode of extension, which we can either perceive or imagine, or have any idea of, can be really inherent in matter—not to mention the peculiar difficulty there must be in conceiving a material substance, prior to and distinct from extension, to be the substratum of extension. Be the sensible quality what it will—figure, or sound, or colour, it seems alike impossible it should subsist in that which doth not perceive it.

H. I give up the point for the present, reserving still a right to retract my opinion, in case I shall hereafter discover any false step in my progress to it.

P. That is a right you cannot be denied. Figures and extension being dispatched, we proceed next to motion. Can a real motion in any external body be at the same time both very swift and very slow?

H. It cannot.

P. Is not the motion of a body swift in a reciprocal proportion to the time it takes up in describing any given space? Thus a body that describes a mile in an hour moves three times faster than it would in case it described only a mile in three hours.

H. I agree with you.

P. And is not time measured by the succession of ideas in our minds?

H. It is.

P. And is it not possible ideas should succeed one another twice as fast in your mind as they do in mine, or in that of some spirit of another kind?

H. I own it.

P. Consequently the same body may to another seem to perform its motion over any space in half the time that it doth to you. And the same reasoning will hold as to any other proportion: that is to say, according to your principles (since the motions perceived are both really in the object) it is possible one and the same body shall be really moved the same way at once, both very swift and very slow. How is this consistent either with common sense, or with what you just now granted?

H. I have nothing to say to it.

P. Then as for solidity: either you do not mean any sensible quality by that word, and so it is beside our inquiry: or if you do, it must be either hardness or resistance. But both the one and the other are plainly relative to our senses: it being evident that what seems hard to one animal may appear soft to another, who hath greater force and firmness of limbs. Nor is it less plain that the resistance I feel is not in the body.

H. I own the very sensation of resistance, which is all you immediately perceive, is not in the *body*, but the cause of that sensation is.

P. But the causes of our sensations are not things immediately perceived, and therefore not sensible. This point I thought had been already determined.

H. I own it was; but you will pardon me if I seem a little embarrassed: I know not how to quit my old notions.

P. To help you out, do but consider that if extension be once acknowledged to have no existence without the mind, the same must necessarily be granted of motion, solidity, and gravity, since they all evidently suppose extension. It is therefore superfluous to inquire particularly concerning each of them. In denying extension, you have denied them all to have any real existence.

The Second Dialogue

HYLAS. I beg your pardon, Philonous, for not meeting you sooner. All this morning my head was so filled with our late conversation that I had not leisure to think of the time of the day, or indeed of anything else.

PHILONOUS. I am glad you were so intent upon it, in hopes if there were any mistakes in your concessions, or fallacies in my reasonings from them, you will now discover them to me.

H. I assure you, I have done nothing ever since I saw you but search after mistakes and fallacies, and with that view have minutely examined the whole series of yesterday's discourse: but all in vain, for the notions it led me into, upon review, appear still more clear and evident; and the more I consider them, the more irresistibly do they force my assent. [. . .]

P. Well then, are you at length satisfied that no sensible things have a real existence, and that you are in truth an arrant skeptic?

H. It is too plain to be denied. [. . .] My comfort is, you are as much a skeptic as I am.

P. There, Hylas, I must beg leave to differ from you.

H. What! Have you all along agreed to the premises, and do you now deny the conclusion, and leave me to maintain those paradoxes by myself which you led me into? This surely is not fair.

P. I deny that I agreed with you in those notions that led to skepticism. You indeed said, the reality of sensible things consisted in an absolute existence out of the minds of spirits, or distinct from their being perceived. And pursuant to this notion of reality, you are obliged to deny sensible things any real existence: that is, according to your own definition, you profess yourself a skeptic. But I neither said nor thought the reality of sensible things was to be defined after that manner. To me it is evident, for the reasons you allow of, that sensible things cannot exist otherwise than in a mind or spirit. Whence I conclude, not that they have no real existence, but

that seeing they depend not on my thought, and have all existence distinct from being perceived by me, there must be some other mind wherein they exist. As sure therefore as the sensible world really exists, so sure is there an infinite, omnipresent Spirit who contains and supports it.

H. What! This is no more than I and all Christians hold; nay, and all others too who believe there is a God, and that he knows and comprehends all things.

P. Ay, but here lies the difference. Men commonly believe that all things are known or perceived by God

because they believe the being of a God, whereas I, on the other side, immediately and necessarily conclude the being of a God because all sensible things must be perceived by him.

H. But so long as we all believe the same thing, what matter is it how we come by that belief?

P. But neither do we agree in the same opinion. For philosophers, though they acknowledge all corporeal beings to be perceived by God, yet they attribute to them an absolute subsistence distinct from their being perceived by any mind whatever, which I do not. [. . .]

Study Questions

1. Explain the distinction between primary and secondary qualities. Describe one of Philonous's arguments against the mind-independent existence of sensible things that is directed toward a secondary quality. Describe one of Philonous's arguments against the mind-independent existence of sensible things that is directed toward a primary quality.
2. What positions do Berkeley's two characters, Philonous and Hylas, take regarding the claim "Esse est percipi" ("To be is to be perceived")?
3. Why is Berkeley considered to be an idealist—specifically, a "subjective idealist"?
4. We have all heard that crazy-sounding question: If a tree falls in the forest and no one is around to hear, does it make a sound? Making reference to the distinction between primary and secondary qualities, explain why this question actually has philosophical depth. What is your response? Justify your response by addressing a likely objection.

David Hume

4/8/15

David Hume (1711–76) is considered today to be history's greatest English-language philosopher, but during his lifetime, his reputation as an atheist and skeptic blocked him from holding any academic posts. Employment as Librarian to the Edinburgh Faculty of Advocates during the 1750s offered Hume the opportunity to write his six-volume *History of England*, which became a best-seller and provided him with financial independence. While serving in a diplomatic capacity in Paris during the 1760s, Hume associated with philosophes such as Diderot, d'Holbach, and Rousseau. Hume's *Enquiry* recasts arguments from the *Treatise of Human Understanding* in order to reach a wider audience.

From An Enquiry Concerning Human Understanding

Of the Origin of Ideas [Section 2]

Everyone will readily allow that there is a considerable difference between the perceptions of the mind when a man feels the pain of excessive

heat or the pleasure of moderate warmth, and when he afterwards recalls to his memory this sensation or anticipates it by his imagination. These faculties may mimic or copy the perceptions of the senses, but they never can entirely reach the force and

vivacity of the original sentiment. The utmost we say of them, even when they operate with greatest vigour, is that they represent their object in so lively a manner that we could *almost* say we feel or see it. But, except the mind be disordered by disease or madness, they never can arrive at such a pitch of vivacity as to render these perceptions altogether undistinguishable. All the colours of poetry, however splendid, can never paint natural objects in such a manner as to make the description be taken for a real landscape. The most lively thought is still inferior to the dullest sensation.

We may observe a like distinction to run through all the other perceptions of the mind. A man in a fit of anger is actuated in a very different manner from one who only thinks of that emotion. If you tell me that any person is in love, I easily understand your meaning and form a just conception of his situation, but never can mistake that conception for the real disorders and agitations of the passion. When we reflect on our past sentiments and affections, our thought is a faithful mirror and copies its objects truly, but the colours which it employs are faint and dull, in comparison of those in which our original perceptions were clothed. It requires no nice discernment or metaphysical head to mark the distinction between them.

Here therefore we may divide all the perceptions of the mind into two classes or species, which are distinguished by their different degrees of force and vivacity. The less forcible and lively are commonly denominated *Thoughts or Ideas*. The other species want a name in our language and in most others—I suppose because it was not requisite for any but philosophical purposes to rank them under a general term or appellation. Let us, therefore, use a little freedom and call them *Impressions*, employing that word in a sense somewhat different from the usual. By the term *impression*, then, I mean all our more lively perceptions, when we hear, or see, or feel, or love, or hate, or desire, or will. And impressions are distinguished from ideas, which are the less lively perceptions, of which we are conscious when we reflect on any of those sensations or movements above mentioned.

Nothing, at first view, may seem more unbounded than the thought of man, which not only escapes all human power and authority but is not even restrained within the limits of nature and reality. To form monsters and join incongruous shapes and appearances costs the imagination no more trouble than to conceive the most natural and familiar objects. And while the body is confined to one planet, along which it creeps with pain and difficulty, the thought can in an instant transport us into the most distant regions of the universe, or even beyond the universe into the unbounded chaos, where nature is supposed to lie in total confusion. What never was seen or heard of may yet be conceived; nor is anything beyond the power of thought, except what implies an absolute contradiction.

But though our thought seems to possess this unbounded liberty, we shall find upon a nearer examination that it is really confined within very narrow limits, and that all this creative power of the mind amounts to no more than the faculty of compounding, transposing, augmenting, or diminishing the materials afforded us by the senses and experience. When we think of a golden mountain, we only join two consistent ideas, *gold* and *mountain*, with which we were formerly acquainted. A virtuous horse we can conceive because from our own feeling, we can conceive virtue, and this we may unite to the figure and shape of a horse, which is an animal familiar to us. In short, all the materials of thinking are derived either from our outward or inward sentiment: the mixture and composition of these belongs alone to the mind and will. Or, to express myself in philosophical language, all our ideas or more feeble perceptions are copies of our impressions or more lively ones.

To prove this, the two following arguments will, I hope, be sufficient. First, when we analyze our thoughts or ideas, however compounded or sublime, we always find that they resolve themselves into such simple ideas as were copied from a precedent feeling or sentiment. Even those ideas, which at first view seem the most wide of this origin, are found, upon a nearer scrutiny, to be derived from it. The idea of

God, as meaning an infinitely intelligent, wise, and good Being, arises from reflecting on the operations of our own mind, and augmenting, without limit, those qualities of goodness and wisdom. We may prosecute this enquiry to what length we please, where we shall always find that every idea which we examine is copied from a similar impression. Those who would assert that this position is not universally true nor without exception have only one easy method of refuting it: by producing that idea, which, in their opinion, is not derived from this source. It will then be incumbent on us, if we would maintain our doctrine, to produce the impression, or lively perception, which corresponds to it.

Secondly. If it happens, from a defect of the organ, that a man is not susceptible of any species of sensation, we always find that he is as little susceptible of the correspondent ideas. A blind man can form no notion of colours, a deaf man of sounds. Restore either of them that sense in which he is deficient, by opening this new inlet for his sensations, you also open an inlet for the ideas, and he finds no difficulty in conceiving these objects. The case is the same if the object proper for exciting any sensation has never been applied to the organ. A Laplander or Negro has no notion of the relish of wine. And though there are few or no instances of a like deficiency in the mind where a person has never felt or is wholly incapable of a sentiment or passion that belongs to his species, yet we find the same observation to take place in a lesser degree. A man of mild manners can form no idea of inveterate revenge or cruelty, nor can a selfish heart easily conceive the heights of friendship and generosity.

There is, however, one contradictory phenomenon, which may prove that it is not absolutely impossible for ideas to arise independent of their correspondent impressions. I believe it will readily be allowed that the several distinct ideas of colour which enter by the eye or those of sound which are conveyed by the ear are really different from each other, though, at the same time, resembling. Now, if this be true of different colours, it must be no less so of the different shades of the same colour—that

each shade produces a distinct idea, independent of the rest. For if this should be denied, it is possible by the continual gradation of shades to run a colour insensibly into what is most remote from it; and if you will not allow any of the means to be different, you cannot without absurdity deny the extremes to be the same. Suppose, therefore, a person to have enjoyed his sight for thirty years, and to have become perfectly acquainted with colours of all kinds except one particular shade of blue, for instance, which it never has been his fortune to meet with. Let all the different shades of that colour except that single one be placed before him, descending gradually from the deepest to the lightest; it is plain that he will perceive a blank where that shade is wanting and will be sensible that there is a greater distance in that place between the contiguous colours than in any other. Now I ask, whether it be possible for him, from his own imagination, to supply this deficiency and raise up to himself the idea of that particular shade, though it had never been conveyed to him by his senses? I believe there are few but will be of opinion that he can. This may serve as a proof that the simple ideas are not always, in every instance, derived from the correspondent impressions; though this instance is so singular that it is scarcely worth our observing, and does not merit that for it alone we should alter our general maxim.

Here, therefore, is a proposition, which not only seems in itself simple and intelligible, but, if a proper use were made of it, might render every dispute equally intelligible and banish all that jargon which has so long taken possession of metaphysical reasonings and drawn disgrace upon them. All ideas, especially abstract ones, are naturally faint and obscure; the mind has but a slender hold of them; they are apt to be confounded with other resembling ideas; and when we have often employed any term, though without a distinct meaning, we are apt to imagine it has a determinate idea annexed to it. On the contrary, all impressions, that is, all sensations, either outward or inward, are strong and vivid; the limits between them are more exactly determined; nor is it easy to fall into any error or mistake with regard to

them. When we entertain, therefore, any suspicion that a philosophical term is employed without any meaning or idea (as is but too frequent), we need but enquire, from what impression is that supposed idea derived? And if it be impossible to assign any, this will serve to confirm our suspicion. By bringing ideas into so clear a light, we may reasonably hope to remove all dispute which may arise concerning their nature and reality.

Skeptical Doubts Concerning the Operations of the Understanding [Section 4]

Part I. On Cause and Effect

All the objects of human reason or enquiry may naturally be divided into two kinds, to wit, Relations of Ideas and Matters of Fact. Of the first kind are the sciences of Geometry, Algebra, and Arithmetic; and in short, every affirmation which is either intuitively or demonstratively certain. That the square of the hypotenuse is equal to the square of the two sides is a proposition which expresses a relation between these figures. That three times five is equal to the half of thirty expresses a relation between these numbers. Propositions of this kind are discoverable by the mere operation of thought, without dependence on what is anywhere existent in the universe. Though there never were a circle or triangle in nature, the truths demonstrated by Euclid would forever retain their certainty and evidence.

Matters of fact, which are the second objects of human reason, are not ascertained in the same manner; nor is our evidence of their truth, however great, of a like nature with the foregoing. The contrary of every matter of fact is still possible because it can never imply a contradiction, and is conceived by the mind with the same facility and distinctness as if ever so conformable to reality. That the sun will not rise tomorrow is no less intelligible a proposition, and implies no more contradiction than the affirmation, that it will rise. We should in vain, therefore, attempt to demonstrate its falsehood. Were it demonstratively false, it would imply

a contradiction, and could never be distinctly conceived by the mind.

It may, therefore, be a subject worthy of curiosity to enquire what is the nature of that evidence which assures us of any real existence and matter of fact beyond the present testimony of our senses, or the records of our memory.

All reasonings concerning matter of fact seem to be founded on the relation of Cause and Effect. By means of that relation alone we can go beyond the evidence of our memory and senses. If you were to ask a man why he believes any matter of fact which is absent—for instance, that his friend is in the country or in France, he would give you a reason, and this reason would be some other fact—as a letter received from him, or the knowledge of his former resolutions and promises. A man finding a watch or any other machine in a desert island would conclude that there had once been men in that island. [...] If we anatomize all the other reasonings of this nature, we shall find that they are founded on the relation of cause and effect, and that this relation is either near or remote, direct or collateral. Heat and light are collateral effects of fire, and the one effect may justly be inferred from the other.

If we would satisfy ourselves, therefore, concerning the nature of that evidence which assures us of matters of fact, we must enquire how we arrive at the knowledge of cause and effect.

I shall venture to affirm, as a general proposition which admits of no exception, that the knowledge of this relation is not in any instance attained by reasonings a priori, but arises entirely from experience when we find that any particular objects are constantly conjoined with each other. Let an object be presented to a man of ever so strong natural reason and abilities, and if that object be entirely new to him, he will not be able, by the most accurate examination of its sensible qualities, to discover any of its causes or effects. Adam, though his rational faculties be supposed at the very first entirely perfect, could not have inferred from the fluidity and transparency of water that it would suffocate him, or from the light and warmth of fire that it would consume him.

himself as having a "constructionist orientation".

...euge?

Anjan Chakravartty

Anjan Chakravartty is a professor of philosophy at the University of Notre Dame, having moved there from the University of Toronto's Institute for the History and Philosophy of Science and Technology. Chakravartty earned his PhD and MPhil in history and philosophy of science at the University of Cambridge and an MA in philosophy and BSC in biophysics at the University of Toronto. *A Metaphysics for Scientific Realism* won the Canadian Philosophical Association Book Prize in 2009.

Realism and Anti-realism; Metaphysics and Empiricism

From *A Metaphysics for Scientific Realism: Knowing the Unobservable*

1. The Trouble with Common Sense

Hanging in my office is a framed photograph of an armillary sphere, which resides in the Whipple Museum of the History of Science in Cambridge, England. An armillary sphere is a celestial globe. It is made up of a spherical model of the planet Earth (the sort we all played with as children), but the model is surrounded by an intricate skeleton of graduated rings, representing the most important celestial circles. Armillary spheres were devised in ancient Greece and developed as instruments for teaching and astronomical calculation. During the same period, heavenly bodies were widely conceived as fixed to the surfaces of concentrically arranged crystalline spheres, which rotate around the Earth at their centre.

This particular armillary sphere has, I expect, many fascinating historical stories to tell, but there is a specific reason I framed the picture. Once upon a time, astronomers speculated about the causes and

mechanisms of the motions of the planets and stars, and their ontology of crystalline spheres was a central feature of astronomical theory for hundreds of years. But crystalline spheres are not the sorts of things one can observe, at least not with the naked eye from the surface of the Earth. Even if it had turned out that they exist, it is doubtful one would have been able to devise an instrument to detect them before the days of satellites and space shuttles. Much of the energy of the sciences is consumed in the attempt to work out and describe things that are inaccessible to the unaided senses, whether in practice or in principle. My armillary sphere, with its glorious and complicated mess of interwoven circles, is a reminder of past testaments to that obsession.

In describing the notion of a crystalline sphere, I have already made some distinctions. There are things that one can, under favourable circumstances, perceive with one's unaided senses. Let us call them "observables," though this is to privilege vision

over the other senses for the sake of terminological convenience. Unobservables, then, are things one cannot perceive with one's unaided senses, and this category divides into two subcategories. Some unobservables are nonetheless detectable through the use of instruments with which one hopes to "extend" one's senses, and others are simply undetectable. These distinctions are important, because major controversies about how to interpret the claims of the sciences revolve around them. In this chapter, I will briefly outline the most important positions engaged in these controversies, and consider how the tension between speculative metaphysics and empiricism has kept them alive.

There are occasional disputes about what counts as science—concerning how best to exclude astronomy but include astronomy, about what to say to creationists unhappy with the teaching of evolutionary biology in schools, etc. I leave these disputes to one side here, and begin simply with what are commonly regarded as sciences today. It is widely held that the sciences are not merely knowledge-producing endeavours, but the means of knowledge production *par excellence*. Scientific inquiry is our best hope for gaining knowledge of the world, the things that compose it, its structure, its laws, and so on. And the more one investigates, the better it gets. Scientific knowledge is progressive; it renders the natural world with increasing accuracy.

Scientific realism, to a rough, first approximation, is the view that scientific theories correctly describe the nature of a mind-independent world. Outside of philosophy, realism is usually regarded as common sense, but philosophers enjoy subjecting common-place views to thorough scrutiny, and this one certainly requires it. The main consideration in favour of realism is ancient, but more recently referred to as the "miracle argument" (or "no-miracles argument") after the memorable slogan coined by Hilary Putnam that realism "is the only philosophy that doesn't make the success of science a miracle" (73). Scientific theories are amazingly successful in that they allow us to predict, manipulate, and participate in worldly phenomena, and the most straightforward explanation of this is that they correctly describe the nature of the

world, or something close by. In the absence of this explanation the success afforded by the sciences might well seem miraculous, and, given the choice, one should always choose common sense over miracles.

Some have questioned the need for an explanation of the success of science at all. Bas van Fraassen (23–5, 34–40), for example, suggests that successful scientific theories are analogous to well-adapted organisms. There is no need to explain the success of organisms, he says. Only well-adapted organisms survive, just as only well-adapted theories survive, where "well-adapted" in the latter case means adequate to the tasks to which one puts theories. These tasks are generally thought to include predictions and retrodictions (predictions concerning past phenomena), and perhaps most impressively novel predictions (ones about classes of things or phenomena one has yet to observe). A well-adapted theory is one whose predictions, retrodictions, and novel predictions, if any, are borne out in the course of observation and experimentation. But saying that successful theories are ones that are well-adapted may be tantamount to the tautology that successful theories are successful, which is not saying much. Whatever the merits of the Darwinian analogy for theories generally, one might still wonder why any given theory (organism) survives for the time it does, and this may require a more specific consideration of the properties of the theory (organism) in virtue of which it is well adapted. [...] The attempt to satisfy the desire for an explanation of scientific success has produced the bulk of the literature on scientific realism. As arguments go, the miracle argument is surprisingly poor, all things considered, and consequently alternatives to realism have flourished. The poverty of the miracle argument and consequent flourishing of rivals to realism stem from difficulties presented by three general issues, which I will mention only briefly:

1. the use of abductive inference, or inference to the best explanation (IBE)
2. the underdetermination of theory choice by data or evidence (UTD)
3. discontinuities in scientific theories over time, yielding a pessimistic induction (PI)

Abduction is a form of inference famous from the writings of Charles Sanders Peirce, inspiring what is now generally called "inference to the best explanation" (some use the term synonymously with "abduction" while others, more strictly, distinguish it from Peirce's version). IBE offers the following advice to inference makers: infer the hypothesis that, if true, would provide the best explanation for whatever it is you hope to explain. Note that the miracle argument itself is an abductive argument. Why are scientific theories so successful at making predictions and accounting for empirical data? One answer is that they are true, and this seems, to the realist at any rate, the best explanation. One might even think it the only conceivable explanation, but as we shall see, in light of UTD and PI, this is highly contestable.

First, however, let us turn from the particular case of the miracle argument to the merits of IBE as a form of inference in general. There is little doubt that this sort of inferential practice is fundamental to everyday and scientific reasoning. The decision to adopt one theory as opposed to its rivals, for example, is generally a complex process involving many factors, but IBE will most certainly figure at some stage.

Antirealists are quick to point out that in order for an instance of IBE to yield the truth, two conditions must be met. Firstly, one must rank the rival hypotheses under consideration correctly with respect to the likelihood that they are true. Secondly, the truth must be among the hypotheses one is considering. But can one ensure that these conditions are met? Regarding the first, it is difficult to say what features a truth-like explanation should have. Beyond the minimum criterion of some impressive measure of agreement with outcomes of observation and experiment, possible indicators of good explanations have been widely discussed. Some hold that theories characterized by features such as simplicity, elegance, and unity (with other theories or domains of inquiry) are preferable. Quite apart from the matter of describing what these virtues are, however, it is not immediately obvious that such virtues have anything to do with truth. There is no *a priori* reason, one might argue, to reject the possibility that

natural phenomena are rather complex, inelegant, and disjoint. And regarding the second condition for successful IBE, in most cases it is difficult to see how one could know in advance that the true hypothesis is among those considered. (A case in which one does have this knowledge is where rival hypotheses are contradictions.)

In practice it is often difficult to produce even one theory that explains the empirical data, let alone rivals. This, however, does not diminish the seriousness of the problem. In fact, it turns out that it may be irrelevant whether one ever has a choice to make between rival theories in practice. For some maintain that rival theories are always possible, whether or not one has thought of them, and this is sufficient to raise concerns about IBE. Confidence in the possibility of rivals stems from the underdetermination thesis, or UTP. Its canonical formulation due to Pierre Duhem, later expressed in rather different terms by W.V.O. Quine (hence also called the "Duhem-Quine thesis"), goes this way. Theoretical hypotheses rarely if ever yield predictions by themselves. Rather, they must be conjoined with auxiliary hypotheses—background theories, related theories, theories about the measurement of relevant parameters, etc.—in order to yield predictions. If observation and experimentation produce data that are not as one predicts, one has a choice to make concerning which of the prediction-yielding hypotheses is culpable. One can always preserve a favoured hypothesis at the expense of something else. Since there are different ways of choosing how to account for recalcitrant data, different overall theories or conjunctives of hypotheses may be used to account for the empirical evidence. Thus, in general, there is always more than one overall theory consistent with the data.

In more contemporary discussions, UTD is usually explicated differently. Given a theory, T_1 , it is always possible to generate an empirically equivalent but different theory, T_2 . T_2 is a theory that makes precisely the same claims regarding observable phenomena as T_1 , but differs in other respects. T_2 might, for example, exclude all of the unobservable entities and processes of T_1 , or replace some or all of these with others, or simply alter them, but in such

a way as to produce exactly the same observable predictions. Given that this sort of manoeuvring is always possible, how does one decide between rival theories so constructed? Here again the realist must find a way to infer to a particular theory at the expense of its rivals, with the various difficulties this engenders.

In addition to challenges concerning IBE and UTD, at least one anti-realist argument aspires to the status of an empirical refutation of realism. PI, or as it is often called, the “pessimistic meta-induction,” can be summarized as follows. Consider the history of scientific theories in any particular domain. From the perspective of the present, most past theories are considered false, strictly speaking. There is evidence of severe discontinuity over time, regarding both the entities and processes described. This evidence makes up a catalogue of instability in the things to which theories refer. By induction based on these past cases, it is likely that present-day theories are also false and will be recognized as such in the future. Realists are generally keen to respond that not even they believe that theories are true *simpliciter*. Scientific theorizing is a complex business, replete with things like approximation, abstraction, and idealization. What is important is that successive theories get better with respect to the truth, coming closer to it over time. It is the progress sciences make in describing nature with increasing accuracy that fuels realism. Good theories, they say, are normally “approximately true,” and more so as the sciences progress. Giving a precise account of what “approximate truth” means, however, is no easy task.

So much for common sense. The promise of scientific realism is very much open to debate, and in light of IBE, UTD, and PI, this debate has spawned many positions. Let us take a look at the main players, so as to gain a better understanding of the context of realism.

2. A Conceptual Taxonomy

Earlier I described realism as the view that scientific theories correctly describe the nature of a

mind-independent world. This is shorthand for the various and more nuanced commitments realists tend to make. For example, many add that they are not realists about all theories, just ones that are genuinely successful. The clarification is supplied to dissolve the potential worry that realists must embrace theories that seem artificially successful—those that do not make novel predictions and simply incorporate past empirical data on an *ad hoc* basis, for instance. Realists often say that their position extends only to theories that are sufficiently “mature.” Maturity is an admittedly vague notion, meant to convey the idea that a theory has withstood serious testing in application to its domain over some significant period of time, and some correlate the maturity of disciplines more generally with the extent to which their theories make successful, novel predictions. Finally, as I have already mentioned, it is also standard to qualify that which theories are supposed to deliver: it is said that theoretical descriptions may not be true, *per se*, but that they are nearly or approximately true, or at least more so than earlier descriptions.

With these caveats in mind it may be instructive to situate scientific realism in a broader context, as a species of the genus of positions historically described as realisms. Traditionally, “realism” simply denotes a belief in the reality of something—an existence that does not depend on minds, human or otherwise. Consider an increasingly ambitious sequence of items about which one might be a realist. One could begin with the objects of one’s perceptions (goldfish, fishbowls), move on to objects beyond one’s sensory abilities to detect (genes, electrons), and further still, beyond the realm of the concrete to the realm of the abstract, to non-spatiotemporal things such as numbers, sets, universals, and propositions. The sort of realist one is, if at all, can be gauged from the sorts of things one takes to qualify for mind-independent existence. Though I have just described these commitments as forming a sequence, it should be understood that realism at any given stage does not necessarily entail realism about anything prior to that stage. Some Platonists, for example, appear to hold that ultimately, the only

real objects are abstract ones, the Forms, or that the Forms are in some sense "more real" than observables. Scientific realism, in committing to something approaching the truth of scientific theories, makes a commitment to their subject matter: entities and processes involving their interactions, at the level of both the observable and the unobservable. Anything more detailed is a matter for negotiation, and realists have many opposing views beyond this shared, minimal commitment. [...]

I said that "realism" traditionally denotes a belief in the reality of something, but in the context of scientific realism the term has broader connotations. The most perspicuous way of understanding these aspects is in terms of three lines of inquiry: ontological, semantic, and epistemological. Ontologically, scientific realism is committed to the existence of a mind-independent world or reality. A realist semantics implies that theoretical claims about this reality have truth values, and should be construed literally, whether true or false. I will consider an example of what it might mean to construe claims in a non-literal way momentarily. Finally, the epistemological commitment is to the idea that these theoretical claims give us knowledge of the world. That is, predictively successful (mature, non-*ad hoc*) theories, taken literally as describing the nature of a mind-independent reality are (approximately) true. The things our best scientific theories tell us about entities and processes are

decent descriptions of the way the world really is. Henceforth I will use the term "realism" to refer to this scientific variety only. We are now ready to locate it and various other positions in a conceptual space.

If by "anti-realism" one means any view opposed to realism, many different positions will fit the bill. Exploiting the differences in commitments along our three lines of inquiry, one may construct a taxonomy of views discussed in connection with these debates. Table 1 lists the most prominent of these, and for each notes how it stands on the existence of a mind-independent world, on whether theoretical statements should be taken literally, and on whether such claims yield knowledge of their putative subject matter. This is a blunt instrument; an impressive array of viewpoints is not adequately reflected in this simple classificatory scheme, and the reflections present are imprecise. There are many ways, for example, in which to be a sceptic. But the core views sketched in Table 1 offer some basic categories for locating families of related commitments.

Traditionally and especially in the early twentieth century, around the time of the birth of modern analytic philosophy, realist positions were contrasted with idealism, according to which there is no world external to and thus independent of the mental. The classic statement of this position is credited to Bishop George Berkeley, for whom reality is constituted by thoughts and ultimately sustained by

Table 1 Scientific realism and anti-realisms

	The ontological question: mind-independent reality?	The semantic question: theories literally construed?	The epistemological question: knowledge?
Realism	yes	yes	yes
Constructive empiricism	yes	yes	observables: yes unobservables: no
Skepticism	yes	yes	no
Logical positivism/empiricism	yes/no?	observables: yes unobservables: no	yes
Traditional instrumentalism	yes	observables: yes unobservables: no	observables: yes unobservables: no
Idealism	no	no	yes

the mind of God. Idealism need not invoke a deity, though. A phenomenalist, for instance, might be an idealist without appealing to the divine. Given an idealist ontology, it is no surprise that scientific claims cannot be construed literally, since they are not about what they seem to describe at face value, but this of course does not preclude knowledge of a mind-dependent reality. As Table 1 shows, idealism is the only position considered here to take an unambiguous antirealist stand with respect to ontology.

Instrumentalism is a view shared by a number of positions, all of which have the following contention in common: theories are merely instruments for predicting observable phenomena or systematizing observation reports. Traditional instrumentalism is an even stronger view according to which, furthermore, claims involving unobservable entities and processes have no meaning at all. Such "theoretical claims," as they are called ("claims about unobservables" is better, I think, since theories describe observables too), do not have truth values. They are not even capable of being true or false; rather, they are mere tools for prediction. In common usage, however, some now employ the term in a weaker sense, to describe views that grant truth values to claims involving unobservables while maintaining that one is not in a position, for whatever reason, to determine what these truth values are. In this latter, weaker sense, constructive empiricism is sometimes described as a form of instrumentalism. And though I have represented instrumentalists in Table 1 as subscribing to realism in ontology, some would include those who do not.

Logical positivism, famously associated with the philosophers and scientists of the **Vienna Circle**, and its later incarnation, logical empiricism, are similar to traditional instrumentalism in having a strict policy regarding the unobservable. But where traditional instrumentalism holds that claims about unobservables are meaningless, logical empiricism assigns meaning to some of these claims by interpreting them non-literally. Rather than taking these claims at face value as describing the things they

appear to describe, claims about unobservables are meaningful for logical empiricists if and only if their unobservable terms are linked in an appropriate way to observable terms. The unobservable vocabulary is then treated as nothing more than a shorthand for the observation reports to which they are tied. "Electron," for example, might be shorthand in some contexts for "white streak in a cloud chamber," given the path of water droplets one actually sees in a cloud chamber experiment, along what is theoretically described as the trajectory of an electron. It is by means of such "correspondence rules" or "bridge principles" that talk of the unobservable realm is interpreted. Given a translation manual of this sort, theories construed non-literally are thought to yield knowledge of the world. The label "logical positivism/empiricism" covers vast ground, however, and views regarding the ontological status of the world described by science are far from univocal here. Rudolf Carnap, for instance, held that while theories furnish frameworks for systematizing knowledge, ontological questions "external" to such frameworks are meaningless, or have no cognitive content.

While traditional instrumentalism banishes meaningful talk about unobservables altogether and logical empiricism interprets it non-literally, constructive empiricism, the view advocated by van Fraassen, adopts a realist semantics. The anti-realism of this latter position is thus wholly manifested in its epistemology. For the constructive empiricist the observable-unobservable distinction is extremely important, but only in the realm of knowledge, and this feature marks the position as an interesting halfway house between realism and various kinds of skepticism. By skepticism here, I intend any position that agrees with the realist concerning ontology and semantics, but offers epistemic considerations to suggest that one does not have knowledge of the world, or at least that one is not in a position to know that one does. Constructive empiricism goes along with the skeptic part way, denying that one can have knowledge of the unobservable, but also with the realist part way, accepting that one can have knowledge of the observable. (More strictly,

constructive empiricism is the view that the aim of science is true claims about observables, not truth more generally, but this is usually interpreted in the way I have suggested.) By adopting a realist semantics, constructive empiricism avoids the semantic difficulties that were in large part responsible for the demise of logical empiricism in the latter half of the twentieth century, and has taken its place as the main rival to realism today.

Table 1 does not exhaust the list of “isms” opposed to realism. It does, however, provide a fairly comprehensive list of the reasons and motivations one might have for being an anti-realist. For example, the discipline known as the sociology of scientific knowledge is predominantly anti-realist. This is not a logical consequence, however, of the desire to study science from a sociological perspective. Sociologists who are anti-realists are usually so inclined because of commitments they share with one or more of the antirealist positions outlined in Table 1. Though I will not consider this approach to thinking about the sciences in any detail here, it is important to appreciate its influence. Sociological and related methodologies, which attempt to explicate scientific practice and its social, political, and economic relations, both internal and external, represent the major alternative approach to the study of the sciences today, contrasting with the more straightforwardly philosophical approach of realism and constructive empiricism.

3. Metaphysics, Empiricism, and Scientific Knowledge

Armed with a basic summary of realism and its principal rivals, let us turn to the central focus of this work. Earlier I said that much of the controversy surrounding these positions concerns the question of how one should understand scientific claims, in light of the distinctions between the observable and unobservable on one hand, and between the two categories of the unobservable, the detectable and undetectable, on the other. By examining these distinctions one may begin to shed some light on the roles that metaphysics and empiricism play in the

interpretation of scientific claims, and the dialectic between them. The first distinction, between observables and unobservables, concerns things that one can under favourable circumstances perceive with one's unaided senses, and things one cannot. Note that this use of “observable” and “unobservable” is different from what is often the case in the sciences themselves. In scientific practice the label “observable” is usually applied permissively to anything with which one can forge some sort of causal contact, as one does when one uses instruments (such as microscopes) for detection. In the present discussion, however, observables are strictly things one can perceive with the unmediated senses. As Table 1 attests, almost everyone thinks one can have knowledge of the observable. This is not to say, however, that interpreting claims about observables is necessarily straightforward. It may be, for example, that the categories of objects and processes one employs to express one's knowledge of the observable are interestingly shaped by the theories one adopts. Indeed, that this is the case for both observables and unobservables is a central tenet of the influential views of **Thomas Kuhn**, the sociological approaches that followed him, and even some of the logical empiricists who preceded him, who held that “conceptual schemes” shape one's knowledge of the world. [. . .]

It is the status of the unobservable that has proved most controversial. Logical positivism was, in effect, the founding movement of modern philosophy of science, and the radical empiricism of the positivists has had a lasting impact. It will be useful [. . .] to clarify my second distinction, between unobservables that are detectable and those that are not. Let me reserve the word “detectable” for unobservables one can detect using instruments but not otherwise, and “undetectable” for those one cannot detect at all (see Figure 1). The mitochondrion, for example, is a cellular organelle in which substances are oxidized to produce energy. Though unobservable, one can detect mitochondria using microscopy. A celebrated historical example of a more indirect case of detection is the neutrino, a subatomic particle originally posited by Wolfgang Pauli and theorized about

by Enrico Fermi in the 1930s. The neutrino was hypothesized to allow for the conservation of mass-energy and angular momentum in certain subatomic interactions, such as the β -decay of radium-210, and detections of such interactions might thus be viewed as indirect detections of neutrinos. It was not until 1956 that Frederick Reines and Clyde Cowan successfully performed an experiment in which neutrinos were detected more directly. Now consider unobservables whose putative existence cannot be the subject of empirical investigation, whether in practice or in principle. Examples include Newton's conceptions of position and velocity with respect to absolute space, and causally inefficacious entities such as mathematical objects. Even if they exist, such things are undetectable.

Historically, the most pressing challenges to realism have come from those adopting some form of empiricism. This is not to say, however, that all empiricists are anti-realists! It may be helpful here to note that empiricism is traditionally associated with two strands of thought which often come together, interwoven. One strand is the idea that sensory experience is the *source* of all knowledge

of the world, and this by itself does not preclude an empiricist from being a realist. A realist might accept this first strand while further believing that one can infer the existence of certain unobservables on the basis of the evidence of one's senses. The second strand of empiricism is the idea that all knowledge of the world is *about* experience, and it is this tenet that conflicts with realism, since realists believe claims about things that transcend experience in addition to claims about observables. So an empiricist of the first strand alone may be a realist, but not one of the first and second strands combined. The most adamant critiques of realism stem from those who are committed to the second strand of empiricism, and to violate this commitment is to engage in what its advocates view as a fruitless and misconceived philosophical activity: speculative metaphysics.

What is this metaphysics, then, of which so many empiricists disapprove? To say that there is a conflict between metaphysics and empiricism *simpliciter* is too strong, since many empiricists do metaphysics as it is understood most broadly, as the study of the first or basic principles of philosophy,

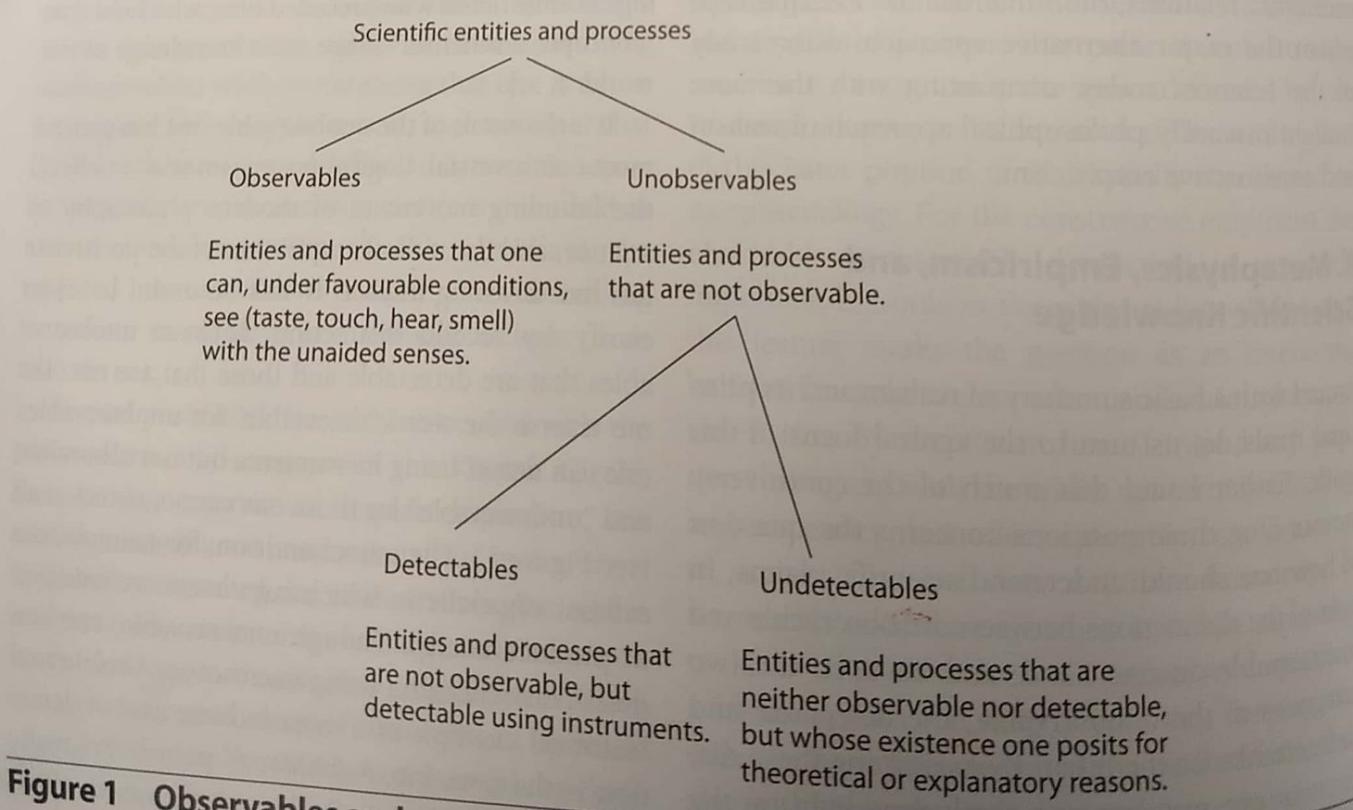


Figure 1 Observables and unobservables

being *qua* being, and the natures of things that exist. The metaphysics that empiricists disavow concerns the unobservable, and thus any position that endorses speculation of this sort, leading to substantive beliefs about detectables or undetectables, is unacceptable to them. This includes not only speculations about things like universals and causal necessity, which are familiar topics within metaphysics, but also speculations about mitochondria and neutrinos, which are familiar topics within the sciences. But empiricists are generally happy to do metaphysics so long as it does not involve believing speculations about the unobservable. Thus Hume gives an account of causation, not in terms of undetectable necessary connections, but solely in terms of observable events that follow one another. And thus nominalists speak of properties, not as abstract entities like universals, but as sets of observable things to which the predicates associated with these properties apply. The unobservable is likewise an anathema to many empiricist accounts of science. The scientific realist, in maintaining that one can have knowledge of scientific unobservables, engages in the very sort of metaphysical speculation these empiricists reject. [. . .]

The metaphysics of the sciences concerns the observable and unobservable parts of the world described by scientific theories, both explicitly and implicitly. The epistemology of the sciences concerns the specific methods used to generate scientific claims, the justification or confirmation of these claims, whether they constitute knowledge, and if so, what sort. The influence of logical positivism during the birth of the philosophy of science as a separate discipline in the late nineteenth century, and throughout most of the twentieth century, led to a vestigial neglect of metaphysical questions in connection with realism. Those investigating problems such as the nature of causation, laws of nature, and conceptions of natural kinds have done so largely in isolation from debates between realists and antirealists. Metaphysical issues have been the purview of the philosophy of particular sciences: space and time, evolutionary biology, quantum mechanics, and so on. The neglect of metaphysics in the context of realism, however, is a mistake. For there is a sense in which the metaphysics of science is a precursor to its epistemology. One cannot fully appreciate what it might mean to be a realist until one has a clear picture of what one is being invited to be a realist about. [. . .]

Works Cited

- Carnap, R. "Empiricism, Semantics and Ontology." *Revue Internationale de Philosophie* 4 (1950): 20–40.
- Kuhn, T.S. *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press, 1970/1962.
- Putnam, H. *Mathematics, Matter and Method*. Cambridge: Cambridge University Press, 1975.
- van Fraassen, B.C. *The Scientific Image*. Oxford: Oxford University Press, 1980.

Study Questions

- Chakravartty notes that although scientific realism seems like common sense, it is a matter of philosophical debate. Outline the debate between Putnam and van Fraassen over whether the success of science justifies scientific realism.
- What challenges do inference to the best explanation, under-determination of theory choice by data, and pessimistic induction pose for scientific realism?
- Describe the ontological, semantic, and epistemological commitments of scientific realism. Explain how each of these commitments is challenged by some variety of anti-realism.
- How does Chakravartty distinguish between "observables" and "unobservables"? Making reference to this distinction, assess the empiricist challenge to metaphysics.

~ PERSONAL IDENTITY ~

John Locke

John Locke (1632–1704) was an English philosopher and physician. He is especially known for his 1689 *An Essay Concerning Human Understanding*. In the first two books of the *Essay*, Locke defends empiricism and the theory of the mind as a *tabula rasa*, or blank slate: knowledge is not innate but arises from the mind's capacities to act on ideas gained entirely through the senses. In this chapter on personal identity, added to the second edition of the *Essay*, Locke rejects Descartes' equation of the person with the immaterial soul.—TR

Of Identity and Diversity

From *An Essay Concerning Human Understanding*

4. [...] **Principium Individuationis.** The principium individuationis [...], it is plain, is existence itself [...]: e.g., let us suppose an atom [...] existing in a determined time and place; it is evident that considered in any instant of its existence, it is in that instant the same with itself. For, being at that instant what it is and nothing else, it is the same, and so must continue as long as its existence is continued; for so long it will be the same, and no other. In like manner, if two or more atoms be joined together into the same mass, every one of those atoms will be the same by the foregoing rule: and whilst they exist united together, the mass, consisting of the same atoms, must be the same mass, or the same body, let the parts be ever so differently jumbled. But if one of these atoms be taken away, or one new one added, it is no longer the same mass or the same body. In the state of living creatures, their identity depends not on a mass of the same particles, but on something else. For in them the variation of great parcels of matter alters not the identity: an oak growing from a plant to a great tree, and then lopped, is still the same oak; and a colt grown up to a horse, sometimes fat, sometimes lean, is all the while the same horse: though, in both these cases, there may be a manifest change of the parts, so that truly they are not either of them the same masses of matter, though they be truly one of them the same oak, and the other the same horse. The reason whereof is that, in these two

cases—a mass of matter and a living body, identity is not applied to the same thing.

5. **Identity of Vegetables.** We must therefore consider wherein an oak differs from a mass of matter, and that seems to me to be in this, that the one is only the cohesion of particles of matter anyhow united, the other such a disposition of them as constitutes the parts of an oak, and such an organization of those parts as is fit to receive and distribute nourishment, so as to continue and frame the wood, bark, and leaves, etc., of an oak, in which consists the vegetable life. That being then one plant which has such an organization of parts in one coherent body, partaking of one common life, it continues to be the same plant as long as it partakes of the same life, though that life be communicated to new particles of matter vitally united to the living plant, in a like continued organization conformable to that sort of plants. For this organization, being at any one instant in any one collection of matter, is in that particular concrete distinguished from all other, and is that individual life, which existing constantly from that moment both forwards and backwards, in the same continuity of insensibly succeeding parts united to the living body of the plant, it has that identity which makes the same plant, and all the parts of it parts of the same plant, during all the time that they exist united in that continued organization which is fit to convey that common life to all the parts so united.

6. Identity of Animals. The case is not so much different in brutes but that anyone may hence see what makes an animal and continues it the same. Something we have like this in machines, and may serve to illustrate it. For example, what is a watch? It is plain it is nothing but a fit organization or construction of parts to a certain end, which, when a sufficient force is added to it, it is capable to attain. If we would suppose this machine one continued body, all whose organized parts were repaired, increased, or diminished by a constant addition or separation of insensible parts, with one common life, we should have something very much like the body of an animal—with this difference: that in an animal, the fitness of the organization and the motion wherein life consists begin together, the motion coming from within; but in machines, the force coming sensibly from without is often away when the organ is in order, and well fitted to receive it.

7. The Identity of Man. This also shows wherein the identity of the same man consists: viz., in nothing but a participation of the same continued life, by constantly fleeting particles of matter, in succession vitally united to the same organized body. He that shall place the identity of man in anything else but, like that of other animals, in one fitly organized body taken in any one instant, and from thence continued under one organization of life in several successively fleeting particles of matter united to it, will find it hard to make an embryo, one of years, mad and sober the same man by any supposition that will not make it possible for Seth, Ishmael, Socrates, Pilate, St Austin [St Augustine], and Caesar Borgia to be the same man. For if the identity of soul alone makes the same man, and there be nothing in the nature of matter why the same individual spirit may not be united to different bodies, it will be possible that those men, living in distant ages and of different tempers, may have been the same man: which way of speaking must be from a very strange use of the word man, applied to an idea out of which body and shape are excluded. And that way of speaking would agree yet worse with the notions of those philosophers who allow of transmigration, and are of opinion that the souls

of men may, for their miscarriages, be detrued into the bodies of beasts, as fit habitations, with organs suited to the satisfaction of their brutal inclinations. But yet I think nobody could be sure that the soul of Heliogabalus were in one of his hogs, would yet say that hog were a man or Heliogabalus.

8. Idea of Identity Suited to the Idea It Is Applied To. It is not therefore unity of substance that comprehends all sorts of identity, or will determine it in every case; but to conceive and judge of it aright, we must consider what idea the word it is applied to stands for: it being one thing to be the same substance, another the same man, and a third the same person, if person, man, and substance are three names standing for three different ideas;—for such as is the idea belonging to that name, such must be the identity; which, if it had been a little more carefully attended to, would possibly have prevented a great deal of that confusion which often occurs about this matter, with no small seeming difficulties, especially concerning personal identity, which therefore we shall in the next place a little consider. [. . .]

11. Personal Identity. This being premised, to find wherein personal identity consists, we must consider what person stands for;—which, I think, is a thinking intelligent being, that has reason and reflection, and can consider itself as itself, the same thinking thing, in different times and places; which it does only by that consciousness which is inseparable from thinking, and, as it seems to me, essential to it: it being impossible for anyone to perceive without perceiving that he does perceive. When we see, hear, smell, taste, feel, meditate, or will anything, we know that we do so. Thus it is always as to our present sensations and perceptions: and by this everyone is to himself that which he calls self:—it not being considered, in this case, whether the same self be continued in the same or diverse substances. For, since consciousness always accompanies thinking, and it is that which makes everyone to be what he calls self, and thereby distinguishes himself from all other thinking things, in this alone consists personal identity, i.e., the sameness of a rational being: and as far as this consciousness can be extended backwards to any past action or thought, so far reaches

the identity of that person; it is the same self now it was then; and it is by the same self with this present one that now reflects on it that that action was done.

12. Consciousness Makes Personal Identity. But it is further inquired, whether it be the same identical substance. This few would think they had reason to doubt of, if these perceptions, with their consciousness, always remained present in the mind, whereby the same thinking thing would be always consciously present, and, as would be thought, evidently the same to itself. But that which seems to make the difficulty is this, that this consciousness being interrupted always by forgetfulness, there being no moment of our lives wherein we have the whole train of all our past actions before our eyes in one view, but even the best memories losing the sight of one part whilst they are viewing another; and we sometimes, and that the greatest part of our lives, not reflecting on our past selves, being intent on our present thoughts, and in sound sleep having no thoughts at all, or at least none with that consciousness which remarks our waking thoughts,—I say, in all these cases, our consciousness being interrupted, and we losing the sight of our past selves, doubts are raised whether we are the same thinking thing, i.e., the same substance or no. Which, however reasonable or unreasonable, concerns not personal identity at all, the question being what makes the same person, and not whether it be the same identical substance which always thinks in the same person, which, in this case, matters not at all: different substances, by the same consciousness (where they do partake in it) being united into one person, as well as different bodies by the same life are united into one animal, whose identity is preserved in that change of substances by the unity of one continued life. For, it being the same consciousness that makes a man be himself to himself, personal identity depends on that only, whether it be annexed solely to one individual substance, or can be continued in a succession of several substances. For as far as any intelligent being can repeat the idea of any past action with the same consciousness it had of it at first, and with the same consciousness it has of any present action, so far it is the same personal self. For it is by the consciousness

it has of its present thoughts and actions that it is self to itself now, and so will be the same self as far as the same consciousness can extend to actions past or to come, and would be, by distance of time or change of substance, no more two persons than a man be two men by wearing other clothes today than he did yesterday, with a long or a short sleep between: the same consciousness uniting those distant actions into the same person, whatever substances contributed to their production.

13. Personal Identity in Change of Substance. That this is so, we have some kind of evidence in our very bodies: all whose particles, whilst vitally united to this same thinking conscious self, so that we feel when they are touched, and are affected by, and conscious of good or harm that happens to them, are a part of ourselves—i.e., of our thinking conscious self. Thus, the limbs of his body are to everyone a part of himself; he sympathizes and is concerned for them. Cut off a hand, and thereby separate it from that consciousness he had of its heat, cold, and other affections, and it is then no longer a part of that which is himself, anymore than the remotest part of matter. Thus, we see the substance whereof personal self consisted at one time may be varied at another, without the change of personal identity, there being no question about the same person, though the limbs which but now were a part of it be cut off.

14. Personality in Change of Substance. But the question is, whether if the same substance which thinks be changed, it can be the same person, or, remaining the same, it can be different persons?

And to this I answer: First, this can be no question at all to those who place thought in a purely material animal constitution, void of an immaterial substance. For, whether their supposition be true or no, it is plain they conceive personal identity preserved in something else than identity of substance, as animal identity is preserved in identity of life, and not of substance.

15. Whether in Change of Thinking Substances There Can Be One Person. But next, as to the first part of the question, whether, if the same thinking substance (supposing immaterial substances only to think) be changed, it can be the same person?

I answer, that cannot be resolved but by those who know what kind of substances they are that do think, and whether the consciousness of past actions can be transferred from one thinking substance to another. I grant were the same consciousness the same individual action it could not: but it being a present representation of a past action, why it may not be possible that that may be represented to the mind to have been which really never was will remain to be shown. [...] Why one intellectual substance may not have represented to it, as done by itself, what it never did, and was perhaps done by some other agent—why, I say, such a representation may not possibly be without reality of matter of fact as well as several representations in dreams are, which yet whilst dreaming we take for true—will be difficult to conclude from the nature of things. And that it never is so will by us, till we have clearer views of the nature of thinking substances, be best resolved into the goodness of God, who, as far as the happiness or misery of any of his sensible creatures is concerned in it, will not, by a fatal error of theirs, transfer from one to another that consciousness which draws reward or punishment with it. [...] But yet, [...] it must be allowed that if the same consciousness [...] can be transferred from one thinking substance to another, it will be possible that two thinking substances may make but one person. For the same consciousness being preserved, whether in the same or different substances, the personal identity is preserved.

16. Whether, the Same Immortal Substance Remaining, There Can Be Two Persons. As to the second part of the question, whether the same immortal substance remaining there may be two distinct persons, which question seems to me to be built on this,—whether the same immortal being, being conscious of the action of its past duration, may be wholly stripped of all the consciousness of its past existence, and lose it beyond the power of ever retrieving it again: and so, as it were beginning a new account from a new period, have a consciousness that cannot reach beyond this new state. All those who hold pre-existence are evidently of this mind, since they allow the soul to have no remaining consciousness of what it did in that pre-existent

state, either wholly separate from body or informing any other body; and if they should not, it is plain experience would be against them. [...] Let anyone reflect upon himself and conclude that he has in himself an immaterial spirit, which is that which thinks in him and in the constant change of his body keeps him the same, and is that which he calls himself: let him also suppose it to be the same soul that was in Nestor or Thersites at the siege of Troy (for souls being, as far as we know anything of them, in their nature indifferent to any parcel of matter, the supposition has no apparent absurdity in it), which it may have been as well as it is now the soul of any other man: but he now having no consciousness of any of the actions either of Nestor or Thersites, does or can he conceive himself the same person with either of them? Can he be concerned in either of their actions? Attribute them to himself, or think them his own, more than the actions of any other men that ever existed? So that this consciousness, not reaching to any of the actions of either of those men, he is no more one self with either of them than if the soul or immaterial spirit that now informs him had been created and began to exist when it began to inform his present body, though it were never so true that the same spirit that informed Nestor's or Thersites' body were numerically the same that now informs his. For this would no more make him the same person with Nestor than if some of the particles of matter that were once a part of Nestor were now a part of this man—the same immaterial substance, without the same consciousness, no more making the same person by being united to any body than the same particle of matter, without consciousness, united to any body makes the same person. But let him once find himself conscious of any of the actions of Nestor, he then finds himself the same person with Nestor.

17. The Body, as well as the Soul, Goes to the Making of a Man. And thus may we be able, without any difficulty, to conceive the same person at the resurrection, though in a body not exactly in make or parts the same which he had here,—the same consciousness going along with the soul that inhabits it. But yet the soul alone in the change of

bodies would scarce, to anyone but to him that makes the soul the man, be enough to make the same man. For should the soul of a prince, carrying with it the consciousness of the prince's past life, enter and inform the body of a cobbler, as soon as deserted by his own soul, everyone sees he would be the same person with the prince, accountable only for the prince's actions: but who would say it was the same man? The body too goes to the making the man, and would, I guess, to everybody determine the man in this case, wherein the soul, with all its princely thoughts about it, would not make another man: but he would be the same cobbler to everyone besides himself. I know that in the ordinary way of speaking, the same person and the same man stand for one and the same thing. And indeed everyone will always have a liberty to speak as he pleases, and to apply what articulate sounds to what ideas he thinks fit, and change them as often as he pleases. But yet, when we will inquire what makes the same spirit, man, or person, we must fix the ideas of spirit, man, or person in our minds; and having resolved with ourselves what we mean by them, it will not be hard to determine, in either of them, or the like, when it is the same, and when not.

18. Consciousness Alone Unites Actions into the Same Person. But though the same immaterial substance or soul does not alone, wherever it be, and in whatsoever state, make the same man; yet it is plain, consciousness, as far as ever it can be extended—should it be to ages past—unites existences and actions very remote in time into the same person, as well as it does the existences and actions of the immediately preceding moment: so that whatever has the consciousness of present and past actions is the same person to whom they both belong. Had I the same consciousness that I saw the ark and Noah's flood, as that I saw an overflowing of the Thames last winter, or as that I write now, I could no more doubt that I who write this now, that saw the Thames overflowed last winter, and that viewed the flood at the general deluge was the same self—place that self in what substance you please—than that I who write this am the same myself now whilst I write (whether I consist of all the same substance,

material or immaterial, or no) that I was yesterday. For as to this point of being the same self, it matters not whether this present self be made up of the same or other substances—I being as much concerned, and as justly accountable for any action that was done a thousand years since, appropriated to me now by this self-consciousness, as I am for what I did the last moment.

19. Self Depends on Consciousness, Not on Substance. Self is that conscious thinking thing—whatever substance made up of (whether spiritual or material, simple or compounded, it matters not)—which is sensible or conscious of pleasure and pain, capable of happiness or misery, and so is concerned for itself, as far as that consciousness extends. Thus everyone finds that, whilst comprehended under that consciousness, the little finger is as much a part of himself as what is most so. Upon separation of this little finger, should this consciousness go along with the little finger and leave the rest of the body, it is evident the little finger would be the person, the same person; and self then would have nothing to do with the rest of the body. As in this case it is the consciousness that goes along with the substance when one part is separate from another which makes the same person and constitutes this inseparable self, so it is in reference to substances remote in time. That with which the consciousness of this present thinking thing can join itself makes the same person, and is one self with it and with nothing else, and so attributes to itself, and owns all the actions of that thing, as its own, as far as that consciousness reaches, and no further—as everyone who reflects will perceive.

20. Persons, Not Substances, the Objects of Reward and Punishment. In this personal identity is founded all the right and justice of reward and punishment, happiness and misery being that for which everyone is concerned for himself, and not mattering what becomes of any substance not joined to, or affected with that consciousness. For, as it is evident in the instance I gave but now, if the consciousness went along with the little finger when it was cut off, that would be the same self which was concerned for the whole body yesterday, as making part of itself, whose actions then it cannot but admit

as its own now. Though, if the same body should still live, and immediately from the separation of the little finger have its own peculiar consciousness, whereof the little finger knew nothing, it would not at all be concerned for it, as a part of itself, or could own any of its actions, or have any of them imputed to him.

21. Which Shows Wherein Personal Identity Consists. This may show us wherein personal identity consists: not in the identity of substance, but, as I have said, in the identity of consciousness, wherein if Socrates and the present mayor of Queenborough agree, they are the same person: if the same Socrates waking and sleeping do not partake of the same consciousness, Socrates waking and sleeping is not the same person. And to punish Socrates waking for what sleeping Socrates thought, and waking Socrates was never conscious of, would be no more of right than to punish one twin for what his brother-twin did, whereof he knew nothing, because their outsides were so like that they could not be distinguished, for such twins have been seen.

22. Absolute Oblivion Separates What Is Thus Forgotten from the Person, But Not from the Man. But yet possibly it will still be objected,—Suppose

I wholly lose the memory of some parts of my life, beyond a possibility of retrieving them, so that perhaps I shall never be conscious of them again; yet am I not the same person that did those actions, had those thoughts that I once was conscious of, though I have now forgot them? To which I answer, that we must here take notice what the word I is applied to; which, in this case, is the man only. And the same man being presumed to be the same person, I is easily here supposed to stand also for the same person. But if it be possible for the same man to have distinct incommunicable consciousness at different times, it is past doubt the same man would at different times make different persons; which, we see, is the sense of mankind in the solemnest declaration of their opinions, human laws not punishing the mad man for the sober man's actions, nor the sober man for what the mad man did,—thereby making them two persons: which is somewhat explained by our way of speaking in English when we say such an one is "not himself," or is "beside himself"; in which phrases it is insinuated, as if those who now, or at least first used them, thought that self was changed; the selfsame person was no longer in that man. [. . .]

Study Questions

1. Describe Locke's criteria for being the same material substance, the same plant or animal, and the same person.
2. What are Locke's criteria for being *a human* and for being *a person*? In Locke's view, are all humans persons? In Locke's view, are all persons humans?
3. Outline the arguments Locke offers to justify the

independence of consciousness from any substance, material or immaterial, in which it might inhore.

4. Is Locke correct to claim that the law does not punish "the mad man for the sober man's actions, nor the sober man for what the mad man did"?

1128

Oliver Sacks

Oliver Sacks is an Oxford-educated neurologist whose books tell real-life stories of patients with neurological disorders. By substituting clinical tales for case studies, Sacks wants readers to see the person behind the disorder. But what is a person? This chapter from *The Man Who Mistook His Wife for a Hat* questions the role of memory in personal identity.—TR

The Lost Mariner

From *The Man Who Mistook His Wife for a Hat*

You have to begin to lose your memory, if only in bits and pieces, to realise that memory is what makes our lives. Life without memory is no life at all. . . . Our memory is our coherence, our reason, our feeling, even our action. Without it, we are nothing . . . (I can only wait for the final amnesia, the one that can erase an entire life, as it did my mother's . . .)

—Luis Buñuel

This moving and frightening segment in Buñuel's recently translated memoirs raises fundamental questions—clinical, practical, existential, philosophical: what sort of a life (if any), what sort of a world, what sort of a self, can be preserved in a man who has lost the greater part of his memory and, with this, his past, and his moorings in time?

It immediately made me think of a patient of mine in whom these questions are precisely exemplified: charming, intelligent, memoryless Jimmie G., who was admitted to our Home for the Aged near New York City early in 1975, with a cryptic transfer note saying, "Helpless, demented, confused and disoriented."

Jimmie was a fine-looking man, with a curly bush of grey hair, a healthy and handsome forty-nine-year-old. He was cheerful, friendly, and warm.

"Hiya, Doc!" he said. "Nice morning! Do I take this chair here?" He was a genial soul, very ready to talk and to answer any questions I asked him. He told me his name and birth date, and the name of the little town in Connecticut where he was born. He described it in affectionate detail, even drew me a map. He spoke of the houses where his family had

lived—he remembered their phone numbers still. He spoke of school and school days, the friends he'd had, and his special fondness for mathematics and science. He talked with enthusiasm of his days in the navy—he was seventeen, had just graduated from high school when he was drafted in 1943. With his good engineering mind he was a "natural" for radio and electronics, and after a crash course in Texas found himself assistant radio operator on a submarine. He remembered the names of various submarines on which he had served, their missions, where they were stationed, the names of his shipmates. He remembered Morse code, and was still fluent in Morse tapping and touch-typing.

A full and interesting early life, remembered vividly, in detail, with affection. But there, for some reason, his reminiscences stopped. He recalled, and almost relived, his war days and service, the end of the war, and his thoughts for the future. He had come to love the navy, thought he might stay in it. But with the GI Bill, and support, he felt he might do best to go to college. His older brother was in accountancy school and engaged to a girl, a "real beauty," from Oregon.

With recalling, reliving, Jimmie was full of animation; he did not seem to be speaking of the past but of the present, and I was very struck by the change of tense in his recollections as he passed from his school days to his days in the navy. He had been using the past tense, but now used the present—and (it seemed to me) not just the formal or fictitious present tense of recall, but the actual present tense of immediate experience.

A sudden, improbable suspicion seized me.

"What year is this, Mr G.?" I asked, concealing my perplexity under a casual manner.

"Forty-five, man. What do you mean?" He went on, "We've won the war, FDR's dead, Truman's at the helm. There are great times ahead."

"And you, Jimmie, how old would you be?"

Oddly, uncertainly, he hesitated a moment, as if engaged in calculation.

"Why, I guess I'm 19, Doc. I'll be 20 next birthday."

Looking at the grey-haired man before me, I had an impulse for which I have never forgiven myself—it was, or would have been, the height of cruelty had there been any possibility of Jimmie's remembering it.

"Here," I said, and thrust a mirror toward him. "Look in the mirror and tell me what you see. Is that a 19-year-old looking out from the mirror?"

He suddenly turned ashen and gripped the sides of the chair. "Jesus Christ," he whispered. "Christ, what's going on? What's happened to me? Is this a nightmare? Am I crazy? Is this a joke?"—and he became frantic, panicked.

"It's okay, Jimmie," I said soothingly. "It's just a mistake. Nothing to worry about. Hey!" I took him to the window. "Isn't this a lovely spring day. See the kids there playing baseball?" He regained his colour and started to smile, and I stole away, taking the hateful mirror with me.

Two minutes later I re-entered the room. Jimmie was still standing by the window, gazing with pleasure at the kids playing baseball below. He wheeled around as I opened the door, and his face assumed a cheery expression.

"Hiya, Doc!" he said. "Nice morning! You want to talk to me—do I take this chair here?" There was no sign of recognition on his frank, open face.

"Haven't we met before, Mr G.?" I asked casually.

"No, I can't say we have. Quite a beard you got there. I wouldn't forget you, Doc!"

"Why do you call me "Doc"?"

"Well, you are a doc, ain't you?"

"Yes, but if you haven't met me, how do you know what I am?"

"You talk like a doc. I can see you're a doc."

"Well, you're right, I am. I'm the neurologist here."

"Neurologist? Hey, there's something wrong with my nerves? And 'here'—where's 'here'? What is this place anyhow?"

"I was just going to ask you—where do you think you are?"

"I see these beds, and these patients everywhere. Looks like a sort of hospital to me. But hell, what would I be doing in a hospital—and with all these old people, years older than me. I feel good, I'm strong as a bull. Maybe I work here. . . . Do I work? What's my job? . . . No, you're shaking your head, I see in your eyes I don't work here. If I don't work here, I've been put here. Am I a patient, am I sick and don't know it, Doc? It's crazy, it's scary. . . . Is it some sort of joke?"

"You don't know what the matter is? You really don't know? You remember telling me about your childhood, growing up in Connecticut, working as a radio operator on submarines? And how your brother is engaged to a girl from Oregon?"

"Hey, you're right. But I didn't tell you that, I never met you before in my life. You must have read all about me in my chart."

"Okay," I said. "I'll tell you a story. A man went to his doctor complaining of memory lapses. The doctor asked him some routine questions, and then said, "These lapses. What about them?" "What lapses?" the patient replied."

"So that's my problem," Jimmie laughed. "I kinda thought it was. I do find myself forgetting things, once in a while—things that have just happened. The past is clear, though."

"Will you allow me to examine you, to run over some tests?"

"Sure," he said genially. "Whatever you want."

On intelligence testing he showed excellent ability. He was quick-witted, observant, and logical, and had no difficulty solving complex problems and puzzles—no difficulty, that is, if they could be done quickly. If much time was required, he forgot what he was doing. He was quick and good at tic-tac-toe and checkers, and cunning and aggressive—he easily beat me. But he got lost at chess—the moves were too slow.

Homing in on his memory, I found an extreme and extraordinary loss of recent memory—so that whatever was said or shown to him was apt to be forgotten in a few seconds' time. Thus I laid out my watch, my tie, and my glasses on the desk, covered them, and asked him to remember these. Then, after a minute's chat, I asked him what I had put under the cover. He remembered none of them—or, indeed, that I had even asked him to remember. I repeated the test, this time getting him to write down the names of the three objects; again he forgot, and when I showed him the paper with his writing on it he was astounded, and said he had no recollection of writing anything down, though he acknowledged that it was his own writing, and then got a faint "echo" of the fact that he had written them down.

He sometimes retained faint memories, some dim echo or sense of familiarity. Thus five minutes after I had played tic-tac-toe with him, he recollects that "some doctor" had played this with him "a while back"—whether the "while back" was minutes or months ago he had no idea. He then paused and said, "It could have been you?" When I said it was me, he seemed amused. This faint amusement and indifference were very characteristic, as were the involved cogitations to which he was driven by being so disoriented and lost in time. When I asked Jimmie the time of the year, he would immediately look around for some clue—I was careful to remove the calendar from my desk—and would work out the time of year, roughly, by looking through the window.

It was not, apparently, that he failed to register in memory, but that the memory traces were fugitive in the extreme, and were apt to be effaced within a minute, often less, especially if there were distracting or competing stimuli, while his intellectual and perceptual powers were preserved, and highly superior.

Jimmie's scientific knowledge was that of a bright high school graduate with a penchant for mathematics and science. He was superb at arithmetical (and also algebraic) calculations, but only if they could be done with lightning speed. If there were many steps, too much time, involved, he would forget where he was, and even the question. He knew the elements,

compared them, and drew the periodic table—but omitted the transuranic elements.

"Is that complete?" I asked when he'd finished.

"It's complete and up-to-date, sir, as far as I know."

"You wouldn't know any elements beyond uranium?"

"You kidding? There's 92 elements, and uranium's the last."

I paused and flipped through a *National Geographic* on the table. "Tell me the planets," I said, "and something about them." Unhesitatingly, confidently, he gave me the planets—their names, their discovery, their distance from the sun, their estimated mass, character, and gravity.

"What is this?" I asked, showing him a photo in the magazine I was holding.

"It's the moon," he replied.

"No, it's not," I answered. "It's a picture of the earth taken from the moon."

"Doc, you're kidding! Someone would've had to get a camera up there!"

"Naturally."

"Hell! You're joking—how the hell would you do that?"

Unless he was a consummate actor, a fraud simulating an astonishment he did not feel, this was an utterly convincing demonstration that he was still in the past. His words, his feelings, his innocent wonder, his struggle to make sense of what he saw, were precisely those of an intelligent young man in the forties faced with the future, with what had not yet happened, and what was scarcely imaginable. "This more than anything else," I wrote in my notes, "persuades me that his cut-off around 1945 is genuine. . . . What I showed him, and told him, produced the authentic amazement which it would have done in an intelligent young man of the pre-Sputnik era."

I found another photo in the magazine and pushed it over to him.

"That's an aircraft carrier," he said. "Real ultra-modern design. I never saw one quite like that."

"What's it called?" I asked.

He glanced down, looked baffled, and said, "The *Nimitz*!"

"Something the matter?"

"The hell there is!" he replied hotly. "I know 'em all by name, and I don't know a Nimitz. . . . Of course there's an Admiral Nimitz, but I never heard they named a carrier after him."

Angrily he threw the magazine down.

He was becoming fatigued, and somewhat irritable and anxious, under the continuing pressure of anomaly and contradiction, and their fearful implications, to which he could not be entirely oblivious. I had already, unthinkingly, pushed him into panic, and felt it was time to end our session. We wandered over to the window again, and looked down at the sunlit baseball diamond; as he looked his face relaxed, he forgot the Nimitz, the satellite photo, the other horrors and hints, and became absorbed in the game below. Then, as a savoury smell drifted up from the dining room, he smacked his lips, said "Lunch!", smiled, and took his leave.

And I myself was wrung with emotion—it was heartbreaking, it was absurd, it was deeply perplexing, to think of his life lost in limbo, dissolving.

"He is, as it were," I wrote in my notes, "isolated in a single moment of being, with a moat or lacuna of forgetting all round him. . . . He is man without a past (or future), stuck in a constantly changing, meaningless moment." And then, more prosaically, "The remainder of the neurological examination is entirely normal. Impression: probably Korsakoff's syndrome, due to alcoholic degeneration of the mammillary bodies." My note was a strange mixture of facts and observations, carefully noted and itemized, with irrepressible meditations on what such problems might "mean," in regard to who and what and where this poor man was—whether, indeed, one could speak of an "existence," given so absolute a privation of memory or continuity.

I kept wondering, in this and later notes—unscientifically—about "a lost soul," and how one might establish some continuity, some roots, for he was a man without roots, or rooted only in the remote past.

"Only connect"—but how could he connect, and how could we help him to connect? What was life without connection? "I may venture to affirm,"

Hume wrote, "that we are nothing but a bundle or collection of different sensations, which succeed each other with an inconceivable rapidity, and are in a perpetual flux and movement." In some sense, he had been reduced to a "Humean" being—I could not help thinking how fascinated Hume would have been at seeing in Jimmie his own philosophical "chimaera" incarnate, a gruesome reduction of a man to mere disconnected, incoherent flux and change.

Perhaps I could find advice or help in the medical literature—a literature which, for some reason, was largely Russian, from Korsakoff's original thesis about such cases of memory loss, which are still called "Korsakoff's syndrome," to Luria's *Neuropsychology of Memory* (which appeared in translation only a year after I first saw Jimmie). Korsakoff wrote in 1887: "Memory of recent events is disturbed almost exclusively; recent impressions apparently disappear soonest, whereas impressions of long ago are recalled properly, so that the patient's ingenuity, his sharpness of wit, and his resourcefulness remain largely unaffected."

To Korsakoff's brilliant but spare observations, almost a century of further research has been added—the richest and deepest, by far, being Luria's. And in Luria's account science became poetry, and the pathos of radical lostness was evoked. "Gross disturbances of the organization of impressions of events and their sequence in time can always be observed in such patients," he wrote. "In consequence, they lose their integral experience of time and begin to live in a world of isolated impressions." Further, as Luria noted, the eradication of impressions (and their disorder) might spread backward in time—"in the most serious cases—even to relatively distant events."

Most of Luria's patients, as described in this book, had massive and serious cerebral tumours, which had the same effects as Korsakoff's syndrome, but later spread and were often fatal. Luria included no cases of "simple" Korsakoff's syndrome, based on the self-limiting destruction that Korsakoff described—neuron destruction, produced by alcohol, in the tiny but crucial mammillary bodies, the rest of the brain being preserved. And so there was no long-term follow-up of Luria's cases.

I had at first been deeply puzzled, and dubious, even suspicious, about the apparently sharp cut-off in 1945, a point, a date, which was also symbolically so sharp. I wrote in a subsequent note:

There is a great blank. We do not know what happened then—or subsequently. . . . We must fill in these "missing" years—from his brother, or the navy, or hospitals he has been to. . . . Could it be that he sustained some massive trauma at this time, some massive cerebral or emotional trauma in combat, in the war, and that this may have affected him ever since? . . . was the war his "high point," the last time he was really alive, and existence since one long anti-climax?

We did various tests on him (EEG, brain scans), and found no evidence of massive brain damage, although atrophy of the tiny mammillary bodies would not show up on such tests. We received reports from the navy indicating that he had remained in the navy until 1965, and that he was perfectly competent at that time.

Then we turned up a short nasty report from Bellevue Hospital, dated 1971, saying that he was "totally disoriented . . . with an advanced organic brain-syndrome, due to alcohol" (cirrhosis had also developed by this time). From Bellevue he was sent to a wretched dump in the Village, a so-called "nursing home" whence he was rescued—lousy, starving—by our Home in 1975.

We located his brother, whom Jimmie always spoke of as being in accountancy school and engaged to a girl from Oregon. In fact he had married the girl from Oregon, had become a father and grandfather, and been a practising accountant for 30 years.

Where we had hoped for an abundance of information and feeling from his brother, we received a courteous but somewhat meagre letter. It was obvious from reading this—especially reading between the lines—that the brothers had scarcely seen each other since 1943, and gone separate ways, partly through the vicissitudes of location and profession, and partly through deep (though not estranging)

differences of temperament. Jimmie, it seemed, had never "settled down," was "happy-go-lucky," and "always a drinker." The navy, his brother felt, provided a structure, a life, and the real problems started when he left it, in 1965. Without his habitual structure and anchor Jimmie had ceased to work, "gone to pieces," and started to drink heavily. There had been some memory impairment, of the Korsakoff type, in the middle and especially the late sixties, but not so severe that Jimmie couldn't "cope" in his nonchalant fashion. But his drinking grew heavier in 1970.

Around Christmas of that year, his brother understood, he had suddenly "blown his top" and become deliriously excited and confused, and it was at this point he had been taken into Bellevue. During the next month, the excitement and delirium died down, but he was left with deep and bizarre memory lapses, or "deficits," to use the medical jargon. His brother had visited him at this time—they had not met for 20 years—and, to his horror, Jimmie not only failed to recognize him, but said, "Stop joking! You're old enough to be my father. My brother's a young man, just going through accountancy school."

When I received this information, I was more perplexed still: why did Jimmie not remember his later years in the navy, why did he not recall and organize his memories until 1970? I had not heard then that such patients might have a retrograde amnesia [. . .]. "I wonder, increasingly," I wrote at this time, "whether there is not an element of hysterical or fugal amnesia—whether he is not in flight from something too awful to recall," and I suggested he be seen by our psychiatrist. Her report was searching and detailed—the examination had included a sodium amyta test, calculated to "release" any memories which might be repressed. She also attempted to hypnotize Jimmie, in the hope of eliciting memories repressed by hysteria—this tends to work well in cases of hysterical amnesia. But it failed because Jimmie could not be hypnotized, not because of any "resistance," but because of his extreme amnesia, which caused him to lose track of what the hypnotist was saying. [. . .]

"I have no feeling or evidence," the psychiatrist wrote, "of any hysterical or "put-on" deficit. He lacks both the means and the motive to make a facade. His memory deficits are organic and permanent and incorrigible, though it is puzzling they should go back so long." Since, she felt, he was "unconcerned . . . manifested no special anxiety . . . constituted no management problem," there was nothing she could offer, or any therapeutic "entrance" or "lever" she could see.

At this point, persuaded that this was, indeed, "pure" Korsakoff's, uncomplicated by other factors, emotional or organic, I wrote to Luria and asked his opinion. He spoke in his reply of his patient Bel (see Luria 250–2), whose amnesia had retroactively eradicated ten years. He said he saw no reason why such a retrograde amnesia should not thrust backward decades, or almost a whole lifetime. "I can only wait for the final amnesia," Buñuel writes, "the one that can erase an entire life." But Jimmie's amnesia, for whatever reason, had erased memory and time back to 1945—roughly—and then stopped. Occasionally, he would recall something much later, but the recall was fragmentary and dislocated in time. Once, seeing the word "satellite" in a newspaper headline, he said offhandedly that he'd been involved in a project of satellite tracking while on the ship *Chesapeake Bay*, a memory fragment coming from the early or mid-sixties. But, for all practical purposes, his cut-off point was during the mid- (or late) forties, and anything subsequently retrieved was fragmentary, unconnected. This was the case in 1975, and it is still the case now, nine years later.

What could we do? What should we do? "There are no prescriptions," Luria wrote, "in a case like this. Do whatever your ingenuity and your heart suggest. There is little or no hope of any recovery in his memory. But a man does not consist of memory alone. He has feeling, will, sensibilities, moral being—matters of which neuropsychology cannot speak. And it is here, beyond the realm of an impersonal psychology, that you may find ways to touch him, and change him. And the circumstances of your work especially allow this, for you work in a Home, which is like a little world, quite different from the clinics and

institutions where I work. Neuropsychologically, there is little or nothing you can do; but in the realm of the individual, there may be much you can do."

Luria mentioned his patient Kur as manifesting a rare self-awareness, in which hopelessness was mixed with an odd equanimity. "I have no memory of the present," Kur would say. "I do not know what I have just done or from where I have just come. . . . I can recall my past very well, but I have no memory of my present." When asked whether he had ever seen the person testing him, he said, "I cannot say yes or no, I can neither affirm nor deny that I have seen you." This was sometimes the case with Jimmie; and, like Kur, who stayed many months in the same hospital, Jimmie began to form a sense of familiarity; he slowly learned his way around the home—the whereabouts of the dining room, his own room, the elevators, the stairs, and in some sense recognized some of the staff, although he confused them, and perhaps had to do so, with people from the past. He soon became fond of the nursing sister in the Home; he recognized her voice, her footfalls, immediately, but would always say that she had been a fellow pupil at his high school, and was greatly surprised when I addressed her as "Sister."

"Gee!" he exclaimed, "the damnedest things happen. I'd never have guessed you'd become a religious, Sister!"

Since he's been at our Home—that is, since early 1975—Jimmie has never been able to identify anyone in it consistently. The only person he truly recognizes is his brother, whenever he visits from Oregon. These meetings are deeply emotional and moving to observe—the only truly emotional meetings Jimmie has. He loves his brother, he recognizes him, but he cannot understand why he looks so old: "Guess some people age fast," he says. Actually his brother looks much younger than his age, and has the sort of face and build that change little with the years. These are true meetings, Jimmie's only connection of past and present, yet they do nothing to provide any sense of history or continuity. If anything they emphasize—at least to his brother, and to others who see them together—that Jimmie still lives, is fossilized, in the past.

All of us, at first, had high hopes of helping Jimmie—he was so personable, so likable, so quick and intelligent, it was difficult to believe that he might be beyond help. But none of us had ever encountered, even imagined, such a power of amnesia, the possibility of a pit into which everything, every experience, every event, would fathomlessly drop, a bottomless memory-hole that would engulf the whole world.

I suggested, when I first saw him, that he should keep a diary, and be encouraged to keep notes every day of his experiences, his feelings, thoughts, memories, reflections. These attempts were foiled, at first, by his continually losing the diary: it had to be attached to him—somehow. But this too failed to work: he dutifully kept a brief daily notebook but could not recognize his earlier entries in it. He does recognize his own writing, and style, and is always astounded to find that he wrote something the day before.

Astounded—and indifferent—for he was a man who, in effect, had no “day before.” His entries remained unconnected and unconnecting and had no power to provide any sense of time or continuity. Moreover, they were trivial—“Eggs for breakfast,” “Watched ballgame on TV”—and never touched the depths. But were there depths in this unmemoried man, depths of an abiding feeling and thinking, or had he been reduced to a sort of Humean drivel, a mere succession of unrelated impressions and events?

Jimmie both was and wasn't aware of this deep, tragic loss in himself, loss of himself. (If a man has lost a leg or an eye, he knows he has lost a leg or an eye; but if he has lost a self—himself—he cannot know it, because he is longer there to know it.) Therefore I could not question him intellectually about such matters.

He had originally professed bewilderment at finding himself amid patients, when, as he said, he himself didn't feel ill. But what, we wondered, did he feel? He was strongly built and fit, he had a sort of animal strength and energy, but also a strange inertia, passivity, and (as everyone remarked) “unconcern”; he gave all of us an overwhelming sense of

“something missing,” although this, if he realized it, was itself accepted with an odd “unconcern.” One day I asked him not about his memory, or past, but about the simplest and most elemental feelings of all:

“How do you feel?”

“How do I feel,” he repeated, and scratched his head. “I cannot say I feel ill. But I cannot say I feel well. I cannot say I feel anything at all.”

“Are you miserable?” I continued.

“Can't say I am.”

“Do you enjoy life?”

“I can't say I do . . .”

I hesitated, fearing that I was going too far, that I might be stripping a man down to some hidden, unacknowledgable, unbearable despair.

“You don't enjoy life,” I repeated, hesitating somewhat. “How then do you feel about life?”

“I can't say that I feel anything at all.”

“You feel alive though?”

“Feel alive? Not really. I haven't felt alive for a very long time.”

His face wore a look of infinite sadness and resignation.

Later, having noted his aptitude for, and pleasure in, quick games and puzzles, and their power to “hold” him, at least while they lasted, and to allow, for a while, a sense of companionship and competition—he had not complained of loneliness, but he looked so alone; he never expressed sadness, but he looked so sad—I suggested he be brought into our recreation programs at the Home. This worked better—better than the diary. He would become keenly and briefly involved in games, but soon they ceased to offer any challenge: he solved all the puzzles, and could solve them easily; and he was far better and sharper than anyone else at games. And as he found this out, he grew fretful and restless again, and wandered the corridors, uneasy and bored and with a sense of indignity—games and puzzles were for children, a diversion. Clearly, passionately, he wanted something to do: he wanted to do, to be, to feel—and could not; he wanted sense, he wanted purpose—in Freud's words, “Work and Love.”

Could he do “ordinary” work? He had “gone to pieces,” his brother said, when he ceased to work in 1965.

He had two striking skills—Morse code and touch-typing. We could not use Morse, unless we invented a use; but good typing we could use, if he could recover his old skills—and this would be real work, not just a game. Jimmie soon did recover his old skill and came to type very quickly—he could not do it slowly—and found in this some of the challenge and satisfaction of a job. But still this was superficial tapping and typing; it was trivial, it did not reach to the depths. And what he typed, he typed mechanically—he could not hold the thought—the short sentences following one another in a meaningless order.

One tended to speak of him, instinctively, as a spiritual casualty—a “lost soul”: was it possible that he had really been “de-souled” by a disease? “Do you think he has a soul?” I once asked the Sisters. They were outraged by my question, but could see why I asked it. “Watch Jimmie in chapel,” they said, “and judge for yourself.”

I did, and I was moved, profoundly moved and impressed, because I saw here an intensity and steadiness of attention and concentration that I had never seen before in him or conceived him capable of. I watched him kneel and take the Sacrament on his tongue, and could not doubt the fullness and totality of Communion, the perfect alignment of his spirit with the spirit of the Mass. Fully, intensely, quietly, in the quietude of absolute concentration and attention, he entered and partook of the Holy Communion. He was wholly held, absorbed, by a feeling. There was no forgetting, no Korsakov’s then, nor did it seem possible or imaginable that there should be; for he was no longer at the mercy of a faulty and fallible mechanism—that of meaningless sequences and memory traces—but was absorbed in an act, an act of his whole being, which carried feeling and meaning in an organic continuity and unity, a continuity and unity so seamless it could not permit any break.

Clearly Jimmie found himself, found continuity and reality, in the absoluteness of spiritual attention and act. The Sisters were right—he did find his soul here. And so was Luria, whose words now came back to me: “A man does not consist of memory alone. He

has feeling, will, sensibility, moral being. . . . It is here . . . you may touch him, and see a profound change.” Memory, mental activity, mind alone, could not hold him; but moral attention and action could hold him completely.

But perhaps “moral” was too narrow a word—for the aesthetic and dramatic were equally involved. Seeing Jim in the chapel opened my eyes to other realms where the soul is called on, and held, and stilled, in attention and communion. The same depth of absorption and attention was to be seen in relation to music and art: he had no difficulty, I noticed, “following” music or simple dramas, for every moment in music and art refers to, contains, other moments. He liked gardening, and had taken over some of the work in our garden. At first he greeted the garden each day as new, but for some reason this had become more familiar to him than the inside of the Home. He almost never got lost or disoriented in the garden now; he patterned it, I think, on loved and remembered gardens from his youth in Connecticut.

Jimmie, who was so lost in extensional “spatial” time, was perfectly organized in Bergsonian “intentional” time; what was fugitive, unsustainable, as formal structure, was perfectly stable, perfectly held, as art or will. Moreover, there was something that endured and survived. If Jimmie was briefly “held” by a task or puzzle or game or calculation, held in the purely mental challenge of these, he would fall apart as soon as they were done, into the abyss of his nothingness, his amnesia. But if he was held in emotional and spiritual attention—in the contemplation of nature or art, in listening to music, in taking part in the Mass in chapel—the attention, its “mood,” its quietude, would persist for a while, and there would be in him a pensiveness and peace we rarely, if ever, saw during the rest of his life at the Home.

I have known Jimmie now for nine years—and neuropsychologically, he has not changed in the least. He still has the severest, most devastating Korsakov’s, cannot remember isolated items for more than a few seconds, and has a dense amnesia going back to 1945. But, humanly, spiritually, he is at times a different man altogether—no longer fluttering, restless, bored, and lost, but deeply attentive to the beauty

and soul of the world, rich in all the Kierkegaardian categories—[the] aesthetic, the moral, the religious, the dramatic. I had wondered, when I first met him, if he was not condemned to a sort of "Humean" froth, a meaningless fluttering on the surface of life, and whether there was any way of transcending the incoherence of his Humean disease. Empirical science told me there was not—but empirical science, empiricism, takes no account of the soul, no account

of what constitutes and determines personal being. Perhaps there is a philosophical as well as a clinical lesson here: that in Korsakov's, or dementia, or other such catastrophes, however great the organic damage and Humean dissolution, there remains the undiminished possibility of reintegration by art, by communion, by touching the human spirit: and this can be preserved in what seems at first a hopeless state of neurological devastation. [. . .]

Work Cited

Luria, A.R. *The Neuropsychology of Memory*. Washington: V.H. Winston, 1976.

Study Questions

1. Apply Locke's criteria for being the same material substance, the same human, and the same person to the example of Sacks's patient Jimmie G.
2. Why does Sacks (at least initially) remark, "I could not help thinking how fascinated Hume would have been at seeing in Jimmie his own philosophical 'chimaera' incarnate"?
3. Explain why, as he works further with Jimmie G., Sacks becomes skeptical that any empirically based account of personal identity manages adequately to capture what it is to be a person.
4. What theory of personal identity does Sacks himself ultimately formulate?

Derek Parfit

1 / 30

Derek Parfit is interested in problems of personal identity, rationality, and ethics. He is a senior research fellow at All Souls College, Oxford, and visiting professor of philosophy at New York University, Harvard University, and Rutgers University. In this well-known essay, Parfit recommends that we give up talk about personal identity.—TR

Personal Identity

1. Fission Thought Experiment

We can start by considering the much-discussed case (implicit in John Locke's *An Essay Concerning Human Understanding*, Ch. 27 s. [20]) of the man who, like an amoeba, divides.

Wiggins has recently dramatized this case (50). He first referred to the operation imagined by Shoemaker (22). We suppose that my brain is transplanted into someone else's (brainless) body, and that the resulting person has my character and apparent memories of my life. Most of us would agree, after thought, that the resulting person is me. I shall here assume such agreement.

Wiggins then imagined his own operation. My brain is divided, and each half is housed in a new body. Both resulting people have my character and apparent memories of my life.

What happens to me? There seem only three possibilities: (1) I do not survive; (2) I survive as one of the two people; (3) I survive as both.

The trouble with (1) is this. We agreed that I could survive if my brain were successfully transplanted. And people have in fact survived with half their brains destroyed. It seems to follow that I could survive if half my brain were successfully transplanted and the other half were destroyed. But if this is so, how could I not survive if the other half were also successfully transplanted? How could a double success be a failure?

We can move to the second description. Perhaps one success is the maximum score. Perhaps I shall be one of the resulting people.

The trouble here is that in Wiggins' case each half of my brain is exactly similar, and so, to start with, is each resulting person. So how can I survive as only one of the two people? What can make me one of them rather than the other?

It seems clear that both of these descriptions—that I do not survive, and that I survive as one of the people—are highly implausible. Those who have accepted them must have assumed that they were the only possible descriptions.

What about our third description: that I survive as both people?

It might be said, "If 'survive' implies identity, this description makes no sense—you cannot be two people. If it does not, the description is irrelevant to a problem about identity."

I shall later deny the second of these remarks. But there are ways of denying the first. We might say, "What we have called 'the two resulting people' are not two people. They are one person. I do survive Wiggins' operation. Its effect is to give me two bodies and a divided mind."

It would shorten my argument if this were absurd. But I do not think it is. It is worth showing why.

We can, I suggest, imagine a divided mind. We can imagine a man having two simultaneous experiences, in having each of which he is unaware of having the other.

We may not even need to imagine this. Certain actual cases, to which Wiggins referred, seem to be best described in these terms. These involve the cutting of the bridge between the hemispheres of the brain. The aim was to cure epilepsy. But the result appears to be, in the surgeon's words, the creation of "two separate spheres of consciousness" (Sperry 299), each of which controls one half of the patient's body. What is experienced in each is, presumably, experienced by the patient.

There are certain complications in these actual cases. So let us imagine a simpler case.

Suppose that the bridge between my hemispheres is brought under my voluntary control. This would enable me to disconnect my hemispheres as easily as if I were blinking. By doing this I would divide my mind. And we can suppose that when my mind is divided I can, in each half, bring about reunion.

This ability would have obvious uses. To give an example: I am near the end of a math exam, and see two ways of tackling the last problem. I decide to divide my mind, to work, with each half, at one of two calculations, and then to reunite my mind and write a fair copy of the best result.

What shall I experience?

When I disconnect my hemispheres, my consciousness divides into two streams. But this division is not something that I experience. Each of my two streams of consciousness seems to have been straightforwardly continuous with my one stream of consciousness up to the moment of division. The only changes in each stream are the disappearance of half my visual field and the loss of sensation in, and control over, half my body.

Consider my experiences in what we can call my "right-handed" stream. I remember that I assigned my right hand to the longer calculation. This I now begin. In working at this calculation I can see, from the movements of my left hand, that I am also working at the other. But I am not aware of working at the other. So I might, in my right-handed stream, wonder how, in my left-handed stream, I am getting on.

My work is now over. I am about to reunite my mind. What should I, in each stream, expect? Simply that I shall suddenly seem to remember just having thought out two calculations, in thinking out each

of which I was not aware of thinking out the other. This, I submit, we can imagine. And if my mind was divided, these memories are correct.

In describing this episode, I assumed that there were two series of thoughts, and that they were both mine. If my two hands visibly wrote out two calculations, and if I claimed to remember two corresponding series of thoughts, this is surely what we should want to say.

If it is, then a person's mental history need not be like a canal, with only one channel. It could be like a river, with islands, and with separate streams.

To apply this to Wiggins' operation: we mentioned the view that it gives me two bodies and a divided mind. We cannot now call this absurd. But it is, I think, unsatisfactory.

There were two features of the case of the exam that made us want to say that only one person was involved. The mind was soon reunited, and there was only one body. If a mind was permanently divided and its halves developed in different ways, the point of speaking of one person would start to disappear. Wiggins' case, where there are also two bodies, seems to be over the borderline. After I have had his operation, the two "products" each have all the attributes of a person. They could live at opposite ends of the earth. (If they later met, they might even fail to recognize each other.) It would become intolerable to deny that they were different people.

Suppose we admit that they are different people. Could we still claim that I survived as both, using "survive" to imply identity?

We could. For we might suggest that two people could compose a third. We might say, "I do survive Wiggins' operation as two people. They can be different people, and yet be me, in just the way in which the Pope's three crowns are one crown."

This is a possible way of giving sense to the claim that I survive as two different people, using "survive" to imply identity. But it keeps the language of identity only by changing the concept of a person. And there are obvious objections to this change.¹

The alternative, for which I shall argue, is to give up the language of identity. We can suggest that I survive as two different people without implying that I am these people.

When I first mentioned this alternative, I mentioned this objection: "If your new way of talking does not imply identity, it cannot solve our problem. For that is about identity. The problem is that all the possible answers to the question about identity are highly implausible."

We can now answer this objection.

We can start by reminding ourselves that this is an objection only if we have one or both of the [following] beliefs [...].

The first [is] the belief that to any question about personal identity, in any describable case, there must be a true answer. For those with this belief, Wiggins' case is doubly perplexing. If all the possible answers are implausible, it is hard to decide which of them is true, and hard even to keep the belief that one of them must be true. If we give up this belief, as I think we should, these problems disappear. We shall then regard the case as like many others in which, for quite unpuzzling reasons, there is no answer to a question about identity. (Consider "Was England the same nation after 1066?")²

Wiggins' case makes the first belief implausible. It also makes it trivial. For it undermines the second belief. This [is] the belief that important questions turn upon the question about identity. (It is worth pointing out that those who have only this second belief do not think that there must be an answer to this question, but rather that we must decide upon an answer.)

Against this second belief my claim is this. Certain questions do presuppose a question about personal identity. And because these questions are important, Wiggins' case does present a problem. But we cannot solve this problem by answering the question about identity. We can solve this problem only by taking these important questions and prizing them apart from the question about identity. After we have done this, the question about identity (though we might for the sake of neatness decide it) has no further interest.

Because there are several questions which presuppose identity, this claim will take some time to fill out.

We can first return to the question of survival. This is a special case, for survival does not

so much presuppose the retaining of identity as seem equivalent to it. It is thus the general relation which we need to prize apart from identity. We can then consider particular relations, such as those involved in memory and intention.

"Will I survive?" seems, I said, equivalent to "Will there be some person alive who is the same person as me?"

If we treat these questions as equivalent, then the least unsatisfactory description of Wiggins' case is, I think, that I survive with two bodies and a divided mind.

Several writers have chosen to say that I am neither of the resulting people. Given our equivalence, this implies that I do not survive, and hence, presumably, that even if Wiggins' operation is not literally death, I ought, since I will not survive it, to regard it as death. But this seemed absurd.

It is worth repeating why An emotion or attitude can be criticized for resting on a false belief, or for being inconsistent. A man who regarded Wiggins' operation as death must, I suggest, be open to one of these criticisms.

He might believe that his relation to each of the resulting people fails to contain some element which is contained in survival. But how can this be true? We agreed that he would survive if he stood in this very same relation to only one of the resulting people. So it cannot be the nature of this relation which makes it fail, in Wiggins' case, to be survival. It can only be its duplication.

Suppose that our man accepts this, but still regards division as death. His reaction would now seem wildly inconsistent. He would be like a man who, when told of a drug that could double his years of life, regarded the taking of this drug as death. The only difference in the case of division is that the extra years are to run concurrently. This is an interesting difference. But it cannot mean that there are no years to run.

I have argued this for those who think that there must, in Wiggins' case, be a true answer to the question about identity. For them, we might add, "Perhaps the original person does lose his identity. But there may be other ways to do this

than to die. One other way might be to multiply. To regard these as the same is to confuse nought with two."

For those who think that the question of identity is up for decision, it would be clearly absurd to regard Wiggins' operation as death. These people would have to think, "We could have chosen to say that I should be one of the resulting people. If we had, I should not have regarded it as death. But since we have chosen to say that I am neither person, I do." This is hard even to understand.

My first conclusion, then, is this. The relation of the original person to each of the resulting people contains all that interests us—all that matters—in any ordinary case of survival. This is why we need a sense in which one person can survive as two,

One of my aims in the rest of this paper will be to suggest such a sense. But we can first make some general remarks.

2. Psychological Continuity and Identity

Identity is a one-one relation. Wiggins' case serves to show that what matters in survival need not be one-one.

Wiggins' case is of course unlikely to occur. The relations which matter are, in fact, one-one. It is because they are that we can imply the holding of these relations by using the language of identity.

This use of language is convenient. But it can lead us astray. We may assume that what matters is identity and, hence, has the properties of identity.

In the case of the property of being one-one, this mistake is not serious. For what matters is in fact one-one. But in the case of another property, the mistake is serious. Identity is all-or-nothing. Most of the relations which matter in survival are, in fact, relations of degree. If we ignore this, we shall be led into quite ill-grounded attitudes and beliefs.

The claim that I have just made—that most of what matters are relations of degree—I have yet to support. Wiggins' case shows only that these relations need not be one-one. The merit of the case is not that it shows this in particular, but that it makes

the first break between what matters and identity.
The belief that identity is what matters is hard to overcome. This is shown in most discussions of the problem cases which actually occur: cases, say, of amnesia or of brain damage. Once Wiggins' case has made one breach in this belief, the rest should be easier to remove.

To turn to a recent debate: most of the relations which matter can be provisionally referred to under the heading "psychological continuity" (which includes causal continuity). My claim is thus that we use the language of personal identity in order to imply such continuity. This is close to the view that psychological continuity provides a criterion of identity.

Williams has attacked this view with the following argument. Identity is a one-one relation. So any criterion of identity must appeal to a relation which is logically one-one. Psychological continuity is not logically one-one. So it cannot provide a criterion.

Some writers have replied that it is enough if the relation appealed to is always in fact one-one.

I suggest a slightly different reply. Psychological continuity is a ground for speaking of identity when it is one-one.

If psychological continuity took a one-many or branching form, we should need, I have argued, to abandon the language of identity. So this possibility would not count against this view.

We can make a stronger claim. This possibility would count in its favour.

The view might be defended as follows. Judgments of personal identity have great importance. What gives them their importance is the fact that they imply psychological continuity. This is why, whenever there is such continuity, we ought, if we can, to imply it by making a judgment of identity.

If psychological continuity took a branching form, no coherent set of judgments of identity could correspond to, and thus be used to imply, the branching form of this relation. But what we ought to do, in such a case, is take the importance which would attach to a judgment of identity and attach this importance directly to each limb of the branching relation. So this case helps to show that judgments

of personal identity do derive their importance from the fact that they imply psychological continuity. It helps to show that when we can, usefully, speak of identity, this relation is our ground. [. . .]

The criterion might be sketched as follows. "X and Y are the same person if they are psychologically continuous and there is no person who is contemporary with either and psychologically continuous with the other." We should need to explain what we mean by "psychologically continuous" and say how much continuity the criterion requires. We should then, I think, have described a sufficient condition for speaking of identity.³

We need to say something more. If we admit that psychological continuity might not be one-one, we need to say what we ought to do if it were not one-one. Otherwise our account would be open to the objections that it is incomplete and arbitrary.

I have suggested that if psychological continuity took a branching form, we ought to speak in a new way, regarding what we describe as having the same significance as identity. This answers these objections.

We can now return to our discussion. We have three remaining aims. One is to suggest a sense of "survive" which does not imply identity. Another is to show that most of what matters in survival are relations of degree. A third is to show that none of these relations needs to be described in a way that presupposes identity.

We can take these aims in the reverse order.

3. Survival Relations and Identity

The most important particular relation is that involved in memory. This is because it is so easy to believe that its description must refer to identity. This belief about memory is an important cause of the view that personal identity has a special nature. But it has been well discussed by Shoemaker and by Wiggins. So we can be brief.

It may be a logical truth that we can only remember our own experiences. But we can frame a new concept for which this is not a logical truth. Let us call this "q-memory."

To sketch a definition I am *q*-remembering an experience if (1) I have a belief about a past experience which seems in itself like a memory belief, (2) someone did have such an experience, and (3) my belief is dependent upon this experience in the same way (whatever that is) in which a memory of an experience is dependent upon it.

According to (1) *q*-memories seem like memories. So I *q*-remember having experiences.

This may seem to make *q*-memory presuppose identity. One might say, "My apparent memory of having an experience is an apparent memory of my having an experience. So how could I *q*-remember my having other people's experiences?"

This objection rests on a mistake. When I seem to remember an experience, I do indeed seem to remember having it. But it cannot be a part of what I seem to remember about this experience that I, the person who now seems to remember it, am the person who had this experience.⁴ That I am is something that I automatically assume. (My apparent memories sometimes come to me simply as the belief that I had a certain experience.) But it is something that I am justified in assuming only because I do not in fact have *q*-memories of other people's experiences.

Suppose that I did start to have such *q*-memories. If I did, I should cease to assume that my apparent memories must be about my own experiences. I should come to assess an apparent memory by asking two questions: (1) Does it tell me about a past experience? (2) If so, whose?

Moreover (and this is a crucial point) my apparent memories would now come to me as *q*-memories. Consider those of my apparent memories which do come to me simply as beliefs about my past: for example, "I did that." If I knew that I could *q*-remember other people's experiences, these beliefs would come to me in a more guarded form: for example, "Someone—probably I—did that." I might have to work out who it was.

I have suggested that the concept of *q*-memory is coherent. Wiggins' case provides an illustration. The resulting people, in his case, both have apparent memories of living the life of the original person.

If they agree that they are not this person, they will have to regard these as only *q*-memories. And when they are asked a question like "Have you heard this music before?" they might have to answer "I am sure that I *q*-remember hearing it. But I am not sure whether I remember hearing it. I am not sure whether it was I who heard it, or the original person."

We can next point out that on our definition every memory is also a *q*-memory. Memories are, simply, *q*-memories of one's own experiences. Since this is so, we could afford now to drop the concept of memory and use in its place the wider concept *q*-memory. If we did, we should describe the relation between an experience and what we now call a "memory" of this experience in a way which does not presuppose that they are had by the same person.

This way of describing this relation has certain merits. It vindicates the "memory criterion" of personal identity against the charge of circularity. And it might, I think, help with the problem of other minds.

But we must move on. We can next take the relation between an intention and a later action. It may be a logical truth that we can intend to perform only our own actions. But intentions can be redescribed as *q*-intentions. And one person could *q*-intend to perform another person's actions.

Wiggins' case again provides the illustration. We are supposing that neither of the resulting people is the original person. If so, we shall have to agree that the original person can, before the operation, *q*-intend to perform their actions. He might, for example, *q*-intend, as one of them, to continue his present career, and, as the other, to try something new. (I say "*q*-intend as one of them" because the phrase "*q*-intend that one of them" would not convey the directness of the relation which is involved. If I intend that someone else should do something, I cannot get him to do it simply by forming this intention. But if I am the original person, and he is one of the resulting people, I can.)

The phrase "*q*-intend as one of them" reminds us that we need a sense in which one person can survive as two. But we can first point out that the concepts of *q*-memory and *q*-intention give us our

model for the others that we need: thus, a man who can *q-remember* could *q-recognize*, and be a *q-witness of*, what he has never seen; and a man who can *q-intend* could have *q-ambitions*, make *q-promises*, and be *q-responsible* for.

To put this claim in general terms: many different relations are included within, or are a consequence of, *psychological continuity*. We describe these relations in ways which presuppose the continued existence of one person. But we could describe them in new ways which do not.

This suggests a bolder claim. It might be possible to think of experiences in a wholly "impersonal" way. I shall not develop this claim here. What I shall try to describe is a way of thinking of our own identity through time which is more flexible, and less misleading, than the way in which we now think.

This way of thinking will allow for a sense in which one person can survive as two. A more important feature is that it treats survival as a matter of degree.

4. Degrees of Survival

We must first show the need for this second feature. I shall use two imaginary examples.

The first is the converse of Wiggins' case: fusion. Just as division serves to show that what matters in survival need not be one-one, so fusion serves to show that it can be a question of degree.

Physically, fusion is easy to describe. Two people come together. While they are unconscious, their two bodies grow into one. One person then wakes up.

The psychology of fusion is more complex. One detail we have already dealt with in the case of the exam. When my mind was reunited, I remembered just having thought out two calculations. The one person who results from a fusion can, similarly, *q-remember* living the lives of the two original people. None of their *q-memories* need be lost.

But some things must be lost. For any two people who fuse together will have different characteristics, different desires, and different intentions. How can these be combined?

We might suggest the following. Some of these will be compatible. These can coexist in the one

resulting person. Some will be incompatible. These, if of equal strength, can cancel out, and if of different strengths, the stronger can be made weaker. And all these effects might be predictable.

To give examples—first, of compatibility: I like **Palladio** and intend to visit Venice. I am about to fuse with a person who likes **Giotto** and intends to visit Padua. I can know that the one person we shall become will have both tastes and both intentions. Second, of incompatibility: I hate red hair, and always vote Labour. The other person loves red hair, and always votes Conservative. I can know that the one person we shall become will be indifferent to red hair, and a floating voter.

If we were about to undergo a fusion of this kind, would we regard it as death?

Some of us might. This is less absurd than regarding division as death. For after my division the two resulting people will be in every way like me, while after my fusion the one resulting person will not be wholly similar. This makes it easier to say, when faced with fusion, "I shall not survive," thus continuing to regard survival as a matter of all-or-nothing.

This reaction is less absurd. But here are two analogies which tell against it.

First, fusion would involve the changing of some of our characteristics and some of our desires. But only the very self-satisfied would think of this as death. Many people welcome treatments with these effects.

Second, someone who is about to fuse can have, beforehand, just as much "intentional control" over the actions of the resulting individual as someone who is about to marry can have, beforehand, over the actions of the resulting couple. And the choice of a partner for fusion can be just as well considered as the choice of a marriage partner. The two original people can make sure (perhaps by "trial fusion") that they do have compatible characters, desires, and intentions.

I have suggested that fusion, while not clearly survival, is not clearly failure to survive, and hence that what matters in survival can have degrees.

To reinforce this claim we can now turn to a second example. This is provided by certain imaginary beings. These beings are just like ourselves except that they reproduce by a process of natural division.

We can illustrate the histories of these imagined beings with the aid of a diagram [see below]. The lines on the diagram represent the spatiotemporal paths which would be traced out by the bodies of these beings. We can call each single line (like the double line) a "branch"; and we can call the whole structure a "tree." And let us suppose that each "branch" corresponds to what is thought of as the life of one individual. These individuals are referred to as "A," "B + 1," and so forth.

Now, each single division is an instance of Wiggins' case. So A's relation to both B + 1 and B + 2 is just as good as survival. But what of A's relation to B + 30?

I said earlier that what matters in survival could be provisionally referred to as "psychological continuity." I must now distinguish this relation from another, which I shall call "psychological connectedness."

Let us say that the relation between a *q*-memory and the experience *q*-remembered is a "direct" relation. Another "direct" relation is that which holds between a *q*-intention and the *q*-intended action. A third is that which holds between different expressions of some lasting *q*-characteristic.

"Psychological connectedness," as I define it, requires the holding of these direct psychological relations. "Connectedness" is not transitive, since these relations are not transitive. Thus, if X *q*-remembers most of Y's life, and Y *q*-remembers most of Z's life, it does not follow that X *q*-remembers most of Z's life. And if X carries out the *q*-intentions of Y, and Y carries out the *q*-intentions of Z, it does not follow that X carries out the *q*-intentions of Z.

"Psychological continuity," in contrast, only requires overlapping chains of direct psychological relations. So "continuity" is transitive.

To return to our diagram. A is psychologically continuous with B + 30. There are between the two continuous chains of overlapping relations. Thus, A has *q*-intentional control over B + 2, B + 2 has *q*-intentional control over B + 6, and so on up to B + 30. Or B + 30 can *q*-remember the life of B + 14, B + 14 can *q*-remember the life of B + 6, and so on back to A.⁵

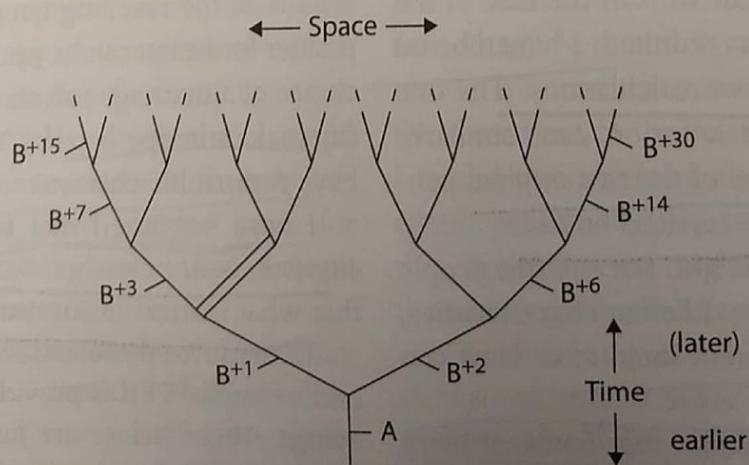
A, however, need not be psychologically connected to B + 30. Connectedness requires direct relations. And if these beings are like us, A cannot stand in such relations to every individual in his indefinitely long "tree." *Q*-memories will weaken with the passage of time, and then fade away. *Q*-ambitions, once fulfilled, will be replaced by others. *Q*-characteristics will gradually change. In general, A stands in fewer and fewer direct psychological relations to an individual in his "tree" the more remote that individual is. And if the individual is (like B + 30) sufficiently remote, there may be between the two no direct psychological relations.

Now that we have distinguished the general relations of psychological continuity and psychological connectedness, I suggest that connectedness is a more important element in survival. As a claim about our own survival, this would need more arguments than I have space to give. But it seems clearly true for my imagined beings. A is as close psychologically to B + 1 as I today am to myself

$$X = Y$$

$$Y = Z$$

$$X \neq Z$$



tomorrow. A is as distant from B + 30 as I am from my great-great-grandson.

Even if connectedness is not more important than continuity, the fact that one of these is a relation of degree is enough to show that what matters in survival can have degrees. And in any case the two relations are quite different. So our imagined beings would need a way of thinking in which this difference is recognized.

5. Survival without Identity

What I propose is this.

First, A can think of any individual, anywhere in his "tree," as "a descendant self." This phrase implies psychological continuity. Similarly, any later individual can think of any earlier individual on the single path which connects him to A as "an ancestral self."

Since psychological continuity is transitive, "being an ancestral self of" and "being a descendant self of" are also transitive.

To imply psychological connectedness I suggest the phrases "one of my future selves" and "one of my past selves."

These are the phrases with which we can describe Wiggins' case. For having past and future selves is, what we needed, a way of continuing to exist which does not imply identity through time. The original person does, in this sense, survive Wiggins' operation: the two resulting people are his later selves. And they can each refer to him as "my past self." (They can share a past self without being the same self as each other.)

Since psychological connectedness is not transitive, and is a matter of degree, the relations "being a past self of" and "being a future self of" should themselves be treated as relations of degree. We allow for this series of descriptions: "my most recent self," "one of my earlier selves," "one of my distant selves," "hardly one of my past selves (I can only *quasi-remember* a few of his experiences)," and, finally, "not in any way one of my past selves—just an ancestral self." [. . .]

Notes

1. Suppose the resulting people fight a duel. Are there three people fighting, one on each side, and one on both? And suppose one of the bullets kills. Are there two acts, one murder and one suicide? How many people are left alive? One? Two? (We could hardly say, "One and a half.")
2. 1066 is the date of the Norman invasion of England: after English king Harold's defeat and death in the Battle of Hastings, William of Normandy claimed the nation's crown.—LG
3. But not a necessary condition, for in the absence of psychological continuity bodily identity might be sufficient.
4. This is what so many writers have overlooked. Cf. Thomas Reid: "My memory testifies not only that this was done, but that it was done by me who now remember it" (203).
5. The chain of continuity must run in one direction of time. B + 2 is not, in the sense I intend, psychologically continuous with B + 1.

Works Cited

- Reid, T. "Of Identity." *Essays on the Intellectual Powers of Man*. Ed. A.D. Woozley. London: Macmillan, 1941. 200–6.
- Shoemaker, S. *Self-Knowledge and Self-Identity*. Ithaca: Cornell University Press, 1963.
- Sperry, R.W. "Brain Bisection and Mechanisms of Consciousness." *Brain and Conscious Experience*.
- Ed. J.C. Eccles. New York: Springer-Verlag, 1966. 298–313.
- Wiggins, D. *Identity and Spatio-Temporal Continuity*. Oxford: Blackwell, 1967.
- Williams, B.A.O. "Bodily Continuity and Personal Identity." *Analysis* 21, 2 (1960): 43–8.

Study Questions

1. What three possible responses to the fission thought experiment does Parfit identify?
2. Parfit's ultimate response to the fission thought experiment is that it is possible to *survive* as two people without *being* two people. What role do "q-memory" and "q-intention" play in this argument?
3. How does Parfit's fusion thought experiment contribute to his argument that survival, unlike identity, is a matter of degree?
4. How does Parfit distinguish between "psychological continuity" and "psychological connectedness," and why does he consider the latter more important than the former when it comes to a person's survival?

Marya Schechtman

Hilary Putnam

2/20/15

Hilary Putnam, a well-known American philosopher, received his PhD at the University of California, Los Angeles, in 1951 and spent most of his career at Harvard University, retiring in 2000. He has done comprehensive work in many areas, including philosophy of language and science. In this article, Putnam defends semantic externalism: the position that meanings of words are fixed by physical facts and linguistic practices, not internal mental states.—GS

Meaning and Reference

1. Introduction

Unclear as it is, the traditional doctrine that the notion "meaning" possesses the extension/intension ambiguity has certain typical consequences. The doctrine that the meaning of a term is a concept carried the implication that meanings are mental entities. Frege, however, rebelled against this "psychologism." Feeling that meanings are public property—that the same meaning can be "grasped" by more than one person and by persons at different times—he identified concepts (and hence "intensions" or meanings) with abstract entities rather than mental entities. However, "grasping" these abstract entities was still an individual psychological act. None of these philosophers doubted that understanding a word (knowing its intension) was just a matter of being in a certain psychological state (somewhat in the way in which knowing how to factor numbers in one's head is just a matter of being in a certain very complex psychological state).

Secondly, the timeworn example of the two terms 'creature with a kidney' and 'creature with a heart' does show that two terms can have the same extension and yet differ in intension. But it was taken to be obvious that the reverse is impossible: two terms cannot differ in extension and have the same intension. Interestingly, no argument for this impossibility was ever offered. Probably it reflects the tradition of the ancient and medieval philosophers, who assumed that the concept corresponding to a term was just a conjunction of predicates, and hence that the concept corresponding to a term must always

provide a necessary and sufficient condition for falling into the extension of the term. For philosophers like Carnap, who accepted the verifiability theory of meaning, the concept corresponding to a term provided (in the ideal case, where the term had "complete meaning") a criterion for belonging to the extension (not just in the sense of "necessary and sufficient condition," but in the strong sense of way of recognizing whether a given thing falls into the extension or not). So theory of meaning came to rest on two unchallenged assumptions:

1. That knowing the meaning of a term is just a matter of being in a certain psychological state (in the sense of "psychological state," in which states of memory and belief are "psychological states"; no one thought that knowing the meaning of a word was a continuous state of consciousness, of course).
2. That the meaning of a term determines its extension (in the sense that sameness of intension entails sameness of extension).

I shall argue that these two assumptions are not jointly satisfied by any notion, let alone any notion of meaning. The traditional concept of meaning is a concept which rests on a false theory.

2. Are Meanings in the Head?

For the purpose of the following science-fiction examples, we shall suppose that somewhere there is a planet we shall call Twin Earth. Twin Earth is very

much like Earth: in fact, people on Twin Earth even speak English. In fact, apart from the differences we shall specify in our science-fiction examples, the reader may suppose that Twin Earth is *exactly* like Earth. He may even suppose that he has a *Doppelganger*—an identical copy—on Twin Earth, if he wishes, although my stories will not depend on this.

Although some of the people on Twin Earth (say, those who call themselves “Americans” and those who call themselves “Canadians” and those who call themselves “Englishmen,” etc.) speak English, there are, not surprisingly, a few tiny differences between the dialects of English spoken on Twin Earth and standard English.

One of the peculiarities of Twin Earth is that the liquid called “water” is not H_2O but a different liquid whose chemical formula is very long and complicated. I shall abbreviate this chemical formula simply as XYZ. I shall suppose that XYZ is indistinguishable from water at normal temperatures and pressures. Also, I shall suppose that the oceans and lakes and seas of Twin Earth contain XYZ and not water, that it rains XYZ on Twin Earth and not water, etc.

If a space ship from Earth ever visits Twin Earth, then the supposition at first will be that ‘water’ has the same meaning on Earth and on Twin Earth. This supposition will be corrected when it is discovered that “water” on Twin Earth is XYZ, and the Earthian space ship will report somewhat as follows.

“On Twin Earth the word ‘water’ means XYZ.”

Symmetrically, if a space ship from Twin Earth ever visits Earth, then the supposition at first will be that the word ‘water’ has the same meaning on Twin Earth and on Earth. This supposition will be corrected when it is discovered that “water” on Earth is H_2O , and the Twin Earthian space ship will report:

“On Earth the word ‘water’ means H_2O .”

Note that there is no problem about the extension of the term ‘water’: the word simply has two different meanings (as we say); in the sense in which it is used on Twin Earth, the sense of water_{TE}, what we call “water” simply isn’t water, while in the sense in which it is used on Earth, the sense of water_E, what the Twin Earthians call “water” simple isn’t

water. The extension of ‘water’ in the sense of water_E is the set of all wholes consisting of H_2O molecules, or something like that; the extension of water in the sense of water_{TE} is the set of all wholes consisting of XYZ molecules, or something like that.

Now let us roll the time back to about 1750.¹ The typical Earthian speaker of English did not know that water consisted of hydrogen and oxygen, and the typical Twin-Earthian speaker of English did not know that “water” consisted of XYZ. Let Oscar₁ be such a typical Earthian English speaker, and let Oscar₂ be his counterpart on Twin Earth. You may suppose that there is no belief that Oscar₁ had about water that Oscar₂ did not have about “water.” If you like, you may even suppose that Oscar₁ and Oscar₂ were exact duplicates in appearance, feelings, thoughts, interior monologue, etc. Yet the extension of the term ‘water’ was just as much H_2O on Earth in 1750 as in 1950; and the extension of the term ‘water’ was just as much XYZ on Twin Earth in 1750 as in 1950. Oscar₁ and Oscar₂ understood the term ‘water’ differently in 1750 *although they were in the same psychological state*, and although, given the state of science at the time, it would have taken their scientific communities about fifty years to discover that they understood the term ‘water’ differently. Thus the extension of the term ‘water’ (and, in fact, its “meaning” in the intuitive pre-analytical usage of that term) is *not* a function of the psychological state of the speaker by itself.²

But, it might be objected, why should we accept it that the term ‘water’ had the same extension in 1750 and in 1950 (on both Earths)? Suppose I point to a glass of water and say “this liquid is called water.” My “ostensive definition” of water has the following empirical presupposition: that the body of liquid I am pointing to bears a certain sameness relation (say, *x* is the same liquid as *y*, or *x* is the same_L as *y*) to most of the stuff I and other speakers in my linguistic community have on other occasions called “water.” If this presupposition is false because, say, I am—unknown to me—pointing to a glass of gin and not a glass of water, then I do not intend my ostensive definition to be accepted. Thus the ostensive definition conveys what might be called a “defeasible”

necessary and sufficient condition: the necessary and sufficient condition for being water is bearing the relation *same_L* to the stuff in the glass; but this is the necessary and sufficient condition only if the empirical presupposition is satisfied. If it is not satisfied, then one of a series of, so to speak, "fallback" conditions becomes activated.

The key point is that the relation *same_L* is a theoretical relation; whether something is or is not the same liquid as this may take an indeterminate amount of scientific investigation to determine. Thus, the fact that an English speaker in 1750 might have called XYZ "water," whereas he or his successors would not have called XYZ water in 1800 or 1850 does not mean that the "meaning" of 'water' changed for the average speaker in the interval. In 1750 or in 1850 or in 1950 one might have pointed to, say, the liquid in Lake Michigan as an example of "water." What changed was that in 1750 we would have mistakenly thought that XYZ bore the relation *same_L* to the liquid in Lake Michigan, whereas in 1800 or 1850 we would have known that it did not.

Let us now modify our science-fiction story. I shall suppose that molybdenum pots and pans can't be distinguished from aluminum pots and pans save by an expert. (This could be true for all I know, and, *a fortiori*, it could be true for all I know by virtue of "knowing the meaning" of the words *aluminum* and *molybdenum*.) We will now suppose that molybdenum is as common on Twin Earth as aluminum is on Earth, and that aluminum is as rare on Twin Earth as molybdenum is on Earth. In particular, we shall assume that "aluminum" pots and pans are made of molybdenum on Twin Earth. Finally, we shall assume that the words 'aluminum' and 'molybdenum' are switched on Twin Earth: 'aluminum' is the name of *molybdenum*, and 'molybdenum' is the name of *aluminum*. If a space ship from Earth visited Twin Earth, the visitors from Earth probably would not suspect that the "aluminum" pots and pans on Twin Earth were not made of aluminum, especially when the Twin Earthians said they were. But there is one important difference between the two cases. An Earthian metallurgist could tell very easily that "aluminum" was molybdenum, and a Twin Earthian

metallurgist could tell equally easily that aluminum was "molybdenum." (The shudder quotes in the preceding sentence indicate Twin Earthian usages.) Whereas in 1750 no one on either Earth or Twin Earth could have distinguished water from "water," the confusion of aluminum with "aluminum" involves only a part of the linguistic communities involved.

This example makes the same point as the preceding example. If Oscar₁ and Oscar₂ are standard speakers of Earthian English and Twin Earthian English, respectively, and neither is chemically or metallurgically sophisticated, then there may be no difference at all in their psychological states when they use the word 'aluminum'; nevertheless, we have to say that 'aluminum' has the extension *aluminum* in the idiolect of Oscar₁ and the extension *molybdenum* in the idiolect of Oscar₂. (Also we have to say that Oscar₁ and Oscar₂ mean different things by 'aluminum'; that 'aluminum' has a different meaning on Earth than it does on Twin Earth, etc.) Again we see that the psychological state of the speaker does not determine the extension (or the "meaning," speaking preanalytically) of the word.

Before discussing this example further, let me introduce a non-science-fiction example. Suppose you are like me and cannot tell an elm from a beech tree. We still say that the extension of 'elm' in my idiolect is the same as the extension of 'elm' in anyone else's, viz., the set of all elm trees, and that the set of all beech trees is the extension of 'beech' in both of our idiolects. Thus 'elm' in my idiolect has a different extension from 'beech' in your idiolect (as it should). Is it really credible that this difference in extension is brought about by some difference in our concepts? My concept of an elm tree is exactly the same as my concept of a beech tree (I blush to confess). If someone heroically attempts to maintain that the difference between the extension of 'elm' and the extension of 'beech' in my idiolect is explained by a difference in my psychological state, then we can always refute him by constructing a "Twin Earth" example—just let the words 'elm' and 'beech' be switched on Twin Earth (the way 'aluminum' and 'molybdenum' were in the previous example).

Moreover, suppose I have a *Doppelganger* on Twin Earth who is molecule for molecule "identical" with me. If you are a dualist, then also suppose my Doppelganger thinks the same verbalized thoughts I do, has the same sense data, the same dispositions, etc. It is absurd to think his psychological state is one bit different from mine: yet he "means" *beech* when he says "elm," and I "mean" *elm* when I say "elm." Cut the pie any way you like, "meanings just ain't in the head!"

3. A Sociolinguistic Hypothesis

The last two examples depend upon a fact about language that seems, surprisingly, never to have been pointed out: that there is division of linguistic labour. We could hardly use such words as 'elm' and 'aluminum' if no one possessed a way of recognizing elm trees and aluminum metal; but not everyone to whom the distinction is important has to be able to make the distinction. Let us shift the example; consider *gold*. Gold is important for many reasons: it is a precious metal; it is a monetary metal; it has symbolic value (it is important to most people that the "gold" wedding ring they wear *really* consist of gold and not just *look* gold); etc. Consider our community as a "factory": in this "factory" some people have the "job" of *wearing gold wedding rings*; other people have the "job" of *selling gold wedding rings*; still other people have the job of *telling whether or not something is really gold*. It is not at all necessary or efficient that every one who wears a gold ring (or a gold cufflink, etc.), or discusses the "gold standard," etc., engage in buying and selling gold. Nor is it necessary or efficient that everyone who buys and sells gold be able to tell whether or not something is really gold in a society where this form of dishonesty is uncommon (selling fake gold) and in which one can easily consult an expert in case of doubt. And it is *certainly* not necessary or efficient that everyone who has occasion to buy or wear gold be able to tell with any reliability whether or not something is really gold.

The foregoing facts are just examples of mundane division of labour (in a wide sense). But they engender

a division of linguistic labour: everyone to whom gold is important for any reason has to *acquire* the word 'gold'; but he does not have to acquire the *method* of recognizing whether something is or is not gold. He can rely on a special subclass of speakers. The features that are generally thought to be present in connection with a general name—necessary and sufficient conditions for membership in the extension, ways of recognizing whether something is in the extension, etc.—are all present in the linguistic community considered as a *collective body*; but that collective body divides the "labour" of knowing and employing these various parts of the "meaning" of 'gold'.

This division of linguistic labour rests upon and presupposes the division of *nonlinguistic* labour, of course. If only the people who know how to tell whether some metal is really gold or not have any reason to have the word 'gold' in their vocabulary, then the word 'gold' will be as the word 'water' was in 1750 with respect to that subclass of speakers, and the other speakers just won't acquire it at all. And some words do not exhibit any division of linguistic labour: 'chair', for example. But with the increase of division of labour in the society and the rise of science, more and more words begin to exhibit this kind of division of labour. 'Water', for example, did not exhibit it at all before the rise of chemistry. Today it is obviously necessary for every speaker to be able to recognize water (reliably under normal conditions), and probably most adult speakers even know the necessary and sufficient condition "water is H_2O ," but only a few adult speakers could distinguish water from liquids that superficially resembled water. In case of doubt, other speakers would rely on the judgment of these "expert" speakers. Thus the way of recognizing possessed by these "expert" speakers is also, through them, possessed by the collective linguistic body, even though it is not possessed by each individual member of the body, and in this way the most *recherché* fact about water may become part of the social meaning of the word although unknown to almost all speakers who acquire the word.

It seems to me that this phenomenon of division of linguistic labour is one that it will be very important for sociolinguistics to investigate. In connection

with it, I should like to propose the following hypothesis:

HYPOTHESIS OF THE UNIVERSALITY OF THE DIVISION OF LINGUISTIC LABOUR: Every linguistic community exemplifies the sort of division of linguistic labour just described; that is, it possesses at least some terms whose associated "criteria" are known only to a subset of the speakers who acquire the terms, and whose use by the other speakers depends upon a structured cooperation between them and the speakers in the relevant subsets.

It is easy to see how this phenomenon accounts for some of the examples given above of the failure of the assumptions (1 and 2). When a term is subject to the division of linguistic labour, the "average" speaker who acquires it does not acquire anything that fixes its extension. In particular, his individual psychological state *certainly* does not fix its extension; it is only the sociolinguistic state of the collective linguistic body to which the speaker belongs that fixes the extension.

We may summarize this discussion by pointing out that there are two sorts of tools in the world: there are tools like a hammer or a screwdriver which can be used by one person; and there are tools like a steamship which require the cooperative activity of a number of persons to use. Words have been thought of too much on the model of the first sort of tool.

4. Indexicality and Rigidity

The first of our science-fiction examples—'water' on Earth and on Twin Earth in 1750—does not involve division of linguistic labour, or at least does not involve it in the same way the examples of 'aluminum' and 'elm' do. There were not (in our story, anyway) any "experts" on water on Earth in 1750, nor any experts on "water" on Twin Earth. The example does involve things which are of fundamental importance to the theory of reference and also to the theory of necessary truth, which we shall now discuss.

Let W_1 and W_2 be two possible worlds in which I exist and in which this glass exists and in which

I am giving a meaning explanation by pointing to this glass and saying "This is water." Let us suppose that in W_1 the glass is full of H_2O and in W_2 the glass is full of XYZ. We shall also suppose that W_1 is the *actual* world, and that XYZ is the stuff typically called "water" in the world W_2 (so that the relation between English speakers in W_1 and English speakers in W_2 is exactly the same as the relation between English speakers on Earth and English speakers on Twin Earth). Then there are two theories one might have concerning the meaning of 'water':

1. One might hold that 'water' was world-relative but *constant* in meaning (i.e., the word has a constant relative meaning). On this theory, 'water' means the same in W_1 and W_2 ; it's just that water is H_2O in W_1 , and water is XYZ in W_2 .
2. One might hold that water is H_2O in all worlds (the stuff called "water" in W_2 isn't water), but 'water' doesn't have the same meaning in W_1 and W_2 .

If what was said before about the Twin Earth case was correct, then (2) is clearly the correct theory. When I say "this (liquid) is water," the "this" is, so to speak, a *de re* "this"—i.e., the force of my explanation is that "water" is whatever bears a certain equivalence relation (the relation we called "same" above) to the piece of liquid referred to as "this" in the *actual* world. [...]

Kripke calls a designator "rigid" (in a given sentence) if (in that sentence) it refers to the same individual in every possible world in which the designator designates. If we extend this notion of rigidity to substance names, then we may express Kripke's theory and mine by saying that the term 'water' is rigid.

The rigidity of the term 'water' follows from the fact that when I give the "ostensive definition": "this (liquid) is water," I intend (2) and not (1).

We may also say, following Kripke, that when I give the ostensive definition "this (liquid) is water," the demonstrative 'this' is rigid. [...]

Words like 'now', 'this', 'here' have long been recognized to be *indexical*, or *token-reflexive*—i.e.,

to have an extension which varies from context to context or token to token. For these words, no one has ever suggested the traditional theory that "intension determines extension." To take our Twin Earth example: if I have a *Doppelganger* on Twin Earth, then when I think "I have a headache," he thinks "I have a headache." But the extension of the particular token of 'T' in his verbalized thought is himself (or his unit class, to be precise), while the extension of the token of 'T' in my verbalized thought is *me* (or my unit class, to be precise). So the same word, 'T', has two different extensions in two different idiolects; but it does not follow that the concept I have of myself is in any way different from the concept my *Doppelganger* has of himself.

Now then, we have maintained that indexicality extends beyond the *obviously* indexical words and morphemes (e.g., the tenses of verbs). Our theory can be summarized as saying that words like 'water' have an unnoticed indexical component: "water" is stuff that bears a certain similarity relation to the water *around here*. Water at another time or in another place or even in another possible world has to bear the relation *same₁* to our "water" *in order to be water*. Thus the theory that (1) words have "intensions," which are something like concepts associated with the words by speakers; and (2) intension determines extension—cannot be true of natural-kind words like 'water' for the same reason it cannot be true of obviously indexical words like 'T'.

The theory that natural-kind words like 'water' are indexical leaves it open, however, whether to say that 'water' in the Twin Earth dialect of English has the same *meaning* as 'water' in the Earth dialect and a different extension—which is what we normally say about 'T' in different idiolects—thereby giving up the doctrine that "meaning (intension) determines extension," or to say, as we have chosen to do, that difference in extension is *ipso facto* a difference in meaning for natural-kind words, thereby giving up the doctrine that meanings are concepts, or, indeed, mental entities of *any kind*.³

It should be clear, however, that Kripke's doctrine that natural-kind words are rigid designators and our doctrine that they are indexical are but two ways of making the same point.

We have now seen that the extension of a term is not fixed by a concept that the individual speaker has in his head, and this is true both because extension is, in general, determined *socially*—there is division of linguistic labour as much as of "real" labour—and because extension is, in part, determined *indexically*. The extension of our terms depends upon the actual nature of the particular things that serve as paradigms, and this actual nature is not, in general, fully known to the speaker. Traditional semantic theory leaves out two contributions to the determination of reference—the contribution of society and the contribution of the real world; a better semantic theory must encompass both.

Notes

1. In the early 1780s, it was discovered that water is not a simple substance but composed of two gases, which the French chemist Antoine Lavoisier (1743–94) named oxygen and hydrogen.—LG
2. See note 3 [...] and the corresponding text.
3. Our reasons for rejecting the first option—to say that 'water' has the same meaning on Earth and on Twin Earth, while giving up the doctrine that meaning determines reference—are presented in "The Meaning of 'Meaning'." They may be

illustrated thus: Suppose 'water' has the same meaning on Earth and on Twin Earth. Now, let the word 'water' become phonemically different on Twin Earth—say, it becomes 'quaxel'. Presumably, this is not a change in meaning per se, on any view. So 'water' and 'quaxel' have the same meaning (although they refer to different liquids). But this is highly counterintuitive. Why not say, then, that 'elm' in my idiolect has the same meaning as 'beech' in your idiolect, although they refer to different trees?

Study Questions

1. What is the difference between "intension" and "extension" when it comes to meaning?
2. Describe Putnam's famous Twin Earth example. Explain how the Twin Earth example challenges the traditional theory of meaning and supports Putnam's theory of meaning.
3. With the help of an example, outline Putnam's theory of the division of linguistic labour.
4. What role does indexicality as a property of natural-kind terms play in Putnam's defence of his theory of meaning of such terms?

John Dupré

2/20/15

John Dupré received his PhD at Cambridge in 1981, and taught at Stanford University and Birkbeck College, University of London, before becoming director of Egenis, the Economic and Social Research Council Centre for Genomics in Society at the University of Exeter. Dupré has published on a wide range of topics in philosophy of science, including reductionism, human nature, genomics, Darwin, and, the focus of this essay, natural kinds.—GS

Sally Haslanger

3/2/15

Sally Haslanger acquired her PhD from University of California, Berkeley, in 1985. She is a professor of philosophy and linguistics and director of women's and gender studies at the Massachusetts Institute of Technology. This article's critical analysis of the categories of "gender" and "race" combines Haslanger's philosophical interests in metaphysics and feminist theory.—GS

Gender and Race: (What) Are They? (What) Do We Want Them To Be?

1. Introduction

It is always awkward when someone asks me informally what I'm working on and I answer that I'm trying to figure out what gender is. For outside a rather narrow segment of the academic world, the term "gender" has come to function as the polite way to talk about the sexes. And one thing people feel pretty confident about is their knowledge of the difference between males and females. Males are those human beings with a range of familiar primary and secondary sex characteristics, most important being the penis; females are those with a different set, most important being the vagina or, perhaps, the uterus. Enough said. Against this background, it isn't clear what could be the point of an inquiry, especially a philosophical inquiry, into "what gender is."

But within that rather narrow segment of the academic world concerned with gender issues, not only is there no simple equation of sex and gender, but the seemingly straightforward anatomical distinction between the sexes has been challenged as well. What began as an effort to note that men and women differ socially as well as anatomically has prompted an explosion of different uses of the term "gender." Within these debates, not only is it unclear what gender is and how we should go about understanding it, but whether it is anything at all.

The situation is similar, if not worse, with respect to race. The self-evidence of racial distinctions in everyday American life is at striking odds with the

uncertainty about the category of race in law and the academy. Work in the biological sciences has informed us that our practices of racial categorization don't map neatly onto any useful biological classification; but that doesn't settle much, if anything. For what should we make of our tendency to classify individuals according to race, apparently on the basis of physical appearance? And what are we to make of the social and economic consequences of such classifications? Is race real or is it not?

2. The Question(s)

[Concerning] the questions, "What is race?" or "What is gender?" we can distinguish [...] three projects with importantly different priorities: conceptual, descriptive, and analytical.

(A)

A conceptual inquiry into race or gender would seek an articulation of our concepts of race or gender. To answer the conceptual question, one way to proceed would be to use the method of reflective equilibrium. (Although within the context of analytic philosophy this might be seen as a call for a conceptual analysis of the term(s), I want to reserve the term "analytical" for a different sort of project, described below.)

In contrast to the conceptual project, a descriptive project is not concerned with exploring the nuances of our concepts (or anyone else's for that matter); it focuses instead on their extension. Here, the task is to develop potentially more accurate concepts

through careful consideration of the phenomena, usually relying on empirical or quasi-empirical methods. Paradigm descriptive projects occur in studying natural phenomena. [...] However, a descriptive approach need not be confined to a search for natural or physical kinds; inquiry into what it is to be, e.g., a human right, a citizen, a democracy, might begin by considering the full range of what has counted as such to determine whether there is an underlying (possibly social) kind that explains the temptation to group the cases together. Just as natural science can enrich our “folk” conceptualization of natural phenomena, social sciences (as well as the arts and humanities) can enrich our “folk” conceptualization of social phenomena. So, a descriptive inquiry into race and gender need not presuppose that race and gender are biological kinds; instead it might ask whether our uses of race and gender vocabularies are tracking social kinds, and if so which ones.

The third sort of project takes an analytical approach to the question, “What is gender?” or “What is race?” On this approach,

So, on an analytical approach, the questions “What is gender?” or “What is race?” require us to consider what work we want these concepts to do for us; why do we need them at all? The responsibility is ours to define them for our purposes. In doing so we will want to be responsive to some aspects of ordinary usage (and to aspects of both the connotation and extension of the terms). However, neither ordinary usage nor empirical investigation is overriding, for there is a stipulative element to the project: *this* is the phenomenon we need to be thinking about. Let the term in question refer to it. On this approach, the world by itself can’t tell us what gender is, or what race is; it is up to us to decide what in the world, if anything, they are.

This essay pursues an analytical approach to defining race and gender. [...]

The goal of the project is to consider what work the concepts of gender and race might do for us in a critical—specifically feminist and antiracist—social theory, and to suggest concepts that can accomplish

of men and women. My strategy is to offer a focal analysis that defines gender, in the primary sense, as a social class. A focal analysis undertakes to explain a variety of connected phenomena in terms of their relations to one that is theorized as the central or core phenomenon. As I see it, the core phenomenon to be addressed is the pattern of social relations that constitute the social classes of men as dominant and women as subordinate. [...] I see my emphasis as falling within, though not following uncritically, the tradition of materialist feminism. [...]

The main strategy of materialist feminist accounts of gender has been to define gender in terms of women's subordinate position in systems of male dominance. Although there are materialist feminist roots in Marxism, contemporary versions resist the thought that all social phenomena can be explained in or reduced to economic terms; and although materialist feminists emphasize the role of language and culture in women's oppression, there is a wariness of extreme forms of linguistic constructivism and a commitment to staying grounded in the material realities of women's lives. In effect, there is a concerted effort to show how gender oppression is jointly sustained by both cultural and material forces.

Critiques of universalizing feminisms have taught us to be attentive to the variety of forms gender takes and the concrete social positions females occupy. However it is compatible with these commitments to treat the category of gender as a genus that is realized in different ways in different contexts; doing so enables us to recognize significant patterns in the ways that gender is instituted and embodied. Working at the most general level, then, the materialist strategy offers us three basic principles to guide us in understanding gender:

- (i) Gender categories are defined in terms of how one is socially positioned, where this is a function of, e.g., how one is viewed, how one is treated, and how one's life is structured socially, legally, and economically; gender is not defined in terms of an individual's intrinsic physical or psychological features.

(This allows that there may be other categories—such as sex—that are defined in terms of intrinsic physical features. Note, however, that once we focus our attention on gender as social position, we must allow that one can be a woman without ever (in the ordinary sense) "acting like a woman," "feeling like a woman," or even having a female body.)

- (ii) Gender categories are defined hierarchically within a broader complex of oppressive relations; one group (viz., women) is socially positioned as subordinate to the other (viz., men), typically within the context of other forms of economic and social oppression.
- (iii) Sexual difference functions as the physical marker to distinguish the two groups, and is used in the justification of viewing and treating the members of each group differently.

(Tentatively) we can capture these main points in the following analyses:

- S is a woman iff_{df} [if and only if, by definition] S is systematically subordinated along some dimension (economic, political, legal, social, etc.), and S is "marked" as a target for this treatment by observed or imagined bodily features presumed to be evidence of a female's biological role in reproduction.
- S is a man iff_{df} S is systematically privileged along some dimension (economic, political, legal, social, etc.), and S is "marked" as a target for this treatment by observed or imagined bodily features presumed to be evidence of a male's biological role in reproduction.

It is a virtue, I believe, of these accounts, that depending on context, one's sex may have a very different meaning and it may position one in very different kinds of hierarchies. The variation will clearly occur from culture to culture (and subculture to subculture); so e.g., to be a Chinese woman of the 1790s, a Brazilian woman of the 1890s, or an American woman of the 1990s may involve very different social relations, and very different kinds

of oppression. Yet on the analysis suggested, these groups count as women insofar as their subordinate positions are marked and justified by reference to (female) sex. Similarly, this account allows that the substantive import of gender varies even from individual to individual within a culture depending on how the meaning of sex interacts with other socially salient characteristics (e.g., race, class, sexuality, etc.). For example, a privileged white woman and a black woman of the underclass will both be women insofar as their social positions are affected by the social meanings of being female; and yet the social implications of being female vary for each because sexism is intertwined with race and class oppression.

4. What Is Oppression?

There are points in the proposed analysis that require clarification, however. What does it mean to say that someone is “systematically subordinated” or “privileged,” and further, that the subordination occurs “on the basis of” certain features? The background idea is that women are *oppressed*, and that they are oppressed *as women*. But we still need to ask: What does it mean to say that women are oppressed, and what does the qualification “as women” add?

Marilyn Frye’s account of oppression with Iris Young’s elaborations provides a valuable starting point. [...] Oppression in the intended sense is a structural phenomenon that positions certain groups as disadvantaged and others as advantaged or privileged in relation to them. Oppression consists of, “an enclosing structure of forces and barriers which tends to the immobilization and reduction of a group or category of people” (Frye 11). Importantly, such structures, at least as we know them, are not designed and policed by those in power, rather, “... oppression refers to the vast and deep injustices some groups suffer as a consequence of often unconscious assumptions and reactions of well-meaning people in ordinary interactions, media and cultural stereotypes, and structural features of bureaucratic hierarchies and market mechanisms—in short, the normal processes of everyday life” (Young 41).

Developing this concept of oppression, Young specifies five forms it can take: exploitation, marginalization, powerlessness, cultural imperialism, and (systematic) violence. The key point for us is that oppression comes in different forms, and even if one is privileged along some dimension (e.g., in income or respect), one might be oppressed in others. In fact, one might be systematically subordinated along some social axis, and yet still be tremendously privileged in one’s *overall social position*.

It is clear that women are oppressed in the sense that women are members of groups that suffer exploitation, marginalization, etc. But how should we understand the claim that women are oppressed *as women*. Frye explains this as follows:

One is marked for application of oppressive pressures by one’s membership in some group or category . . . In the case at hand, it is the category, *woman*. . . . If a woman has little or no economic or political power, or achieves little of what she wants to achieve, a major causal factor in this is that she is a woman. For any woman of any race or economic class, being a woman is significantly attached to whatever disadvantages and deprivations she suffers, be they great or small. . . . [In contrast,] being male is something [a man] has going for him, even if race or class or age or disability is going against him. (15–6)

But given the diffusion of power in a model of structural oppression how are we to make sense of one’s being “marked” and the “application” of pressures? In the context of oppression, certain properties of individuals are socially meaningful. This is to say that the properties play a role in a broadly accepted (though usually not fully explicit) representation of the world that functions to justify and motivate particular forms of social intercourse. The significant properties in question—in the cases at hand, assumed or actual properties of the body—mark you “for application of oppressive pressures” insofar as the attribution of these properties is interpreted as adequate, in light of this background

representation, to explain and/or justify your position in a structure of oppressive social relations. In the case of women, the idea is that societies are guided by representations that link being female with other facts that have implications for how one should be viewed and treated; insofar as we structure our social life to accommodate the cultural meanings of the female (and male) body, females occupy an oppressed social position.

Although I agree with Frye that in sexist societies social institutions are structured in ways that on the whole disadvantage females and advantage males, we must keep in mind that societies are not monolithic and that sexism is not the only source of oppression. For example, in the contemporary US, there are contexts in which being black and male marks one as a target for certain forms of systematic violence (e.g., by the police). In those contexts, contrary to Frye's suggestion, being male is not something that a man "has going for him"; though there are other contexts (also in the contemporary US) in which black males benefit from being male. In examples of this sort, the systematic violence against males as males is emasculating (and may be intended as such); but there are important differences between an emasculated man and a woman. On the sort of view we're considering, a woman is someone whose subordinated status is marked by reference to (assumed) female anatomy; someone marked for subordination by reference to (assumed) male anatomy does not qualify as a woman, but also, in the particular context, is not socially positioned as a man.

5. What Is Race?

One advantage of this account of gender is the parallel it offers for race. To begin, let me review a couple of points that I take to be matters of established fact: First, there are no racial genes responsible for the complex morphologies and cultural patterns we associate with different races. Second, in different contexts racial distinctions are drawn on the basis of different characteristics, e.g., the Brazilian and US classification schemes for who counts as "black" differ. For these reasons and others, it appears that

race, like gender, could be fruitfully understood as a position within a broad social network.

Although suggestive, this idea is not easy to develop. It is one thing to acknowledge that race is socially real, even if a biological fiction; but it is another thing to capture in general terms "the social meaning of colour." There seem to be too many different forms race takes. Note, however, that we encountered a similar problem with gender: is there any prospect for a unified analysis of "the social meaning of sex"? The materialist feminist approach offered a helpful strategy: don't look for an analysis that assumes that the meaning is always and everywhere the same; rather, consider how members of the group are *socially positioned*, and what *physical markers* serve as a supposed basis for such treatment.

How might we extend this strategy to race? Transposing the slogan, we might say that race is the social meaning of the geographically marked body, familiar markers being skin colour, hair type, eye shape, physique. To develop this, I propose the following account [...]:¹

A group is *racialized* iff its members are socially positioned as subordinate or privileged along some dimension (economic, political, legal, social, etc.), and the group is "marked" as a target for this treatment by observed or imagined bodily features presumed to be evidence of ancestral links to a certain geographical region.

[...] In other words, races are those groups demarcated by the geographical associations accompanying perceived body type, when those associations take on evaluative significance concerning how members of the group should be viewed and treated. As in the case of gender, the ideology need not use physical morphology or geography as the entire basis for "appropriate" treatment; these features may instead simply be "markers" of other characteristics that the ideology uses to justify the treatment in question.

Given this definition, we can say that S is of the white (black, Asian . . .) race iff [if and only if] whites

(blacks, Asians . . .) are a racialized group, and S is a member. [. . .] [W]hether a group is racialized, and so how and whether an individual is raced, is not an absolute fact, but will depend on context. For example, blacks, whites, Asians, Native Americans, are currently racialized in the US insofar as these are all groups defined in terms of physical features associated with places of origin, and insofar as membership in the group functions socially as a basis for evaluation. However, some groups are not currently racialized in the US, but have been so in the past and possibly could be again (and in other contexts are), e.g., the Italians, the Germans, the Irish.

It is useful to note a possible contrast between race and ethnicity. I don't have a theory of ethnicity to offer; these are some preliminary comparisons. One's ethnicity concerns one's ancestral links to a certain geographical region (perhaps together with participation in the cultural practices of that region); often ethnicity is associated with characteristic physical features. For our purposes, however, it might be useful to employ the notion of "ethnicity" for those groups that are like races as I've defined them except that they do not experience systematic subordination or privilege in the context in question.² Ethnic groups can be (and are) racialized, however, and when they are, one's membership in the group positions one in a social hierarchy; but (on the view I'm sketching) the occurrence of this hierarchical positioning means that the group has gone beyond simply being an ethnic group and functions in that context as a race. In short, we can distinguish between grouping individuals on the basis of their (assumed) origins, and grouping them hierarchically on the basis of their (assumed) origins, and the contrast between race and ethnicity might be a useful way to capture this distinction.

6. Negotiating Terms

Let me now turn to summarize some of the advantages of the proposed definitions. At this point we could bracket the terminological issues and just consider whether the groups in question are ones that are important to consider given the goals of our inquiry. I hope it is clear from what I've already said how the

analyses can help us identify and critique broad patterns of racial and sexual oppression, and how they accommodate the intersectionality of social categories. But a further and, I think, more interesting question is whether it is useful to think of these groups in these terms: Does it serve both the goal of understanding racial and sexual oppression, and of achieving sexual and racial equality to think of ourselves as men or women, or raced in the ways proposed?

By appropriating the everyday terminology of race and gender, the analyses I've offered invite us to acknowledge the force of oppressive systems in framing our personal and political identities. Each of us has some investment in our race and gender: I am a White woman. On my accounts, this claim locates me within social systems that in some respects privilege and in others subordinate me. Because gender and racial inequality are not simply a matter of public policy but implicate each of us at the heart of our self-understandings, the terminological shift calls us to reconsider who we think we are.

This point highlights why the issue of terminological appropriation is especially sensitive when the terms designate categories of social identity. Writing in response to a NY Times editorial supporting the terminological shift from "black" to "African-American," Trey Ellis responded: "When somebody tries to tell me what to call myself in all its uses just because they come to some decision at a cocktail party to which I wasn't even invited, my mama raised me to tell them to kiss my black ass. In many cases, African-American just won't do." The issue is not just what words we should use, and who gets to say what words to use, but who we take ourselves to be, and so, in some sense, who we are. Terms for social groups can function as descriptive terms: it may be accurate to say that someone is a woman when she satisfies certain conditions. However, terms for social groups serve other rhetorical purposes. Typically the act of classifying someone as a member of a social group invokes a set of "appropriate" (contextually specific) norms and expectations. It positions her in a social framework and makes available certain kinds of evaluation; in short, it carries prescriptive force. Accepting or identifying with the classification

typically involves an endorsement of some norms and expectations, however, not always the socially sanctioned ones. [. . .]

Although “identifying” someone as a member of a social group invokes a set of “appropriate” norms, what these norms are is not fixed. What it means to be a woman, or to be white, or to be Latino, in this sense, is unstable and always open to contest. The instability across time is necessary to maintain the basic structure of gender and race relations through other social changes: as social roles change—prompted by the economy, immigration, political movements, natural disasters, war, etc.—the contents of normative race and gender identities adjust. The flexibility across contexts accommodates the complexity of social life: what norms are assumed to apply depends on the dominant social structure, the ideological context, and other dimensions of one’s identity (such as class, age, ability, sexuality). But this instability and flexibility is exactly what opens the door for groups to redefine themselves in new ways. One strategy is for the group to adopt new names (“African-American,” “womyn”); another is to appropriate old names with a normative twist (“queer”); but in some cases the contest is over the meanings of the standard terms (“Ain’t I a woman?”). [. . .]

Given the normative force and political potential of identifying someone (or self-identifying) in racial or gendered terms, how do we evaluate a terminological appropriation of the kind I’m proposing? For example, isn’t there something disingenuous about appropriating race and gender terminology because it is used to frame how we think of ourselves and each other, in order to use them for new concepts that are *not* part of our self-understandings?

This latter question is especially pressing because the appropriation under consideration intentionally invokes what many find to be positive self-understandings—being Latina, being a white man—and offers analyses of them which emphasize the broader context of injustice. Thus there is an invitation not only to revise one’s understanding of these categories (given their instability, this happens often enough), but to revise one’s relationship to their prescriptive force. By offering these analyses

of our ordinary terms, I call upon us to reject what seemed to be positive social identities. I’m suggesting that we should work to undermine those forces that make being a man, a woman, or a member of a racialized group possible; we should refuse to be gendered man or woman, refuse to be raced. This goes beyond denying essentialist claims about one’s embodiment and involves an active political commitment to live one’s life differently (Stoltenberg). In one sense this appropriation is “just semantics”: I’m asking us to use an old term in a new way. But it is also politics: I’m asking us to understand ourselves and those around us as deeply moulded by injustice and to draw the appropriate prescriptive inference. This, I hope, will contribute to empowering critical social agents. However, whether the terminological shift I’m suggesting is politically useful will depend on the contexts in which it is employed and the individuals employing it. The point is not to legislate what terms to use in all contexts, but to offer resources that should be used judiciously.

7. Lingering Concerns, Promising Alternatives

There is, nonetheless, a broader concern one might have about the strategy I’ve employed: Why build hierarchy into the definitions? Why not define gender and race as those social positions motivated and justified by cultural responses to the body, without requiring that the social positions are hierarchical? Wouldn’t that provide what we need without implying (implausibly) that women are, by definition, subordinate, men, by definition, privileged, and races, by definition, hierarchically positioned?

If we were to remove hierarchy from the definitions, then there would be two other benefits: first, by providing a place in our model for cultural representations of the body besides those that contribute to maintaining subordination and privilege, we could better acknowledge that there are positive aspects to having a gender and a race. And second, the accounts would provide a framework for envisioning the sorts of constructive changes needed to create a more just world. The suggestion that we

must eliminate race and gender may be a powerful rallying call to those who identify with radical causes, but it is not at all clear that societies can or should avoid giving meanings to the body, or organizing themselves to take sexual and reproductive differences into account. [. .]

Consider gender. I am sympathetic to radical rethinkings of sex and gender. In particular, I believe that we should refuse to use anatomy as a primary basis for classifying individuals and that any distinctions between kinds of sexual and reproductive bodies are importantly political and open to contest. Some authors have argued that we should acknowledge the continuum of anatomical differences and recognize at least five sexes (Fausto-Sterling). And if sexual distinctions become more complex, we would also need to rethink sexuality, given that sexual desire would not fit neatly within existing homosexual/heterosexual paradigms.

However, one can encourage the proliferation of sexual and reproductive options without maintaining that we can or should eliminate *all* social implications of anatomical sex and reproduction. Given that as a species there are substantial differences in what human bodies contribute to reproduction, and what sorts of bodies bear the main physical burdens of reproduction, and given further that reproduction cannot really help but be a socially significant fact (it does, after all, produce children), it can seem difficult to imagine a functioning society, more specifically, a functioning *feminist* society, that doesn't acknowledge in some way the difference between those kinds of bodies that are likely able to bear children, and those that aren't. [. .]

I will not debate here the degree to which a just society must be attentive to sexual and reproductive differences. Whether we, as feminists, ought to recommend the construction of (new) non-hierarchical genders or work to abolish gender entirely is a normative issue I leave for another occasion. Nonetheless, at the very least it would help to have terminology to debate these issues. I propose that we use the definitions of *man* and *woman* offered above: it is clear that these dominant nodes of our current gender structures are hierarchical. But borrowing strategies employed before, we can define gender in

generic terms under which the previous definitions of man and women fall; thus allowing the possibility of non-hierarchical genders and breaking the binary opposition between man and woman.

A group *G* is a *gender* relative to context *C* iff_{df} members of *G* are (all and only) those:

- (i) who are regularly observed or imagined to have certain bodily features presumed in *C* to be evidence of their reproductive capacities;
- (ii) whose having (or being imagined to have) these features marks them within the context of the ideology in *C* as motivating and justifying some aspect(s) of their social position; and
- (iii) whose satisfying (i) and (ii) plays (or would play) a role in *C* in their social position's having one or another of these designated aspects.

I offer this analysis as a way of capturing the standard slogan: gender is the social meaning of sex. Note, however, that in imagining "alternative" genders we should be careful not to take for granted that the relevant biological divisions will correspond to what we consider "sex." (Alternative groupings could include: "pregnant persons," "lactating persons," "menstruating persons," "infertile persons," [. . .]). Neither should we assume that membership in a gender will constitute one's personal or psychological identity to any significant degree. Recall that on the accounts of gender and race I am proposing, both are to be understood first and foremost as social groups defined within a structure of social relations; whatever links there might be to identities and norms are highly contingent and would depend on the details of the picture. For example, we might imagine that "after the revolution" gender is a component of one's overall social position because, for example, there are legal protections or medical entitlements granted to individuals classified as having a certain sort of "sexed" body; but this need not have broad implications for psychological identity or everyday social interactions, for the "sex" of bodies might not even be publicly marked.

Turning briefly to race, the parallel issue arises: Do we need a concept of non-hierarchical "races" in order to frame and debate different visions

of a "racially" just society? It would seem that we have the terminological resources available without a further definition: let races be, as previously defined, those hierarchically organized groups that are defined (roughly) by physical features and (assumed) geographical origins, and call those that aren't hierarchically organized (in the context in question) "ethnicities." Admittedly, ethnicity as we know it does have implications for social status and power, so my proposal is to employ the term for a somewhat idealized conception.

As in the case of gender, the question arises whether it ought to be part of an anti-racist project to recommend the preservation of existing ethnic groups or the formation of "new" ethnicities. And more generally, we need to ask whether a feminist anti-racism should treat genders and ethno-racial groups in the same way over the long term. Should we seek, e.g., to eliminate all genders and ethnora-

cial groupings; to preserve and proliferate them; to eliminate gender but not ethnicity (or vice versa)? These questions deserve careful attention but I cannot address them here.

8. Conclusion

On the accounts I've offered, there are striking parallels between race and gender. Both gender and race are real, and both are social categories. Neither gender nor race is chosen, but the forms they take can be resisted or mutated. Both race and gender (as we know it) are hierarchical, but the systems that sustain the hierarchy are contingent. And although the ideologies of race and gender and the hierarchical structures they sustain are substantively very different, they are intertwined.

There are many different types of human bodies; it is not the case that there is a unique "right" way of

classifying them, though certain classifications will be more useful for some purposes than others. How we classify bodies can and does matter politically, for our laws, social institutions, and personal identities are profoundly linked to understandings of the body and its possibilities. This is compatible with the idea that what possibilities a human body has is not wholly a function of our understandings of it. Our bodies often outdo us, and undo us, in spite of the meanings we give them.

Within the framework I've sketched, there is room for theoretical categories such as *man*, *woman*, and *race* (and particular racial groups), that take hierarchy to be a constitutive element, and those such as *gender* and *ethnicity* that do not. As I have suggested before, I am willing to grant that there are other ways to define race or gender, man or woman, that are useful to answer different questions, motivated by different concerns and priorities. I'm sure we need several concepts to do all the work needed in understanding the complex systems of racial and gender subordination.

In short, (speaking of my analyses) I'm less committed to saying that *this* is what gender is and what race is, than to saying that *these* are important categories that a feminist antiracist theory needs. [...] To return to the point made much earlier in characterizing analytic projects: it is our responsibility to define gender and race for our theoretical purposes. The world itself can't tell us what gender is. The same is true for race. [...] Of course, in defining our terms, we must keep clearly in mind our political aims both in analyzing the past and present, and in envisioning alternative futures. But rather than worrying, "what is gender, really?" or "what is race, really?" I think we should begin by asking (both in the theoretical and political sense) what, if anything, we want them to be.

Notes

1. On this I am deeply indebted to Stevens, Ch. 4, and Omi and Winant, esp. pp. 53–61.
2. We may want to allow there to be kinds of social stratification among ethnic groups that fall short

of the kind of systematic subordination constitutive of race. My account remains vague on this point.

Works Cited

- Ellis, T. *Village Voice*, 13 June 1989.
- Fausto-Sterling, A. "The Five Sexes: Why Male and Female Are Not Enough." *The Sciences* 33, 2 (1993): 20–4.
- Frye, M. *The Politics of Reality*. Freedom, CA: Crossing Press, 1983.
- Omi, M., and H. Winant. *Racial Formation in the United States*. New York: Routledge, 1994.
- Stevens, J. *Reproducing the State*. Princeton: Princeton University Press, 1999.
- Young, I. *Justice and the Politics of Difference*. Princeton: Princeton University Press, 1990.

Study Questions

- Haslanger outlines three different approaches that philosophers might take in response to the questions "What is gender?" and "What is race?" Compare these approaches using an example other than gender or race.
- On Haslanger's material feminist approach, it is within a matrix of social relations that we become an individual of a particular gender and race. Explain how this happens by attending to the specifics of Haslanger's account and using your own gender and race as an example.
- What norms and expectations does society prescribe for individuals of the same gender and race as you? Can you identify any ways in which these norms and expectations incorporate the oppressive relations Haslanger argues are basic to classification by gender and race?
- Haslanger suggests that "we should work to undermine those forces that make being a man, a woman, or a member of a racialized group possible; we should refuse to be gendered man or woman, refuse to be raced." Do you agree or disagree?

Michel Foucault

Plato

Euthyphro is considered to be one of Plato's (429–347 BCE) early dialogues, also known as the "Socratic" dialogues because in these we encounter the historical Socrates, a philosopher whom Aristotle described as focused on questions about how to lead a good life, preoccupied with seeking, though without success, definitions of ethical concepts—here, of "holiness" or "piety."

Euthyphro (or On Holiness, a Tentative Dialogue)

Characters: Euthyphro, Socrates

EUTHYPHRO. What strange thing has happened, Socrates, that you have left your accustomed haunts in the Lyceum and are now haunting the portico where the king archon sits? For it cannot be that you have an action before the king, as I have.

SOCRATES. Our Athenians, Euthyphro, do not call it an action, but an indictment.

E. What? Somebody has, it seems, brought an indictment against you; for I don't accuse you of having brought one against anyone else.

S. Certainly not.

E. But someone else against you?

S. Quite so.

E. Who is he?

S. I don't know the man very well myself, Euthyphro, for he seems to be a young and unknown person. His name, however, is Meletus, I believe. And he is of the deme of Pitthus, if you remember any Pitthian Meletus, with long hair and only a little beard, but with a hooked nose.

E. I don't remember him, Socrates. But what sort of an indictment has he brought against you?

S. What sort? No mean one, it seems to me: for the fact that, young as he is, he has apprehended so important a matter reflects no small credit upon him. For he says he knows how the youth are corrupted and who those are who corrupt them. He must be a wise man; who, seeing my lack of wisdom and that I am corrupting his fellows, comes to the State, as a boy runs to his mother, to accuse me. And he seems to me to be the only one of the public men who begins in the right way; for the right way is to take care of the young men first, to make them as good as possible, just as a good husbandman will naturally take care of the young plants first and afterwards of the rest. And so Meletus, perhaps, is first clearing away us who corrupt the young plants, as he says; then after this, when he has turned his attention to the older men, he will bring countless most precious blessings upon the State—at least, that is the natural outcome of the beginning he has made.

E. I hope it may be so, Socrates; but I fear the opposite may result. For it seems to me that he begins by injuring the State at its very heart when he undertakes to harm you. Now tell me, what does he say you do that corrupts the young?

S. Absurd things, my friend, at first hearing. For he says I am a maker of gods; and because I make new gods and do not believe in the old ones, he indicted me for the sake of these old ones, as he says.

E. I understand, Socrates; it is because you say the divine monitor keeps coming to you. So he has brought the indictment against you for making innovations in religion, and he is going into court to slander you, knowing that slanders on such subjects are readily accepted by the people. Why, they even laugh at me and say I am crazy when I say anything in the assembly about divine things and foretell the future to them. And yet there is not one of the things I have foretold that is not true; but they are jealous of all such men as you and I are. However, we must not be disturbed, but must come to close quarters with them.

S. My dear Euthyphro, their ridicule is perhaps of no consequence. For the Athenians, I fancy, are

E. I don't much desire to test their sentiments toward me in this matter.

S. No, for perhaps they think that you are reserved and unwilling to impart your wisdom. But I fear that because of my love of men they think that I not only pour myself out copiously to anyone and everyone without payment, but that I would even pay something myself, if anyone would listen to me. Now if, as I was saying just now, they were to laugh at me, as you say they do at you, it would not be at all unpleasant to pass the time in the court with jests and laughter; but if they are in earnest, then only soothsayers like you can tell how this will end.

E. Well, Socrates, perhaps it won't amount to much, and you will bring your case to a satisfactory ending, as I think I shall mine.

E. Prosecuting.

S. What is your case, Euthyphro? Are you defending or prosecuting?

E. Such a man that they think I am insane because I am prosecuting him.

S. Why? Are you prosecuting one who has wings to fly away with?

E. No flying for him at his ripe old age.

S. Who is he?

E. My father.

S. Your father, my dear man?

E. Certainly.

S. But what is the charge, and what is the suit about?

E. Murder. Socrates.

not much concerned if they think a man is clever, provided he does not impart his clever notions to others; but when they think he makes others to be like himself, they are angry with him, either through jealousy, as you say, or for some other reason.

S. Heracles! Surely, Euthyphro, most people do not know where the right lies; for I fancy it is not everyone who can rightly do what you are doing, but only one who is already very far advanced in wisdom.

E. Very far, indeed, Socrates, by Zeus.

S. Is the one who was killed by your father a relative? But of course he was, for you would not bring a charge of murder against him on a stranger's account.

E. It is ridiculous, Socrates, that you think it matters whether the man who was killed was a stranger or a relative, and do not see that the only thing to consider is whether the action of the slayer was justified or not, and that if it was justified one ought to let him alone, and if not, one ought to proceed against him, even if he share one's hearth and eat at one's table. For the pollution is the same if you associate knowingly with such a man and do not purify yourself and him by proceeding against him. In this case, the man who was killed was a hired workman of mine, and when we were farming at Naxos, he was working there on our land. Now he got drunk, got angry with one of our house slaves, and butchered him. So my father bound him hand and foot, threw him into a ditch, and sent a man here to Athens to ask the religious adviser what he ought to do. In the meantime he paid no attention to the man as he lay there bound, and neglected him, thinking that he was a murderer and it did not matter if he were to die. And that is just what happened to him. For he died of hunger and cold and his bonds before the messenger came back from the

adviser: Now my father and the rest of my relatives

are angry with me, because for the sake of this mur-

derer I am prosecuting my father for murder. For

they say he did not kill him, and if he had killed

him never so much, yet since the dead man was a

murderer, I ought not to trouble myself about such

a fellow because it is unholy for a son to prosecute

his father for murder. Which shows how little they

know what the divine law is in regard to holiness

and unholiness.

S. But, in the name of Zeus, Euthyphro, do you think

your knowledge about divine laws and holiness and

unholiness is so exact that, when the facts are as you say, you are not afraid of doing something unholy yourself in prosecuting your father for murder?

E. I should be of no use, Socrates, and Euthyphro would be in no way different from other men, if I did not have exact knowledge about all such things.

S. Then the best thing for me, my admirable Euthyphro, is to become your pupil and, before the suit with Meletus comes on, to challenge him and say that I always thought it very important before to know about divine matters and that now, since he says I am doing wrong by acting carelessly and making innovations in matters of religion, I have become your pupil. And "Meletus," I should say, "if you acknowledge that Euthyphro is wise in such matters, then believe that I also hold correct opinions, and do not bring me to trial; and if you do not acknowledge that, then bring a suit against him, my teacher, rather than against me, and charge him with corrupting the old, namely, his father and me, which he does by teaching me and by correcting and punishing his father." And if he does not do as I ask and does not release me from the indictment or bring it against you in my stead, I could say in the court the same things I said in my challenge to him, could I not?

E. By Zeus, Socrates, if he should undertake to indict me, I fancy I should find his weak spot, and it would be much more a question about him in court than about me.

S. And I, my dear friend, perceiving this, wish to become your pupil; for I know that neither this fellow Meletus, nor anyone else, seems to notice you at all, but he has seen through me so sharply and so easily that he has indicted me for impiety. Now in the name of Zeus, tell me what you just now asserted that you knew so well. What do you say is the nature of piety and impiety, both in relation to murder and to other things? Is not holiness always the same with itself in every action, and, on the other hand, is not unholiness the opposite of all

is to be unholy possessing some one characteristic quality?

E. Certainly, Socrates.

S. Tell me then, what do you say holiness is, and what unholiness?

E. Well then, I say that holiness is doing what I am doing now, prosecuting the wrongdoer who commits murder or steals from the temples or does any such thing, whether he be your father, or your mother or anyone else, and not prosecuting him is unholy. And, Socrates see what a sure proof I offer you—a proof I have already given to others—that this is established and right and that we ought not to let him who acts impiously go unpunished, no matter who he may be. Men believe that Zeus is the best and most just of the gods, and they acknowledge that he put his father in bonds because he wickedly devoured his children, and he in turn had mutilated his father for similar reasons; but they are incensed against me because I proceed against my father when he has done wrong, and so they are inconsistent in what they say about the gods and about me.

S. Is not this, Euthyphro, the reason why I am being prosecuted, because when people tell such stories about the gods I find it hard to accept them? And therefore, probably, people will say I am wrong. Now if you, who know so much about such things, accept these tales, I suppose I too must give way. For what am I to say, who confess frankly that I know nothing about them? But tell me, in the name of Zeus, the god of friendship, do you really believe these things happened?

E. Yes, and still more wonderful things than these, Socrates, which most people do not know.

S. And so you believe that there was really war between the gods, and fearful empires and battles and other things of the sort, such as are told of by the poets and represented in varied designs by the great artists in our sacred places and especially on the robe which is carried up to the Acropolis at the great Panathenaea, for this is covered with such representations? Shall we agree that these things are true, Euthyphro?

E. Not only these things, Socrates; but, as I said just now, I will, if you like, tell you many other things about the gods, which I am sure will amaze you when you hear them.

S. I dare say. But you can tell me those things at your leisure some other time. At present try to tell me more clearly what I asked you just now. For, my friend, you did not give me sufficient information before, when I asked what holiness was, but you told me that this was holy which you are now doing, prosecuting your father for murder.

E. Well, what I said was true, Socrates.

S. Perhaps. But, Euthyphro, you say that many other things are holy, do you not?

E. Why, so they are.

S. Now call to mind that this is not what I asked you, to tell me one or two of the many holy acts, but to tell the essential aspect, by which all holy acts are holy; for you said that all unholy acts were unholy and all holy ones holy by one aspect. Or don't you remember?

E. I remember.

S. Tell me then what this aspect is, that I may keep my eye fixed upon it and employ it as a model and, if anything you or anyone else does agrees with it, may say that the act is holy, and if not, that it is unholy.

E. If you wish me to explain in that way, I will do so. S. I do wish it.

E. Well then, what is clear to the gods is holy, and what is not dear to them is unholy.

S. Excellent, Euthyphro, now you have answered as I asked you to answer. However, whether it is true, I am not yet sure; but you will, of course, show that what you say is true.

E. Certainly.

S. Come then, let us examine our words. The thing and the person that are dear to the gods are holy, and the thing and the person that are hateful to the gods are unholy; and the two are not the same, but the

holy and the unholy are the exact opposites of each other. Is not this what we have said?

E. Yes, just this.

S. And it seems to be correct?

E. I think so, Socrates.

S. Well then, have we said this also, that the gods, Euthyphro, quarrel and disagree with each other, and that there is enmity between them?

E. Yes, we have said that.

S. But what things is the disagreement about, which causes enmity and anger? Let us look at it in this way. If you and I were to disagree about number, for instance, which of two numbers were the greater, would the disagreement about these matters make us enemies and make us angry with each other, or should we not quickly settle it by resorting to arithmetic?

E. Of course we should.

S. Then, too, if we were to disagree about the relative size of things, we should quickly put an end to the disagreement by measuring?

E. Yes.

S. And we should, I suppose, come to terms about relative weights by weighing?

E. Of course.

S. But about what would a disagreement be, which we could not settle and which would cause us to be enemies and be angry with each other? Perhaps you cannot give an answer offhand; but let me suggest it. Is it not about right and wrong, and noble and disgraceful, and good and bad? Are not these the questions about which you and I and other people become enemies, when we do become enemies, because we differ about them and cannot reach any satisfactory agreement?

E. Yes, Socrates, these are the questions about which we should become enemies.

S. And how about the gods, Euthyphro? If they disagree, would they not disagree about these questions?

E. Necessarily.

S. Then, my noble Euthyphro, according to what you say, some of the gods too think some things are right or wrong, and noble or disgraceful, and good or bad, and others disagree; for they would not quarrel with each other if they did not disagree about these matters. Is that the case?

E. You are right.

S. Then the gods in each group love the things which they consider good and right and hate the opposites of these things?

E. Certainly.

S. But you say that the same things are considered right by some of them and wrong by others; and it is because they disagree about these things that they quarrel and wage war with each other. Is not this what you said?

E. It is.

S. Then, as it seems, the same things are hated and loved by the gods, and the same things would be dear and hateful to the gods.

E. So it seems.

S. And then the same things would be both holy and unholy, Euthyphro, according to this statement.

E. I suppose so.

S. Then you did not answer my question, my friend. For I did not ask you what is at once holy and unholy; but, judging from your reply, what is dear to the gods is also hateful to the gods. And so, Euthyphro, it would not be surprising if, in punishing your father as you are doing, you were performing an act that is pleasing to Zeus, but hateful to Cronus and Uranus, and pleasing to Hephaestus, but hateful to Hera, and so forth in respect to the other gods, if any disagree with any other about it.

E. But I think, Socrates, that none of the gods disagrees with any other about this, or holds that he who kills anyone wrongfully ought not to pay the penalty.

S. Well, Euthyphro, to return to men, did you ever hear anybody arguing that he who had killed anyone wrongfully; or had done anything else whatever wrongfully, ought not to pay the penalty?

E. Why, they are always arguing these points, especially in the law courts. For they do very many wrong things, and then there is nothing they will not do or say, in defending themselves, to avoid the penalty.

S. Yes, but do they acknowledge, Euthyphro, that they have done wrong and, although they acknowledge it, nevertheless say that they ought not to pay the penalty?

satisfaction, I will go on, I live.

S. I understand; it is because you think I am slower to understand than the judges, since it is plain that you will show them that such acts are wrong and that all the gods hate them.

S. I then there is something they do not do and say. For they do not, I fancy, dare to say and argue that if they have really done wrong, they ought not to pay the penalty; but, I think, they say they have not done wrong: do they not?

E. You are right.

doer must not pay the penalty; but perhaps they argue about this, who is a wrongdoer, and what he did, and when.

E. That is true.

Then is not the same thing true of the gods, if they quarrel about right and wrong, as you say; and some say others have done wrong, and some say they have not? For surely, my friend, no one, either of gods or men, has the face to say that he who does wrong ought not to pay the penalty.

. Yes, you are right about this, Socrates, in the main.

But I think, Euthyphro, those who dispute, both men and gods, if the gods do dispute, dispute about which separate act. When they differ with one another about any act, some say it was right and others that was wrong. Is it not so?

Certainly

come now, my dear Euthyphro, inform me, that may be made wiser, what proof you have that all

L., what is to hinder, Socrates?

E. What is to hinder Socrates?

S. Nothing, so far as I am concerned, Euthyphro, but consider your own position, whether by adopting this definition you will most easily teach me what

I think that the man lost his life wrongfully,

E. Well, I should say that what all the gods love is holy and, on the other hand, what they all hate is unholy.

S. Then shall we examine this again, Euthyphro, to see if it is correct, or shall we let it go and accept our own statement, and those of others, agreeing that it is so if anyone merely says that it is? Or ought we to inquire into the correctness of the statement?

E. We ought to inquire. However, I think this is now correct.

S. We shall soon know more about this, my friend. Just consider this question: Is that which is holy loved by the gods because it is holy, or is it holy because it is loved by the gods?

E. I don't know what you mean, Socrates.

S. Then I will try to speak more clearly. We speak of being carried and of carrying, of being led and of leading, of being seen and of seeing; and you understand—do you not?—that in all such expressions the two parts differ one from the other in meaning, and how they differ.

E. I think I understand.

S. Then, too, we conceive of a thing being loved and of a thing loving, and the two are different?

E. Of course.

S. Now tell me, is a thing which is carried a carried thing because one carries it, or for some other reason?

E. No, for that reason.

S. And a thing which is led is led because one leads it, and a thing which is seen is so because one sees it?

E. Certainly.

S. Then one does not see it because it is a seen thing, but, on the contrary, it is a seen thing because one sees it; and one does not lead it because it is a led thing, but it is a led thing because one leads it; and one does not carry it because it is a carried thing, but it is a carried thing because one carries it. Is it clear, Euthyphro, what I am trying to say? I am trying to say this, that if anything becomes or undergoes, it

does not become because it is in a state of becoming, but it is in a state of becoming because it becomes, and it does not undergo because it is a thing which undergoes, but because it undergoes it is a thing which undergoes; or do you not agree to this?

E. I agree.

S. Is not that which is beloved a thing which is either becoming or undergoing something?

E. Certainly.

S. And is this case like the former ones: those who love it do not love it because it is a beloved thing, but it is a beloved thing because they love it?

E. Obviously.

S. Now what do you say about that which is holy, Euthyphro? It is loved by all the gods, is it not, according to what you said?

E. Yes.

S. For this reason, because it is holy, or for some other reason?

E. No, for this reason.

S. It is loved because it is holy, not holy because it is loved?

E. I think so.

S. Then that which is dear to the gods and that which is holy are not identical, but differ one from the other.

E. Of course.

S. But that which is dear to the gods is dear to them and beloved by them because they love it.

E. How so, Socrates?

S. Because we are agreed that the holy is loved because it is holy and that it is not holy because it is loved; are we not?

E. Yes.

S. But we are agreed that what is dear to the gods is dear to them because they love it, that is, by reason of this love, not that they love it because it is dear.

E. Very true.

S. But if that which is dear to the gods and that which is holy were identical, my dear Euthyphro, then if the holy were loved because it is holy, that which is dear to the gods would be loved because it is dear; and if that which is dear to the gods is dear because it is loved, then that which is holy would be holy because it is loved; but now you see that the opposite is the case, showing that the two are different from each other. For the one becomes lovable from the fact that it is loved, whereas the other is loved because it is in itself lovable. And, Euthyphro, it seems that when you were asked

what holiness is you were unwilling to make plain its essence, but you mentioned something that has happened to this holiness, namely, that it is loved by the gods. But you did not tell us yet what it really is. So, if you please, do not hide it from me, but begin over again and tell me what holiness is, no matter whether it is loved by the gods or anything else happens to it; for we shall not quarrel about that. But tell me frankly, What is holiness, and what is unholiness?

E. But, Socrates, I do not know how to say what I mean. For whatever statement we advance, somehow or other it moves about and won't stay where we put it.

S. Your statements, Euthyphro, are like works of my ancestor Daedalus,¹ and if I were the one who made or advanced them, you might laugh at me and say that on account of my relationship to him my works in words run away and won't stay where they are put. But now—well, the statements are yours, so some other jest is demanded; for they stay fixed, as you yourself see.

E. I think the jest does very well as it is; for I am not the one who makes these statements move about and

not stay in the same place, but you are the Daedalus, for they would have stayed, so far as I am concerned.

S. Apparently then, my friend, I am a more clever artist than Daedalus, inasmuch as he made only his own works move, whereas I, as it seems, give motion to the works of others as well as to my own. And the most

exquisite thing about my art is that I am clever against my will; for I would rather have my words stay fixed

and stable than possess the wisdom of Daedalus and the wealth of Tantalus besides. But enough of this.

Since you seem to be indolent, I will aid you myself, so that you may instruct me about holiness. And do not give it up beforehand. Just see whether you do not think that everything that is holy is right.

E. I do.

S. But is everything that is right also holy? Or is all which is holy right, and not all which is right holy, but part of it holy and part something else?

E. I can't follow you, Socrates.

S. And yet you are as much younger than I as you are wiser; but, as I said, you are indolent on account of your wealth of wisdom. But exert yourself, my friend; for it is not hard to understand what I mean. What I mean is the opposite of what the poet said, who wrote: "Zeus the creator, him who made all things, thou wilt not name; for where fear is, there also is reverence." Now I disagree with the poet. Shall I tell you how?

E. By all means.

S. It does not seem to me true that where fear is, there also is reverence; for many who fear diseases and poverty and other such things seem to me to fear, but not to reverence at all these things which they fear. Don't you think so, too?

E. Certainly.

S. But I think that where reverence is, there also is fear; for does not everyone who has a feeling of reverence and shame about any act also dread and fear the reputation for wickedness?

E. Yes, he does fear.

S. Then it is not correct to say "where fear is, there also is reverence." On the contrary, where reverence is, there also is fear; but reverence is not everywhere where fear is, since, as I think, fear is more comprehensive than reverence; for reverence is a part of fear just as the odd is a part of number, so that it is not true that where number is, there also is the odd, but

that where the odd is, there also is number. Perhaps you follow me now?

E. Perfectly.

S. It was something of this sort that I meant before, when I asked whether where the right is, there also is holiness, or where holiness is, there also is the right; but holiness is not everywhere where the right is, for holiness is a part of the right. Do we agree to this, or do you dissent?

E. No, I agree; for I think the statement is correct.

S. Now observe the next point. If holiness is a part of the right, we must, apparently, find out what part

of the right holiness is. Now if you asked me about one of the things I just mentioned, as, for example, what part of number the even was, and what kind of a number it was, I should say, "that which is not indivisible by two, but divisible by two"; or don't you agree?

E. I agree.

S. Now try in your turn to teach me what part of the right holiness is, that I may tell Meletus not to wrong me anymore or bring suits against me for impiety, since I have now been duly instructed by you about what is, and what is not, pious and holy. [. . .]

Note

1. Tr.: Socrates was the son of a sculptor and was himself educated to be a sculptor. This is doubtless the reason for his reference to Daedalus as an

ancestor. Daedalus was a half mythical personage whose statues were said to have been so lifelike that they moved their eyes and walked about.

Study Questions

1. Do you think that Euthyphro's father acted wrongly in the death of his servant? If so, do you think that Euthyphro is acting rightly or wrongly in bringing charges against his father?
2. Outline the several ways in which Euthyphro fails in his attempts to define what is holy and unholy, pious and impious?
3. How does Socrates suggest concepts be properly defined?
4. What is your response to Socrates' question, "Is that which is holy [good] loved by the gods because it is holy [good], or is it holy [good] because it is loved by the gods?"

John Stuart Mill

John Stuart Mill (1806–73), considered the most influential English-speaking philosopher of the nineteenth century, was born in London, the son of Scottish political theorist James Mill (1773–1836), who, with his friend Jeremy Bentham (1748–1832), founder of utilitarianism, home-schooled his son with the goal of cultivating the perfect utilitarian mind. Mill came to believe that this education had stunted his emotional growth, contributing to the severe depression he suffered at age 20. In *Utilitarianism*, Mill goes beyond Bentham's hedonistic calculus to count quality as well as quantity of pleasure as morally relevant.

What Utilitarianism Is

From *Utilitarianism*

1. Utility, or the Greatest Happiness Principle

The creed which accepts as the foundation of morals, Utility, or the Greatest Happiness Principle, holds that actions are right in proportion as they tend to promote happiness, wrong as they tend to produce the reverse of happiness. By happiness is intended pleasure and the absence of pain; by unhappiness, pain and the privation of pleasure. To give a clear view of the moral standard set up by the theory, much more requires to be said: in particular, what things it includes in the ideas of pain and pleasure, and to what extent this is left an open question. But these supplementary explanations do not affect the theory of life on which this theory of morality is grounded—namely, that pleasure and freedom from pain are the only things desirable as ends; and that all desirable things (which are as numerous in the utilitarian as in any other scheme) are desirable either for the pleasure inherent in themselves, or as means to the promotion of pleasure and the prevention of pain.

Now, such a theory of life excites in many minds, and among them in some of the most estimable in feeling and purpose, inveterate dislike. To suppose that life has (as they express it) no higher end than pleasure, no better and nobler object of desire and pursuit, they designate as utterly mean and grovelling—as a doctrine worthy only of swine, to

whom the followers of Epicurus were, at a very early period, contemptuously likened; and modern holders of the doctrine are occasionally made the subject of equally polite comparisons by its German, French, and English assailants.

When thus attacked, the Epicureans have always answered that it is not they but their accusers who represent human nature in a degrading light, since the accusation supposes human beings to be capable of no pleasures except those of which swine are capable. If this supposition were true, the charge could not be gainsaid but would then be no longer an imputation, for if the sources of pleasure were precisely the same to human beings and to swine, the rule of life which is good enough for the one would be good enough for the other. The comparison of the Epicurean life to that of beasts is felt as degrading precisely because a beast's pleasures do not satisfy a human being's conceptions of happiness. Human beings have faculties more elevated than the animal appetites and, when once made conscious of them, do not regard anything as happiness which does not include their gratification. I do not, indeed, consider the Epicureans to have been by any means faultless in drawing out their scheme of consequences from the utilitarian principle. To do this in any sufficient manner, many Stoic, as well as Christian, elements require to be included. But there is no known Epicurean theory of life which does not assign to the pleasures of the intellect, of the feelings

and imagination, and of the moral sentiments a much higher value as pleasures than to those of mere sensation. It must be admitted, however, that utilitarian writers in general have placed the superiority of mental over bodily pleasures chiefly in the greater permanency, safety, uncostliness, etc., of the former—that is, in their circumstantial advantages rather than in their intrinsic nature. And on all these points utilitarians have fully proved their case; but they might have taken the other, and, as it may be called, higher, ground with entire consistency. It is quite compatible with the principle of utility to recognize the fact that some kinds of pleasure are more desirable and more valuable than others. It would be absurd that while, in estimating all other things, quality is considered as well as quantity, the estimation of pleasures should be supposed to depend on quantity alone.

2. Higher and Lower Pleasures

If I am asked what I mean by difference of quality in pleasures, or what makes one pleasure more valuable than another, merely as a pleasure, except its being greater in amount, there is but one possible answer. Of two pleasures, if there be one to which all or almost all who have experience of both give a decided preference, irrespective of any feeling of moral obligation to prefer it, that is the more desirable pleasure. If one of the two is, by those who are competently acquainted with both, placed so far above the other that they prefer it, even though knowing it to be attended with a greater amount of discontent, and would not resign it for any quantity of the other pleasure which their nature is capable of, we are justified in ascribing to the preferred enjoyment a superiority in quality, so far outweighing quantity as to render it, in comparison, of small account.

Now it is an unquestionable fact that those who are equally acquainted with, and equally capable of appreciating and enjoying, both do give a most marked preference to the manner of existence which employs their higher faculties. Few human creatures would consent to be changed into any of the lower

animals for a promise of the fullest allowance of a beast's pleasures; no intelligent human being would consent to be a fool, no instructed person would be an ignoramus, no person of feeling and conscience would be selfish and base, even though they should be persuaded that the fool, the dunce, or the rascal is better satisfied with his lot than they are with theirs. They would not resign what they possess more than he for the most complete satisfaction of all the desires which they have in common with him. If they ever fancy they would, it is only in cases of unhappiness so extreme that to escape from it they would exchange their lot for almost any other, however undesirable in their own eyes. A being of higher faculties requires more to make him happy, is capable probably of more acute suffering, and certainly accessible to it at more points, than one of an inferior type; but in spite of these liabilities, he can never really wish to sink into what he feels to be a lower grade of existence. [. . .] Whoever supposes that this preference takes place at a sacrifice of happiness—that the superior being, in anything like equal circumstances, is not happier than the inferior—confounds the two very different ideas, of happiness and content. It is undisputable that the being whose capacities of enjoyment are low has the greatest chance of having them fully satisfied; and a highly endowed being will always feel that any happiness which he can look for, as the world is constituted, is imperfect. But he can learn to bear its imperfections, if they are at all bearable; and they will not make him envy the being who is indeed unconscious of the imperfections, but only because he feels not at all the good which those imperfections qualify. It is better to be a human being dissatisfied than a pig satisfied; better to be Socrates dissatisfied than a fool satisfied. And if the fool, or the pig, are of a different opinion, it is because they only know their own side of the question. The other party to the comparison knows both sides.

It may be objected that many who are capable of the higher pleasures, occasionally, under the influence of temptation, postpone them to the lower. But this is quite compatible with a full appreciation of the intrinsic superiority of the higher. Men often, from

infirmity of character, make their election for the nearer good, though they know it to be the less valuable—and this no less when the choice is between two bodily pleasures than when it is between bodily and mental. They pursue sensual indulgences to the injury of health, though perfectly aware that health is the greater good. It may be further objected that many who begin with youthful enthusiasm for everything noble, as they advance in years, sink into indolence and selfishness. But I do not believe that those who undergo this very common change voluntarily choose the lower description of pleasures in preference to the higher. I believe that before they devote themselves exclusively to the one, they have already become incapable of the other. Capacity for the nobler feelings is in most natures a very tender plant, easily killed not only by hostile influences but by mere want of sustenance; and in the majority of young persons it speedily dies away if the occupations to which their position in life has devoted them, and the society into which it has thrown them, are not favourable to keeping that higher capacity in exercise. [. . .]

From this verdict of the only competent judges, I apprehend there can be no appeal. On a question which is the best worth having of two pleasures, or which of two modes of existence is the most grateful to the feelings, apart from its moral attributes and from its consequences, the judgment of those who are qualified by knowledge of both, or, if they differ, that of the majority among them, must be admitted as final. And there needs be the less hesitation to accept this judgment respecting the quality of pleasures, since there is no other tribunal to be referred to even on the question of quantity. What means are there of determining which is the acutest of two pains, or the intensest of two pleasurable sensations, except the general suffrage of those who are familiar with both? Neither pains nor pleasures are homogeneous, and pain is always heterogeneous with pleasure. What is there to decide whether a particular pleasure is worth purchasing at the cost of a particular pain, except the feelings and judgment of the experienced? When, therefore, those feelings and judgment declare the pleasures derived from

the higher faculties to be preferable in *kind*, apart from the question of intensity, to those of which the animal nature disjoined from the higher faculties is susceptible, they are entitled on this subject to the same regard.

I have dwelt on this point as being a necessary part of a perfectly just conception of Utility or Happiness, considered as the directive rule of human conduct. But it is by no means an indispensable condition to the acceptance of the utilitarian standard; for that standard is not the agent's own greatest happiness but the greatest amount of happiness altogether; and if it may possibly be doubted whether a noble character is always the happier for its nobleness, there can be no doubt that it makes other people happier, and that the world in general is immensely a gainer by it. Utilitarianism, therefore, could only attain its end by the general cultivation of nobleness of character, even if each individual were only benefited by the nobleness of others, and his own, so far as happiness is concerned, were a sheer deduction from the benefit. But the bare enunciation of such an absurdity as this last renders refutation superfluous.

3. Utility as the Standard of Morality

According to the Greatest Happiness Principle, as above explained, the ultimate end, with reference to and for the sake of which all other things are desirable (whether we are considering our own good or that of other people), is an existence exempt as far as possible from pain and as rich as possible in enjoyments, both in point of quantity and quality—the test of quality and the rule for measuring it against quantity being the preference felt by those who in their opportunities of experience, to which must be added their habits of self-consciousness and self-observation, are best furnished with the means of comparison. This being, according to the utilitarian opinion, the end of human action is necessarily also the standard of morality, which may accordingly be defined, the rules and precepts for human conduct. [. . .]

Unquestionably it is possible to do without happiness; it is done involuntarily by nineteen-twentieths of mankind, even in those parts of our present world which are least deep in barbarism; and it often has to be done voluntarily by the hero or the martyr for the sake of something which he prizes more than his individual happiness. But this something, what is it, unless the happiness of others, or some of the requisites of happiness? It is noble to be capable of resigning entirely one's own portion of happiness, or chances of it: but, after all, this self-sacrifice must be for some end; it is not its own end; and if we are told that its end is not happiness but virtue, which is better than happiness, I ask, would the sacrifice be made if the hero or martyr did not believe that it would earn for others immunity from similar sacrifices? Would it be made if he thought that his renunciation of happiness for himself would produce no fruit for any of his fellow creatures but to make their lot like his and place them also in the condition of persons who have renounced happiness? All honour to those who can abnegate for themselves the personal enjoyment of life when by such renunciation they contribute worthily to increase the amount of happiness in the world; but he who does it, or professes to do it, for any other purpose is no more deserving of admiration than the ascetic mounted on his pillar. He may be an inspiriting proof of what men *can* do, but assuredly not an example of what they *should*. [. . .]

I must again repeat what the assailants of utilitarianism seldom have the justice to acknowledge, that the happiness which forms the utilitarian standard of what is right in conduct is not the agent's own happiness but that of all concerned. As between his own happiness and that of others, utilitarianism requires him to be as strictly impartial as a disinterested and benevolent spectator. In the golden rule of Jesus of Nazareth, we read the complete spirit of the ethics of utility. To do as you would be done by, and to love your neighbour as yourself, constitute the ideal perfection of utilitarian morality. As the means of making the nearest approach to this ideal, utility would enjoin: first, that laws and social arrangements should place the happiness, or (as speaking

practically it may be called) the interest, of every individual as nearly as possible in harmony with the interest of the whole; and secondly, that education and opinion, which have so vast a power over human character, should so use that power as to establish in the mind of every individual an indissoluble association between his own happiness and the good of the whole—especially between his own happiness and the practice of such modes of conduct, negative and positive, as regard for the universal happiness prescribes, so that not only he may be unable to conceive the possibility of happiness to himself consistently with conduct opposed to the general good but also that a direct impulse to promote the general good may be in every individual one of the habitual motives of action, and the sentiments connected therewith may fill a large and prominent place in every human being's sentient existence. If the impugners of the utilitarian morality represented it to their own minds in this its true character, I know not what recommendation possessed by any other morality they could possibly affirm to be wanting to it: what more beautiful or more exalted developments of human nature any other ethical system can be supposed to foster, or what springs of action not accessible to the utilitarian such systems rely on for giving effect to their mandates.

4. Responses to Objections

The objectors to utilitarianism cannot always be charged with representing it in a discreditable light. On the contrary, those among them who entertain anything like a just idea of its disinterested character sometimes find fault with its standard as being too high for humanity. They say it is exacting too much to require that people shall always act from the inducement of promoting the general interests of society. But this is to mistake the very meaning of a standard of morals and confound the rule of action with the motive of it. It is the business of ethics to tell us what are our duties, or by what test we may know them; but no system of ethics requires that the sole motive of all we do shall be a feeling of duty; on the contrary, ninety-nine-hundredths

of all our actions are done from other motives, and rightly so done, if the rule of duty does not condemn them. It is the more unjust to utilitarianism that this particular misapprehension should be made a ground of objection to it, inasmuch as utilitarian moralists have gone beyond almost all others in affirming that the motive has nothing to do with the morality of the action, though much with the worth of the agent. He who saves a fellow creature from drowning does what is morally right, whether his motive be duty or the hope of being paid for his trouble; he who betrays the friend that trusts him is guilty of a crime, even if his object be to serve another friend to whom he is under greater obligations. But to speak only of actions done from the motive of duty and in direct obedience to principle: it is a misapprehension of the utilitarian mode of thought to conceive it as implying that people should fix their minds upon so wide a generality as the world, or society at large. The great majority of good actions are intended not for the benefit of the world but for that of individuals, of which the good of the world is made up; and the thoughts of the most virtuous man need not on these occasions travel beyond the particular persons concerned, except so far as is necessary to assure himself that in benefiting them he is not violating the rights, that is, the legitimate and authorized expectations, of anyone else. The multiplication of happiness is, according to the utilitarian ethics, the object of virtue: the occasions on which any person (except one in a thousand) has it in his power to do this on an extended scale, in other words to be a public benefactor, are but exceptional; and on these occasions alone is he called on to consider public utility; in every other case, private utility, the interest or happiness of some few persons, is all he has to attend to. Those alone the influence of whose actions extends to society in general need concern themselves habitually about so large an object. [. . .]

The same considerations dispose of another reproach against the doctrine of utility, founded on a still grosser misconception of the purpose of a standard of morality and of the very meaning of the words right and wrong. It is often affirmed that utilitarianism

renders men cold and unsympathizing; that it chills their moral feelings towards individuals; that it makes them regard only the dry and hard consideration of the consequences of actions, not taking into their moral estimate the qualities from which those actions emanate. If the assertion means that they do not allow their judgment respecting the rightness or wrongness of an action to be influenced by their opinion of the qualities of the person who does it, this is a complaint not against utilitarianism but against having any standard of morality at all; for certainly no known ethical standard decides an action to be good or bad because it is done by a good or a bad man, still less because done by an amiable, a brave, or a benevolent man, or the contrary. These considerations are relevant not to the estimation of actions but of persons, and there is nothing in the utilitarian theory inconsistent with the fact that there are other things which interest us in persons besides the rightness and wrongness of their actions. The Stoics, indeed, with the paradoxical misuse of language which was part of their system, and by which they strove to raise themselves above all concern about anything but virtue, were fond of saying that he who has that has everything; that he, and only he, is rich, is beautiful, is a king. But no claim of this description is made for the virtuous man by the utilitarian doctrine. Utilitarians are quite aware that there are other desirable possessions and qualities besides virtue, and are perfectly willing to allow to all of them their full worth. They are also aware that a right action does not necessarily indicate a virtuous character, and that actions which are blameable often proceed from qualities entitled to praise. When this is apparent in any particular case, it modifies their estimation, not certainly of the act but of the agent. I grant that they are, notwithstanding, of opinion that in the long run the best proof of a good character is good actions, and resolutely refuse to consider any mental disposition as good of which the predominant tendency is to produce bad conduct. This makes them unpopular with many people; but it is an unpopularity which they must share with everyone who regards the distinction between right and wrong in a serious light; and the reproach is not one which a conscientious utilitarian need be anxious to repel. [. . .]

It may not be superfluous to notice a few more of the common misapprehensions of utilitarian ethics, even those which are so obvious and gross that it might appear impossible for any person of candour and intelligence to fall into them—since persons, even of considerable mental endowments, often give themselves so little trouble to understand the bearings of any opinion against which they entertain a prejudice, and men are in general so little conscious of this voluntary ignorance as a defect that the vulgarest misunderstandings of ethical doctrines are continually met with in the deliberate writings of persons of the greatest pretensions both to high principle and to philosophy. We not uncommonly hear the doctrine of utility inveighed against as a godless doctrine. If it be necessary to say anything at all against so mere an assumption, we may say that the question depends upon what idea we have formed of the moral character of the Deity. If it be a true belief that God desires, above all things, the happiness of his creatures, and that this was his purpose in their creation, utility is not only not a godless doctrine but more profoundly religious than any other. If it be meant that utilitarianism does not recognize the revealed will of God as the supreme law of morals, I answer that a utilitarian who believes in the perfect goodness and wisdom of God necessarily believes that whatever God has thought fit to reveal on the subject of morals must fulfil the requirements of utility in a supreme degree. But others besides utilitarians have been of opinion that the Christian revelation was intended, and is fitted, to inform the hearts and minds of mankind with a spirit which should enable them to find for themselves what is right and incline them to do it when found, rather than to tell them, except in a very general way, what it is; and that we need a doctrine of ethics, carefully followed out, to interpret to us the will of God. Whether this opinion is correct or not, it is superfluous here to discuss since whatever aid religion, either natural or revealed, can afford to ethical investigation is as open to the utilitarian moralist as to any other. He can use it as the testimony of God to the usefulness or hurtfulness of any given course of action by as

good a right as others can use it for the indication of a transcendental law having no connection with usefulness or with happiness.

Again, Utility is often summarily stigmatized as an immoral doctrine by giving it the name of Expediency, and taking advantage of the popular use of that term to contrast it with Principle. But the Expedient, in the sense in which it is opposed to the Right, generally means that which is expedient for the particular interest of the agent himself, as when a minister sacrifices the interest of his country to keep himself in place. When it means anything better than this, it means that which is expedient for some immediate object, some temporary purpose, but which violates a rule whose observance is expedient in a much higher degree. The Expedient, in this sense, instead of being the same thing with the useful, is a branch of the hurtful. Thus, it would often be expedient, for the purpose of getting over some momentary embarrassment, or attaining some object immediately useful to ourselves or others, to tell a lie. But inasmuch as the cultivation in ourselves of a sensitive feeling on the subject of veracity is one of the most useful, and the enfeeblement of that feeling one of the most hurtful, things to which our conduct can be instrumental; and inasmuch as any, even unintentional, deviation from truth does that much towards weakening the trustworthiness of human assertion, which is not only the principal support of all present social well-being but the insufficiency of which does more than any one thing that can be named to keep back civilization, virtue, everything on which human happiness on the largest scale depends, we feel that the violation for a present advantage of a rule of such transcendent expediency is not expedient and that he who for the sake of a convenience to himself or to some other individual does what depends on him to deprive mankind of the good and inflict upon them the evil involved in the greater or less reliance which they can place in each other's word acts the part of one of their worst enemies. Yet, that even this rule, sacred as it is, admits of possible exceptions is acknowledged by all moralists: the chief of which is when the withholding of some fact (as of information from a malefactor, or

of bad news from a person dangerously ill) would save an individual (especially an individual other than oneself) from great and unmerited evil, and when the withholding can only be effected by denial. But in order that the exception may not extend itself beyond the need, and may have the least possible effect in weakening reliance on veracity, it ought to be recognized, and, if possible, its limits defined; and if the principle of utility is good for anything, it must be good for weighing these conflicting utilities against one another, and marking out the region within which one or the other preponderates.

Again, defenders of utility often find themselves called upon to reply to such objections as this—that there is not time, previous to action, for calculating and weighing the effects of any line of conduct on the general happiness. This is exactly as if anyone were to say that it is impossible to guide our conduct by Christianity because there is not time, on every occasion on which anything has to be done, to read through the Old and New Testaments. The answer to the objection is that there has been ample time, namely, the whole past duration of the human species. During all that time, mankind have been learning by experience the tendencies of actions—on which experience all the prudence, as well as all the morality of life, are dependent. People talk as if the commencement of this course of experience had hitherto been put off, and as if at the moment when some man feels tempted to meddle with the property or life of another, he had to begin considering for the first time whether murder and theft are injurious to human happiness. Even then I do not think that he would find the question very puzzling; but, at all events, the matter is now done to his hand. It is truly a whimsical supposition that if mankind were agreed in considering utility to be the test of morality, they would remain without any agreement as to what is useful, and would take no measures for having their notions on the subject taught to the young and enforced by law and opinion. There is no difficulty in proving any ethical standard whatever to work ill if we suppose universal idiocy to be conjoined with it; but on any hypothesis short of that, mankind must by this time have acquired positive beliefs as

to the effects of some actions on their happiness; and the beliefs which have thus come down are the rules of morality for the multitude, and for the philosopher until he has succeeded in finding better. That philosophers might easily do this, even now, on many subjects; that the received code of ethics is by no means of divine right; and that mankind have still much to learn as to the effects of actions on the general happiness, I admit, or rather, earnestly maintain. The corollaries from the principle of utility, like the precepts of every practical art, admit of indefinite improvement, and in a progressive state of the human mind, their improvement is perpetually going on. But to consider the rules of morality as improvable is one thing; to pass over the intermediate generalizations entirely, and endeavour to test each individual action directly by the first principle, is another. It is a strange notion that the acknowledgment of a first principle is inconsistent with the admission of secondary ones. To inform a traveller respecting the place of his ultimate destination is not to forbid the use of landmarks and direction-posts on the way. The proposition that happiness is the end and aim of morality does not mean that no road ought to be laid down to that goal, or that persons going thither should not be advised to take one direction rather than another. [. . .]

The remainder of the stock arguments against utilitarianism mostly consist in laying to its charge the common infirmities of human nature and the general difficulties which embarrass conscientious persons in shaping their course through life. We are told that a utilitarian will be apt to make his own particular case an exception to moral rules, and, when under temptation, will see a utility in the breach of a rule greater than he will see in its observance. But is utility the only creed which is able to furnish us with excuses for evil doing and means of cheating our own conscience? They are afforded in abundance by all doctrines which recognize as a fact in morals the existence of conflicting considerations—which all doctrines do that have been believed by sane persons. It is not the fault of any creed but of the complicated nature of human affairs that rules of conduct cannot be so framed as

to require no exceptions, and that hardly any kind of action can safely be laid down as either always obligatory or always condemnable. There is no ethical creed which does not temper the rigidity of its laws by giving a certain latitude, under the moral responsibility of the agent, for accommodation to peculiarities of circumstances; and under every creed, at the opening thus made, self-deception and dishonest casuistry get in. There exists no moral system under which there do not arise unequivocal cases of conflicting obligation. These are the real difficulties, the knotty points both in the theory of

ethics and in the conscientious guidance of personal conduct. They are overcome practically, with greater or with less success, according to the intellect and virtue of the individual; but it can hardly be pretended that anyone will be the less qualified for dealing with them from possessing an ultimate standard to which conflicting rights and duties can be referred. If utility is the ultimate source of moral obligations, utility may be invoked to decide between them when their demands are incompatible. Though the application of the standard may be difficult, it is better than none at all. [. . .]

Study Questions

1. For Mill, what makes an action a good action, and what is good in itself?
2. How are Bentham's and Mill's different approaches to utilitarianism illustrated by their respective words: "Pushpin is as good as poetry" (Bentham) and "It is better to be a human being dissatisfied than a pig satisfied; better to be Socrates dissatisfied than a fool satisfied" (Mill)? Do you find Mill's justification of his position convincing?
3. Why does Mill consider the principle of utility to provide an impartial and disinterested standard of morality? How does he answer the objection that such a standard is "too high for humanity"?
4. It is often feared that without God, there is no morality: in Dostoyevsky's words, "everything is permitted." How does Mill respond to the criticism that utilitarianism is a "godless doctrine"?

...that lie means?

Immanuel Kant

Immanuel Kant (1724–1804) rejected the moral sentimentalism of his Scottish empiricist contemporaries (e.g., Hume): for Kant, actions motivated by feelings have no moral worth. Kant held that the highest principle of practical philosophy is the moral law, which depends only on the nature of reason. By means of our (human) capacity for reason, we discover what the moral law requires of us. Also called *Groundwork of the Metaphysic of Morals*, this is Kant's foundational work on morality.

From *Fundamental Principles of the Metaphysic of Ethics*

First Section. Transition from the Common Rational Knowledge of Morality to the Philosophical

Nothing can possibly be conceived in the world, or even out of it, which can be called good without

qualification except a Good Will. Intelligence, wit, judgment, and the other *talents* of the mind, however they may be named, or courage, resolution, perseverance, as qualities of temperament, are undoubtedly good and desirable in many respects; but these gifts of nature may also become extremely

bad and mischievous if the will which is to make use of them, and which, therefore, constitutes what is called *character*, is not good. It is the same with the gifts of fortune. Power, riches, honour, even health and the general well-being and contentment with one's condition which is called *happiness*, inspire pride and often presumption if there is not a good will to correct the influence of these on the mind, and with this also to rectify the whole principle of acting and adapt it to its end. The sight of a being who is not adorned with a single feature of a pure and good will enjoying unbroken prosperity can never give pleasure to an impartial rational spectator. Thus a good will appears to constitute the indispensable condition even of being worthy of happiness.

There are even some qualities which are of service to this good will itself and may facilitate its action, yet which have no intrinsic unconditional value but always presuppose a good will, and this qualifies the esteem that we justly have for them and does not permit us to regard them as absolutely good. Moderation in the affections and passions, self-control, and calm deliberation are not only good in many respects but even seem to constitute part of the intrinsic worth of the person; but they are far from deserving to be called good without qualification, although they have been so unconditionally praised by the ancients. For without the principles of a good will, they may become extremely bad, and the coolness of a villain not only makes him far more dangerous but also directly makes him more abominable in our eyes than he would have been without it.

A good will is good not because of what it performs or effects, not by its aptness for the attainment of some proposed end, but simply by virtue of the volition—that is, it is good in itself, and considered by itself is to be esteemed much higher than all that can be brought about by it in favour of any inclination, nay, even of the sum-total of all inclinations. Even if it should happen that owing to special disfavour of fortune or the niggardly provision of a step-motherly nature, this will should wholly lack power to accomplish its purpose, if with its greatest efforts it should yet achieve nothing and

there should remain only the good will (not, to be sure, a mere wish, but the summoning of all means in our power), then, like a jewel, it would still shine by its own light, as a thing which has its whole value in itself. [. . .]

We have then to develop the notion of a will which deserves to be highly esteemed for itself and standing requiring rather to be cleared up than to be taught, and which in estimating the value of our actions always takes the first place and constitutes the condition of all the rest. In order to do this we will take the notion of duty, which includes that of a good will, although implying certain subjective restrictions and hindrances. These, however, far rather bring it out by contrast and make it shine forth so much the brighter. [. . .]

[I]t is a duty to maintain ones' life, and, in addition, everyone has also a direct inclination to do so. But on this account the often anxious care which most men take for it has no intrinsic worth, and their maxim has no moral import. They preserve their life as duty requires, no doubt, but not because duty requires. On the other hand, if adversity and hopeless sorrow have completely taken away the relish for life, if the unfortunate one, strong in mind, indignant at his fate rather than desponding or dejected, wishes for death, and yet preserves his life without loving it—not from inclination or fear but from duty—then his maxim has a moral worth.

To be beneficent when we can is a duty; and besides this, there are many minds so sympathetically constituted that, without any other motive of vanity or self-interest, they find a pleasure in spreading joy around them and can take delight in the satisfaction of others so far as it is their own work. But if maintaining that in such a case an action of this kind, however proper, however amiable it may be, has nevertheless no true moral worth, [. . .] deserving[ing] praise and encouragement but not esteem. For the maxim lacks the moral import, namely, that such actions be done from duty, not from inclination. Put the case that the mind of that philanthropist were clouded by sorrow

of his own, extinguishing all sympathy with the lot of others, and that while he still has the power to benefit others in distress, he is not touched by their trouble because he is absorbed with his own; and now suppose that he tears himself out of this dead insensibility and performs the action without any inclination to it but simply from duty, then first has his action its genuine moral worth. Further still, if nature has put little sympathy in the heart of this or that man; if he, supposed to be an upright man, is by temperament cold and indifferent to the sufferings of others, perhaps because in respect of his own he is provided with the special gift of patience and fortitude, and supposes, or even requires, that others should have the same—and such a man would certainly not be the meanest product of nature; but if nature had not specially framed him for a philanthropist, would he not still find in himself a source from whence to give himself a far higher worth than that of a good-natured temperament could be? Unquestionably. It is just in this that the moral worth of the character is brought out which is incomparably the highest of all, namely, that he is beneficent not from inclination but from duty. [. . .]

The second¹ proposition is: That an action done from duty derives its moral worth *not from the purpose which is to be attained by it but from the maxim by which it is determined*, and therefore does not depend on the realization of the object of the action but merely on the principle of volition by which the action has taken place without regard to any object of desire. It is clear from what precedes that the purposes which we may have in view in our actions, or their effects regarded as ends and springs of the will, cannot give to actions any unconditional or moral worth. In what, then, can their worth lie if it is not to consist in the will and in reference to its expected effect? It cannot lie anywhere but in the principle of the will without regard to the ends which can be attained by the action. For the will stands between its a priori principle, which is formal, and its a posteriori spring, which is material, as between two roads, and as it must be determined by something, it follows that it must be determined by the formal

principle of volition when an action is done from duty, in which case every material principle has been withdrawn from it.

The third proposition, which is a consequence of the two preceding, I would express thus: Duty is the necessity of acting from respect for the law. I may have inclination for an object as the effect of my proposed action but I cannot have respect for it, just for this reason, that it is an effect and not an energy of will. Similarly, I cannot have respect for inclination, whether my own or another's; I can at most, if my own, approve it; if another's, sometimes even love it, i.e., look on it as favourable to my own interest. It is only what is connected with my will as a principle, by no means as an effect—what does not subserve my inclination but overpowers it, or at least in case of choice excludes it from its calculation—in other words, simply the law of itself, which can be an object of respect and hence a command. Now an action done from duty must wholly exclude the influence of inclination and, with it, every object of the will, so that nothing remains which can determine the will except objectively the law and subjectively pure respect for this practical law, and consequently the maxim that I should follow this law even to the thwarting of all my inclinations.

Thus the moral worth of an action does not lie in the effect expected from it, nor in any principle of action which requires to borrow its motive from this expected effect. For all these effects—agreeableness of one's condition and even the promotion of the happiness of others—could have been also brought about by other causes, so that for this there would have been no need of the will of a rational being, whereas it is in this alone that the supreme and unconditional good can be found. The preeminent good which we call moral can therefore consist in nothing else than the conception of law in itself, which certainly is only possible in a rational being insofar as this conception, and not the expected effect, determines the will. This is a good which is already present in the person who acts accordingly, and we have not to wait for it to appear first in the result. [. . .]

Second Section. Transition from Popular Moral Philosophy to the Metaphysic of Morals

[...]

The conception of an objective principle, insofar as it is obligatory for a will, is called a command (of reason), and the formula of the command is called an imperative. [...]

Now all imperatives command either hypothetically or categorically. The former represent the practical necessity of a possible action as means to something else that is willed (or at least which one might possibly will). The categorical imperative would be that which represented an action as necessary of itself without reference to another end, i.e., as objectively necessary. [...]

We will first inquire whether the mere conception of a categorical imperative may not perhaps supply us also with the formula of it, containing the proposition which alone can be a categorical imperative. [...]

When I conceive a hypothetical imperative, in general I do not know beforehand what it will contain until I am given the condition. But when I conceive a categorical imperative, I know at once what it contains. For as the imperative contains besides the law only the necessity that the maxims shall conform to this law, while the law contains no conditions restricting it, there remains nothing but the general statement that the maxim of the action should conform to a universal law, and it is this conformity alone that the imperative properly represents as necessary.

There is therefore but one categorical imperative, namely this: *Act only on that maxim whereby thou const at the same time will that it should become a universal law.*

Now if all imperatives of duty can be deduced from this one imperative as from their principle, then, although it should remain undecided whether what is called duty is not merely a vain notion, yet at least we shall be able to show what we understand by it and what this notion means.

Since the universality of the law according to which effects are produced constitutes what is

properly called *nature* in the most general sense (as to form), that is, the existence of things so far as it is determined by general laws, that imperative of duty may be expressed thus: *Act as if the maxim of thy action were to become by thy will a Universal Law of Nature.*

We will now enumerate a few duties, adopting the usual division of them into duties to ourselves and to others, and into perfect and imperfect duties.²

1. A man reduced to despair by a series of misfortunes feels wearied of life, but is still so far in possession of his reason that he can ask himself whether it would not be contrary to his duty to himself to take his own life. Now he inquires whether the maxim of his action could become a universal law of nature. His maxim is: From self-love I adopt it as a principle to shorten my life when its longer duration is likely to bring more evil than satisfaction. It is asked then simply whether this principle founded on self-

love can become a universal law of nature. Now we see at once that a system of nature of which it should be a law to destroy life by means of the very feeling whose special nature it is to impel to the improvement of life would contradict itself, and therefore could not exist as a system of nature; hence that maxim cannot possibly exist as a universal law of nature, and consequently would be wholly inconsistent with the supreme principle of all duty.

2. Another finds himself forced by necessity to borrow money. He knows that he will not be able to repay it, but sees also that nothing will be lent to him unless he promises stoutly to repay it in a definite time. He desires to make this promise, but he has still so much conscience as to ask himself: Is it not unlawful and inconsistent with duty to get out of a difficulty in this way? Suppose, however, that he resolves to do so, then the maxim of his action would be expressed thus: When I think myself in want of money, I will borrow money and promise to repay it, although I know that I never can do so. Now this principle

of self-love or of one's own advantage may perhaps be consistent with my whole future welfare; but the question now is, is it right? I change then the suggestion of self-love into a universal law, and state the question thus: How would it be if my maxim were a universal law? Then I see at once that it could never hold as a universal law of nature but would necessarily contradict itself. For supposing it to be a universal law that everyone when he thinks himself in a difficulty should be able to promise whatever he pleases with the purpose of not keeping his promise, the promise itself would become impossible, as well as the end that one might have in view in it, since no one would consider that anything was promised to him but would ridicule all such statements as vain pretences.

3. A third finds in himself a talent which with the help of some culture might make him a useful man in many respects. But he finds himself in comfortable circumstances, and prefers to indulge in pleasure rather than to take pains in enlarging and improving his happy natural capacities. He asks, however, whether his maxim of neglect of his natural gifts, besides agreeing with his inclination to indulgence, agrees also with what is called duty. He sees then that a system of nature could indeed subsist with such a universal law, although men (like the South Sea islanders) should let their talents rust and resolve to devote their lives merely to idleness, amusement, and propagation of their species—in a word, to enjoyment; but he cannot possibly will that this should be a universal law of nature, or be implanted in us as such by a natural instinct. For, as a rational being, he necessarily wills that his faculties be developed, since they serve him, and have been given him, for all sorts of possible purposes.

4. A fourth, who is in prosperity, while he sees that others have to contend with great wretchedness and that he could help them, thinks: What concern is it of mine? Let everyone be as happy as heaven pleases, or as he can make himself; I will take nothing from him nor even envy him, only I do not wish to contribute anything to his welfare or to his assistance in distress! Now no doubt if

such a mode of thinking were a universal law, the human race might very well subsist, and doubtless even better than in a state in which everyone talks of sympathy and good-will, or even takes care occasionally to put it into practice, but, on the other side, also cheats when he can, betrays the rights of men, or otherwise violates them. But although it is possible that a universal law of nature might exist in accordance with that maxim, it is impossible to will that such a principle should have the universal validity of a law of nature. For a will which resolved this would contradict itself inasmuch as many cases might occur in which one would have need of the love and sympathy of others, and in which, by such a law of nature sprung from his own will, he would deprive himself of all hope of the aid he desires.

These are a few of the many actual duties, or at least what we regard as such, which obviously fall into two classes on the one principle that we have laid down. We must be able to will that a maxim of our action should be a universal law. This is the canon of the moral appreciation of the action generally. Some actions are of such a character that their maxim cannot without contradiction be even conceived as a universal law of nature, far from it being possible that we should will that it should be so. In others this intrinsic impossibility is not found, but still it is impossible to will that their maxim should be raised to the universality of a law of nature, since such a will would contradict itself. It is easily seen that the former violate strict or rigorous (inflexible) duty, the latter only laxer (meritorious) duty. Thus it has been completely shown by these examples how all duties depend as regards the nature of the obligation (not the object of the action) on the same principle. [. . .]

The question then is this: Is it a necessary law for all rational beings that they should always judge of their actions by maxims of which they can themselves will that they should serve as universal laws? If it is so, then it must be connected (altogether a priori) with the very conception of the will of a rational being generally. But in order to discover this connection we must, however reluctantly, take a step

into metaphysic, although into a domain of it which is distinct from speculative philosophy, namely, the metaphysic of morals. [...]

The will is conceived as a faculty of determining oneself to action *in accordance with the conception of certain laws*. And such a faculty can be found only in rational beings. Now that which serves the will as the objective ground of its self-determination is the *end*, and if this is assigned by reason alone, it must hold for all rational beings. On the other hand, that which merely contains the ground of possibility of the action of which the effect is the end, this is called the *means*. The subjective ground of the desire is the *spring*; the objective ground of the volition is the motive: hence the distinction between subjective ends which rest on springs and objective ends which depend on motives valid for every rational being. Practical principles are *formal* when they abstract from all subjective ends; they are *material* when they assume these, and therefore particular springs of action. The ends which a rational being proposes to himself at pleasure as *effects* of his actions (material ends) are all only relative, for it is only their relation to the particular desires of the subject that gives them their worth, which therefore cannot furnish principles universal and necessary for all rational beings and for every volition, that is to say, practical laws. Hence all these relative ends can give rise only to hypothetical imperatives.

Supposing, however, that there were something whose existence has *in itself* an absolute worth, something which, being *an end in itself*, could be a source of definite laws, then in this and this alone would lie the source of a possible categorical imperative, i.e., a practical law.

Now I say: man and generally any rational being exists as an end in himself, *not merely as a means* to be arbitrarily used by this or that will, but in all his actions, whether they concern himself or other rational beings, must be always regarded at the same time as an end. All objects of the inclinations have only a conditional worth, for if the inclinations and the wants founded on them did not exist, then their object would be without value. But the inclinations themselves, being sources of

want, are so far from having an absolute worth for which they should be desired that, on the contrary, it must be the universal wish of every rational being to be wholly free from them. Thus the worth of any object which is *to be acquired* by our action is always conditional. Beings whose existence depends not on our will but on nature's have nevertheless, if they are irrational beings, only a relative value as means, and are therefore called *things*; rational beings, on the contrary, are called *persons* because their very nature points them out as ends in themselves, that is, as something which must not be used merely as means, and so far therefore restricts freedom of action (and is an object of respect). These, therefore, are not merely subjective ends, whose existence has a worth *for us* as an effect of our action, but *objective ends*, that is, things whose existence is an end in itself: an end, moreover, for which no other can be substituted which they should subserve *merely* as means, for otherwise nothing whatever would possess *absolute worth*; but if all worth were conditioned and therefore contingent, then there would be no supreme practical principle of reason whatever.

If then there is a supreme practical principle or, in respect of the human will, a categorical imperative, it must be one which, being drawn from the conception of that which is necessarily an end for everyone because it is *an end in itself*, constitutes an *objective principle of will*, and can therefore serve as a universal practical law. The foundation of this principle is: *rational nature exists as an end in itself*. Man necessarily conceives his own existence as being so: so far then this is a *subjective principle of human actions*. But every other rational being regards its existence similarly, just on the same rational principle that holds for me: so that it is at the same time an *objective principle* from which, as a supreme practical law, all laws of the will must be capable of being deduced. Accordingly the practical imperative will be as follows: *So act as to treat humanity, whether in thine own person or in that of any other, in every case as an end withal, never as means only*. We will now inquire whether this can be practically carried out.

To abide by the previous examples:

Firstly, under the head of necessary duty to oneself: he who contemplates suicide should ask himself whether his action can be consistent with the idea of humanity as *an end in itself*. If he destroys himself in order to escape from painful circumstances, he uses a person merely as *a mean* to maintain a tolerable condition up to the end of life. But a man is not a thing, that is to say, something which can be used merely as means, but must in all his actions be always considered as an end in himself. I cannot, therefore, dispose in any way of a man in my own person so as to mutilate him, to damage or kill him. [. . .]

Secondly, as regards necessary duties, or those of strict obligation, towards others: he who is thinking of making a lying promise to others will see at once that he would be using another man *merely as a mean*, without the latter containing at the same time the end in himself. For he whom I propose by such a promise to use for my own purposes cannot possibly assent to my mode of acting towards him, and therefore cannot himself contain the end of this action. This violation of the principle of humanity in other men is more obvious if we take in examples of attacks on the freedom and property of others. For then it is clear that he who transgresses the rights of men intends to use the person of others merely as means, without considering that as rational beings they ought always to be esteemed also as ends, that is, as beings who must be capable of containing in themselves the end of the very same action.

Thirdly, as regards contingent (meritorious) duties to oneself: it is not enough that the action does not violate humanity in our own person as an end in itself, it must also harmonize with it. Now there are in humanity capacities of greater perfection which belong to the end that nature has in view in regard to humanity in ourselves as the subject: to neglect these might perhaps be consistent with the maintenance of humanity as an end in itself, but not with the advancement of this end.

Fourthly, as regards meritorious duties towards others: the natural end which all men have is their

own happiness. Now humanity might indeed subsist although no one should contribute anything to the happiness of others, provided he did not intentionally withdraw anything from it; but after all, this would only harmonize negatively, not positively, with humanity as *an end in itself* if everyone does not also endeavour, as far as in him lies, to forward the ends of others. For the ends of any subject which is an end in himself ought as far as possible to be *my* ends also, if that conception is to have its full effect with me.

This principle that humanity and generally every rational nature is *an end in itself* (which is the supreme limiting condition of every man's freedom of action) is not borrowed from experience: *firstly*, because it is universal, applying as it does to all rational beings whatever, and experience is not capable of determining anything about them; *secondly*, because it does not present humanity as an end to men (subjectively), that is as an object which men do of themselves actually adopt as an end, but as an objective end, which must as a law constitute the supreme limiting condition of all our subjective ends, let them be what we will; it must therefore spring from pure reason. In fact the objective principle of all practical legislation lies (according to the first principle) in the rule and its form of universality which makes it capable of being a law (say, e.g., a law of nature); but the subjective principle is in the *end*; now by the second principle, the subject of all ends is each rational being, inasmuch as it is an end in itself. Hence follows the third practical principle of the will, which is the ultimate condition of its harmony with the universal practical reason, viz., the idea of the will of every rational being as a universally legislative will.

On this principle all maxims are rejected which are inconsistent with the will being itself universal legislator. Thus the will is not subject simply to the law, but so subject that it must be regarded as *itself giving the law*, and on this ground only, subject to the law (of which it can regard itself as the author). [. . .]

I will therefore call this the principle of Autonomy of the will, in contrast with every other which I accordingly reckon as Heteronomy.

The conception of every rational being as one which must consider itself as giving in all the maxims of its will universal laws, so as to judge itself and its actions from this point of view—this conception leads to another which depends on it and is very fruitful, namely, that of a kingdom of ends.

By a kingdom I understand the union of different rational beings in a system by common laws. Now since it is by laws that ends are determined as regards their universal validity, hence, if we abstract from the personal differences of rational beings, and likewise from all the content of their private ends, we shall be able to conceive all ends combined in a

systematic whole (including both rational beings as ends in themselves, and also the special ends which each may propose to himself), that is to say, we can conceive a kingdom of ends, which on the preceding principles is possible.

For all rational beings come under the law that each of them must treat itself and all others never merely as a means, but in every case at the same time as ends in themselves. Hence results a systematic union of rational beings by common objective laws, i.e., a kingdom which may be called a kingdom of ends, since what these laws have in view is just the relation of these beings to one another as ends and means. It is certainly only an ideal. [. . .]

Notes

1. Tr: The first proposition was that to have moral worth an action must be done from duty.

2. I understand by a perfect duty one that admits no exception in favour of inclination.

Study Questions

1. For Kant, what makes an action a good action, and what is good in itself?
2. Although Kant specifies that there is “but one categorical imperative,” he characterizes it in three different ways in order to help us apply it in practice. Describe the three forms of the categorical imperative, and show how each applies by making reference to an example.

3. Kant’s theory is a “deontological” theory, a term derived from Greek *deon*, which means “duty.” Describe Kant’s classification of duties into perfect vs. imperfect duties and duties to oneself vs. others, providing an example for each of the four classes.
4. On Kant’s account, is a child a person? A fetus? A comatose patient? Someone senile? What moral implications result?

↖ MORAL FOUNDATIONS ↳

Aristotle

Aristotle (384–322 BCE) was born in the northern Greek city of Stagira; his father was physician to King Amyntas III of Macedonia. At age 17, Aristotle went to Athens to study at Plato's Academy. Aristotle left Athens after Plato's death and studied natural history in Mysea and Lesbos before becoming tutor to King Philip II of Macedonia's son Alexander—later Alexander the Great—in Pella. After Alexander became king, Aristotle returned to Athens and opened his own school at the Lyceum. *The Nicomachean Ethics* discusses what is the best kind of life for a human being to live.

From *The Nicomachean Ethics*

Book 1

account of it is not the same as that given by the philosophers. Ordinary people identify it with some

[...] or honour—some say one thing and some another,
2. If [...] among the ends at which our actions aim there be one which we wish for its own sake, while we wish the others only for the sake of this, and if we do not choose everything for the sake of something else (which would obviously result in a process ad infinitum, so that all desire would be futile and vain), it is clear that this one ultimate End must be the Good, and indeed the Supreme Good. Will not then a knowledge of this Supreme Good be also of great practical importance for the conduct of life? Will it not better enable us to attain what is fitting, like archers having a target to aim at? If this be so, we ought to make an attempt to determine [...] what exactly this Supreme Good is, and of which of the theoretical or practical sciences it is the object. [...]

4. [...] [Inasmuch as all studies and undertakings are directed to the attainment of some good, let us discuss what it is that we pronounce to be the aim of Politics, that is, what is the highest of all the goods that action can achieve. As far as the name goes, we may almost say that the great majority of mankind are agreed about this; for both the multitude and persons of refinement speak of it as Happiness,¹ and conceive "the good life" or "doing well" to be the same thing as "being happy." But what constitutes happiness is a matter of dispute; and the popular

task to review all the different opinions that are held. It will suffice to examine those that are most widely prevalent, or that seem to have some argument in their favour. [...]

5. [...] To judge from men's lives, the more or less reasoned conceptions of the Good or Happiness that seem to prevail among them are the following. On the one hand the generality of men and the most vulgar identify the Good with pleasure, and accordingly are content with the Life of Enjoyment—for there are three specially prominent Lives, the one just mentioned, the Life of Politics, and thirdly, the Life of Contemplation. The generality of man

by preferring what is only a life for cattle; but they get a hearing for their view as reasonable because many persons of high position share the feelings of Sardanapalus.²

Men of refinement, on the other hand, and men of action think that the Good is honour—for this may be said to be the end of the Life of Politics. But honour after all seems too superficial to be the Good for which we are seeking; since it appears to depend on those who confer it more than on him upon whom it is conferred, whereas we instinctively feel that the Good must be something proper to its possessor and not easy to be taken away from him. Moreover men's motive in pursuing honour seems to be to assure themselves of their own merit; at least they seek to be honoured by men of judgement and by people who know them, that is, they desire to be honoured on the ground of virtue. It is clear therefore that in the opinion [...] of men of action, virtue is a greater good than honour; and one might perhaps accordingly suppose that virtue rather than honour is the end of the Political Life. But even virtue proves on examination to be too incomplete to be the End; since it appears possible to possess it while you are asleep, or without putting it into practice throughout the whole of your life; and also for the virtuous man to suffer the greatest misery and misfortune—though no one would pronounce a man living a life of misery to be happy, unless for the sake of maintaining a paradox. [...]

The third type of life is the Life of Contemplation, which we shall consider in the sequel.

6. [...] Now there do appear to be several ends at which our actions aim; but as we choose some of them—for instance wealth, or flutes, and instruments generally—as a means to something else, it is clear that not all of them are final ends; whereas the Supreme Good seems to be something final. Consequently if there be some one thing which alone is a final end, this thing—or if there be several final ends, the one among them which is the most final—will be the Good which we are seeking. In speaking of degrees of finality, we mean that a thing pursued as an end in itself is more final than one pursued as a means to something else, and that a

thing never chosen as a means to anything else is more final than things chosen both as ends in themselves and as means to that thing, and accordingly a thing chosen always as an end and never as a means we call absolutely final. Now happiness above all else appears to be absolutely final in this sense, since we always choose it for its own sake and never as a means to something else; whereas honour, pleasure, intelligence, and excellence in its various forms, we choose indeed for their own sakes (since we should be glad to have each of them although no extraneous advantage resulted from it), but we also choose them for the sake of happiness, in the belief that they will be a means to our securing it. But no one chooses happiness for the sake of honour, pleasure, etc., nor as a means to anything whatever other than itself. [...]

To say however that the Supreme Good is happiness will probably appear a truism; we still require a more explicit account of what constitutes happiness. Perhaps then we may arrive at this by ascertaining what is man's function. For the goodness or efficiency of a flute-player or sculptor or craftsman of any sort, and in general of anybody who has some function or business to perform, is thought to reside in that function; and similarly it may be held that the good of man resides in the function of man, if he has a function.

Are we then to suppose that, while the carpenter and the shoemaker have definite functions or businesses belonging to them, man as such has none, and is not designed by nature to fulfil any function? Must we not rather assume that, just as the eye, the hand, the foot and each of the various members of the body manifestly has a certain function of its own, so a human being also has a certain function over and above all the functions of his particular members? What then precisely can this function be? The mere act of living appears to be shared even by plants, whereas we are looking for the function peculiar to man; we must therefore set aside the vital activity of nutrition and growth. Next in the scale will come some form of sentient life; but this too appears to be shared by horses, oxen, and animals generally. There remains therefore what

may be called the practical³ life of the rational part of man. [...] Rational life again has two meanings; let us assume that we are here concerned with the active exercise⁴ of the rational faculty, since this seems to be the more proper sense of the term. If then the function of man is the active exercise of the souls faculties in conformity with rational principle, or at all events not in dissociation from rational principle, and if we acknowledge the function of an individual and of a good individual of the same class (for instance, a harpist and a good harpist, and so generally with all classes) to be generically the same, the qualification of the latter's superiority in excellence being added to the function in his case (I mean that if the function of a harpist is to play the harp, that of a good harpist is to play the harp well): if this is so, and if we declare that the function of man is a certain form of life, and define that form of life as the exercise of the souls faculties and activities in association with rational principle, and say that the function of a good man is to perform these activities well and rightly, and if a function is well performed when it is performed in accordance with its own proper excellence—from these premises it follows that the Good of man is the active exercise of his souls faculties in conformity with excellence or virtue, or if there be several human excellences or virtues, in conformity with the best and most perfect among them. Moreover this activity must occupy a complete lifetime; for one swallow does not make spring, nor does one fine day; and similarly one day or a brief period of happiness does not make a man supremely blessed and happy.

Book 2

1. [...] [M]oral or ethical virtue is the product of habit (*ethos*), and has indeed derived its name (*ethos* for "character"), with a slight variation of form, from that word. And therefore it is clear that none of the moral virtues is engendered in us by nature, for no natural property can be altered by habit. For instance, it is the nature of a stone to move downwards, and it cannot be trained to move upwards, even though you should try to train it to do so by

throwing it up into the air ten thousand times; nor can fire be trained to move downwards, nor can anything else that naturally behaves in one way be trained into a habit of behaving in another way. The virtues therefore are engendered in us neither by nature nor yet in violation of nature; nature gives us the capacity to receive them, and this capacity is brought to maturity by habit.

Moreover, the faculties given us by nature are bestowed on us first in a potential form; we exhibit their actual exercise afterwards. This is clearly so with our senses: we did not acquire the faculty of sight or hearing by repeatedly seeing or repeatedly listening, but the other way about—because we had the senses we began to use them, we did not get them by using them. The virtues on the other hand we acquire by first having actually practised them, just as we do the arts. We learn an art or craft by doing the things that we shall have to do when we have learnt it: for instance, men become builders by building houses, harpists by playing on the harp. Similarly we become just by doing just acts, temperate by doing temperate acts, brave by doing brave acts. [...] Again, the actions from or through which any virtue is produced are the same as those through which it also is destroyed—just as is the case with skill in the arts, for both the good harpists and the bad ones are produced by harping, and similarly with builders and all the other craftsmen: as you will become a good builder from building well, so you will become a bad one from building badly. Were this not so, there would be no need for teachers of the arts, but everybody would be born a good or bad craftsman as the case might be. The same then is true of the virtues. It is by taking part in transactions with our fellow-men that some of us become just and others unjust; by acting in dangerous situations and forming a habit of fear or of confidence we become courageous or cowardly. And the same holds good of our dispositions with regard to the appetites, and anger; some men become temperate and gentle, others profligate and lascivious, by actually comporting themselves in one way or the other in relation to those passions. In a word, our moral dispositions are formed as a result of the

corresponding activities. Hence it is incumbent on us to control the character of our activities, since on the quality of these depends the quality of our dispositions. It is therefore not of small moment whether we are trained from childhood in one set of habits or another; on the contrary it is of very great, or rather of supreme, importance. [...]

4. A difficulty may however be raised as to what we mean by saying that in order to become just men must do just actions, and in order to become temperate they must do temperate actions. For if they do just and temperate actions, they are just and temperate already, just as, if they spell correctly or play in tune, they are scholars or musicians.

But perhaps this is not the case even with the arts. It is possible to spell a word correctly by chance, or because someone else prompts you; hence you will be a scholar only if you spell correctly in the scholar's way, that is, in virtue of the scholarly knowledge which you yourself possess.

Moreover the case of the arts is not really analogous to that of the virtues. Works of art have their merit in themselves, so that it is enough if they are produced having a certain quality of their own; but acts done in conformity with the virtues are not done justly or temperately if they themselves are of a certain sort, but only if the agent also is in a certain state of mind when he does them: first he must act with knowledge; secondly he must deliberately choose the act, and choose it for its own sake; and thirdly the act must spring from a fixed and permanent disposition of character.

For the possession of an art, none of these conditions is included, except the mere qualification of knowledge; but for the possession of the virtues, knowledge is of little or no avail, whereas the other conditions, so far from being of little moment, are all-important, inasmuch as virtue results from the repeated performance of just and temperate actions. Thus although actions are entitled just and temperate when they are such acts as just and temperate men would do, the agent is just and temperate not when he does these acts merely, but when he does them in the way in which just and temperate men do them. It is correct therefore to say that a man

becomes just by doing just actions and temperate by doing temperate actions, and no one can have the remotest chance of becoming good without doing them. But the mass of mankind, instead of doing virtuous acts, have recourse to discussing virtue, and fancy that they are pursuing philosophy and that this will make them good men. In so doing they act like invalids who listen carefully to what the doctor says, but entirely neglect to carry out his prescriptions. That sort of philosophy will no more lead to a healthy state of soul than will the mode of treatment produce health of body.

5. We have next to consider the formal definition of virtue.

A state of the soul is either (1) an emotion, (2) a capacity, or (3) a disposition; virtue therefore must be one of these three things. By the emotions, I mean desire, anger, fear, confidence, envy, joy, friendship, hatred, longing, jealousy, pity; and generally those states of consciousness which are accompanied by pleasure or pain. The capacities are the faculties in virtue of which we can be said to be liable to the emotions, for example, capable of feeling anger or fear or pity. The dispositions are the formed states of character in virtue of which we are well or ill disposed in respect of the emotions; for instance, we have a bad disposition in regard to anger if we are disposed to get angry too violently or not violently enough, a good disposition if we habitually feel a moderate amount of anger; and similarly in respect of the other emotions.

Now the virtues and vices are not emotions because we are not pronounced good or bad according to our emotions, but we are according to our virtues and vices; nor are we either praised or blamed for our emotions—a man is not praised for being frightened or angry, nor is he blamed for being angry merely, but for being angry in a certain way—but we are praised or blamed for our virtues and vices. Again, we are not angry or afraid from choice, but the virtues are certain modes of choice, or at all events involve choice. Moreover, we are said to be “moved” by the emotions, whereas in respect of the virtues and vices we are not said to be “moved” but to be “disposed” in a certain way.

And the same considerations also prove that the virtues and vices are not capacities; since we are not pronounced good or bad, praised or blamed, merely by reason of our capacity for emotion. Again, we possess certain capacities by nature, but we are not born good or bad by nature: of this however we spoke before.

If then the virtues are neither emotions nor capacities, it remains that they are dispositions.

Thus we have stated what virtue is generically.

6. But it is not enough merely to define virtue generically as a disposition; we must also say what species of disposition it is. It must then be premised that all excellence has a twofold effect on the thing to which it belongs: it not only renders the thing itself good, but it also causes it to perform its function well. For example, the effect of excellence in the eye is that the eye is good and functions well; since having good eyes means having good sight. Similarly excellence in a horse makes it a good horse, and also good at galloping, at carrying its rider, and at facing the enemy. If therefore this is true of all things, excellence or virtue in a man will be the disposition which renders him a good man and also which will cause him to perform his function well. We have already indicated what this means; but it will throw more light on the subject if we consider what constitutes the specific nature of virtue.

Now of everything that is continuous and divisible, it is possible to take the larger part, or the smaller part, or an equal part, and these parts may be larger, smaller, and equal either with respect to the thing itself or relatively to us; the equal part being a mean between excess and deficiency. By the mean of the thing I denote a point equally distant from either extreme, which is one and the same for everybody; by the mean relative to us, that amount which is neither too much nor too little, and this is not one and the same for everybody. For example, let 10 be many and 2 few; then one takes the mean with respect to the thing if one takes 6; since $6 - 2 = 10 - 6$, and this is the mean according to arithmetic proportion. But we cannot arrive by this method at the mean relative to us. Suppose that 10 lb. of food is a large ration for anybody and 2 lb. a small

one: it does not follow that a trainer will prescribe 6 lb., for perhaps even this will be a large ration, or a small one, for the particular athlete who is to receive it; it is a small ration for a Milo [a famous wrestler], but a large one for a man just beginning to go in for athletics. And similarly with the amount of running or wrestling exercise to be taken. In the same way then an expert in any art avoids excess and deficiency, and seeks and adopts the mean—the mean, that is, not of the thing but relative to us.

If therefore the way in which every art or science performs its work well is by looking to the mean and applying that as a standard to its productions (hence the common remark about a perfect work of art, that you could not take from it nor add to it)—meaning that excess and deficiency destroy perfection, while adherence to the mean preserves it)—if then, as we say, good craftsmen look to the mean as they work, and if virtue, like nature, is more accurate and better than any form of art, it will follow that virtue has the quality of hitting the mean. [...] For example, one can be frightened or bold, feel desire or anger or pity, and experience pleasure and pain in general, either too much or too little, and in both cases wrongly; whereas to feel these feelings at the right time, on the right occasion, towards the right people, for the right purpose and in the right manner, is to feel the best amount of them, which is the mean amount—and the best amount is of course the mark of virtue. And similarly there can be excess, deficiency, and the due mean in actions. Now feelings and actions are the objects with which virtue is concerned; and in feelings and actions excess and deficiency are errors, while the mean amount is praiseworthy, and constitutes success; and to be praiseworthy, and to be successful are both marks of virtue. [...] Virtue then is a settled disposition of the mind determining the choice of actions and emotions, consisting essentially in the observance of the mean relative to us, this being determined by principle, that is, as the prudent man would determine it.

And it is a mean state between two vices, one of excess and one of defect. Furthermore, it is a mean state in that whereas the vices either fall short of virtue

ascertains and adopts the mean. Hence while in respect of its substance and the definition that states what it really is in essence virtue is the observance of the mean, in point of excellence and rightness it is an extreme. [...]

9. Enough has now been said to show that moral virtue is a mean, and in what sense this is so, namely that it is a mean between two vices, one of excess and the other of defect; and that it is such a mean because it aims at hitting the middle point in feelings and in actions. This is why it is a hard task to be good, for it is hard to find the middle point in anything: for instance, not everybody can find the centre of a circle, but only someone who knows geometry. So also anybody can become angry—that is easy; and so it is to give and spend money; but to be angry with or give money to the right person, and to the right amount, and at the right time, and for the right purpose, and in the right way—this is not within everybody's power and is not easy; so that to do these things properly is rare, praiseworthy, and noble.

Hence the first rule in aiming at the mean is to avoid that extreme which is the more opposed to the mean, as Calypso advises (*Od.* xii, 219)⁵—

Steer the ship clear of yonder spray and surge.

For of the two extremes one is a more serious error than the other. Hence, inasmuch as to hit the mean extremely well is difficult, the second best way to sail⁶ as the saying goes, is to take the least of the evils; and the best way to do this will be the way we enjoin.

The second rule is to notice what are the errors to which we are ourselves most prone (as different men are inclined by nature to different faults)—and we shall discover what these are by observing the pleasure or pain that we experience—; then we must drag ourselves away in the opposite direction, for by steering wide of our besetting error we shall make a middle course. This is the method adopted by carpenters to straighten warped timber.

Thirdly, we must in everything be most of all on our guard against what is pleasant and against pleasure; for when pleasure is on her trial we are

not impartial judges. The right course is therefore to feel towards pleasure as the elders of the people felt towards Helen (*Iliad* iii, 155–160), and to apply their words to her on every occasion, for if we roundly bid her be gone, we shall be less likely to err.

These then, to sum up the matter, are the precautions that will best enable us to hit the mean. [...]

Book 10

[...]

7. But if happiness consists in activity in accordance with virtue, it is reasonable that it should be activity in accordance with the highest virtue; and this will be the virtue of the best part of us. Whether then this be the intellect, or whatever else it be that is thought to rule and lead us by nature, and to have cognizance of what is noble and divine, either as being itself also actually divine, or as being relatively the divinest part of us, it is the activity of this part of us in accordance with the virtue proper to it that will constitute perfect happiness; and [...] this activity is the activity of contemplation.

And that happiness consists in contemplation may be accepted as agreeing both with the results already reached and with the truth. For contemplation is at once the highest form of activity (since the intellect is the highest thing in us, and the objects with which the intellect deals are the highest things that can be known). [...] Also the activity of contemplation may be held to be the only activity that is loved for its own sake: it produces no result beyond the actual act of contemplation, whereas from practical pursuits we look to secure some advantage, greater or smaller, beyond the action itself. [...]

Such a life as this however will be higher than the human level: not in virtue of his humanity will a man achieve it, but in virtue of something within him that is divine; and by as much as this something is superior to his composite nature, by so much is its activity superior to the exercise of the other forms of virtue. If then the intellect is something divine in comparison with man, so is the life of the intellect divine in comparison with human life. Nor ought

we to obey those who enjoin that a man should have man's thoughts and a mortal the thoughts of mortality; but we ought so far as possible to achieve immortality, and do all that man may to live in accordance with the highest thing in him; for though this be small in bulk, in power and value it far surpasses all the rest.

It may even be held that this is the true self of each, inasmuch as it is the dominant and better part; and therefore it would be a strange thing if a man should choose to live not his own life but the life of some other than himself.

Moreover what was said before will apply here also: that which is best and most pleasant for each creature is that which is proper to the nature of each; accordingly the life of the intellect is the best and the pleasantest life for man, inasmuch as the intellect more than anything else is man; therefore this life will be the happiest.

8. The life of moral virtue, on the other hand, is happy only in a secondary degree. For the moral activities are purely human. Justice, I mean, Courage and the other virtues we display in our intercourse with our fellows, when we observe what is due to each in contracts and services and in our various actions, and in our emotions also; and all of these things seem to be purely human affairs. And some moral actions are thought to be the outcome of the physical constitution, and moral virtue is thought to have a close affinity in many respects with the passions. Moreover, Prudence is intimately connected with Moral Virtue, and this with Prudence, inasmuch as the first principles which Prudence employs are determined by the

1. Tr.: This translation of *eudaimonia* can hardly be avoided, but it would perhaps be more accurately rendered by "Well-being" or "Prosperity"; and it will be found that the writer does not interpret it as a state of feeling but as a kind of activity.
2. Tr.: Sardanapalus was a mythical Assyrian king; two versions of his epitaph are recorded by

3. Tr.: "Practice" for Aristotle denotes purposeful conduct, of which only rational beings are capable.

Notes

Moral Virtues, and the right standard for the Moral Virtues is determined by Prudence. But these being also connected with the passions are related to our composite nature; now the virtues of our composite nature are purely human, so therefore also is the life that manifests these virtues, and the happiness that belongs to it. Whereas the happiness that belongs to the intellect is separate [. . .].⁷ And such happiness the politician does as a matter of fact take more trouble about bodily requirements and so forth than the philosopher; for in this respect there may be little difference between them. But for the purpose of their special activities their requirements will differ widely. The liberal man will need wealth in order to do liberal actions, and so indeed will the just man in order to discharge his obligations (since mere intentions are invisible, and even the unjust pretend to wish to act justly); and the brave man will need strength if he is to perform any action displaying his virtue; and the temperate man opportunity for indulgence: otherwise how can he, or the possessor of any other virtue, show that he is virtuous? [. . .] But the student, so far as the pursuit of his activity is concerned, needs no external apparatus on the contrary, worldly goods may almost be said to be a hindrance to contemplation; though it is true that, being a man and living in the society of others, he chooses to engage in virtuous action, and so will need external goods to carry on his life as a human being [. . .].

Athenaeus, one containing the words "Eat, drink, play since all else is not worth that snap of the fingers"; the other ends "I have what I ate; and the delightful deeds of wantonness and love which I did and suffered; whereas all my wealth is vanished."

4. Tr.: In contrast with the mere state of possessing the faculty.
5. Tr.: Really the words are said by Odysseus, conveying to his steersman Circe's advice, to avoid the whirlpool of Charybdis which will engulf them all, and steer nearer to the monster Scylla who will devour only some of them.
6. Tr.: A proverb, meaning to take to the oars when the wind fails.
7. Tr.: In *De anima* III.5, Aristotle distinguishes the active from the passive intellect, and pronounces the former to be "separate or separable (from matter, or the body), unmixed and impassible."

Works Cited

Homer: *The Odyssey*. Trans. Samuel Butler. London: Longmans, Green, and Co., 1900.

Homer: *The Iliad*. Trans. A.T. Murray. Cambridge, MA: Harvard University Press and London: William Heinemann, 1924.

Study Questions

1. Provide an example of an action you do for its own sake and an example of an action you do for the sake of other ends.
2. According to Aristotle, what are three prevailing conceptions of the "good life," what "good" is specific to each, and which of these is truly a happy life?
3. What three kinds of "states of the soul" does Aristotle identify, and which of these is virtue?
4. Aristotle tells us that "the mark of virtue" is to feel or act "at the right time, on the right occasion, toward the right people, for the right purpose and in the right manner." Can you give an example that illustrates what he means?

Immanuel Kant

Immanuel Kant (1724–1804) rejected the moral sentimentalism of his Scottish empiricist contemporaries (e.g., Hume): for Kant, actions motivated by feelings have no moral worth. Kant held that the highest principle of practical philosophy is the moral law, which depends only on the nature of reason. By means of our (human) capacity for reason, we discover what the moral law requires of us. Also called *Groundwork of the Metaphysic of Morals*, this is Kant's foundational work on morality.

Plato

Plato (429–347 BCE) was a citizen of the ancient Greek city-state of Athens, born into a wealthy family with anti-democratic inclinations. When Athens was defeated by Sparta in the Peloponnesian War (431–404 BCE), the oligarchy installed included Plato's mother's cousin Critias and brother Charmides. Pro-democratic forces regained control in 403, and it was at the hands of Athens' newly reconstituted democracy that Socrates was executed in 399. *Crito* covers Socrates' final days, as he refused his friends' help to escape and calmly accepted his death.

Crito (or On Duty)

Characters: Socrates, Crito

SOCRATES. Why have you come at this time, Crito? Or isn't it still early?

CRITO. Yes, very early.

S. About what time?

C. Just before dawn.

S. I am surprised that the watchman of the prison was willing to let you in.

C. He is used to me by this time, Socrates, because I come here so often, and besides I have done something for him.

S. Have you just come, or some time ago?

C. Some little time ago.

S. Then why did you not wake me at once, instead of sitting by me in silence?

C. No, no, by Zeus, Socrates, I only wish I myself were not so sleepless and sorrowful. But I have been wondering at you for some time, seeing how sweetly you sleep; and I purposely refrained from waking you, that you might pass the time as pleasantly as possible. I have often thought throughout your life hitherto that you were of a happy disposition, and I think so more than ever in this present misfortune, since you bear it so easily and calmly.

S. Well, Crito, it would be absurd if at my age I were disturbed because I must die now.

C. Other men as old, Socrates, become involved in similar misfortunes, but their age does not in the least prevent them from being disturbed by their fate.

S. That is true. But why have you come so early?

C. To bring news, Socrates, sad news, though apparently not sad to you, but sad and grievous to me and all your friends, and to few of them, I think, so grievous as to me.

S. What is this news? Has the **ship** come from **Delos**, at the arrival of which I am to die?

C. It has not exactly come, but I think it will come today from the reports of some men who have come from Sunium and left it there. Now it is clear from what they say that it will come today, and so tomorrow, Socrates, your life must end.

S. Well, Crito, good luck be with us! If this is the will of the gods, so be it. However, I do not think it will come today.

C. What is your reason for not thinking so?

S. I will tell you. I must die on the day after the ship comes in, must I not?

C. So those say who have charge of these matters.

S. Well, I think it will not come in today, but tomorrow. And my reason for this is a dream which I had a little while ago in the course of this night. And perhaps you let me sleep just at the right time.

C. What was the dream?

S. I dreamed that a beautiful, fair woman, clothed in white raiment, came to me and called me and said, "Socrates, on the third day thou wouldest come to fertile Phthia" (Homer, *Iliad* ix, 363).

C. A strange dream, Socrates.

S. No, a clear one, at any rate, I think, Crito.

C. Too clear, apparently. But, my dear Socrates, even now listen to me and save yourself. Since, if you die, it will be no mere single misfortune to me, but I shall lose a friend such as I can never find again, and besides, many persons who do not know you and me well will think I could have saved you if I had been willing to spend money, but that I would not take the trouble. And yet what reputation could be more disgraceful than that of considering one's money of more importance than one's friends? For most people will not believe that we were eager to help you to go away from here, but you refused.

S. But, my dear Crito, why do we care so much for what most people think? For the most reasonable men, whose opinion is more worth considering, will think that things were done as they really will be done.

C. But you see it is necessary, Socrates, to care for the opinion of the public, for this very trouble we are in now shows that the public is able to accomplish not by any means the least, but almost the greatest of evils, if one has a bad reputation with it.

S. I only wish, Crito, the people could accomplish the greatest evils, that they might be able to accomplish also the greatest good things. Then all would be well. But now they can do neither of the two; for they are not able to make a man wise or foolish, but they do whatever occurs to them.

C. That may well be. But, Socrates, tell me this: you are not considering me and your other friends, are you, fearing that, if you escape, the informers will make trouble for us by saying that we stole you away, and we shall be forced to lose either all our property or a good deal of money, or be punished in some

other way besides? For if you are afraid of anything of that kind, let it go; since it is right for us to run this risk, and even greater risk than this, if necessary, provided we save you. Now please do as I ask.

S. I am considering this, Crito, and many other things.

C. Well, do not fear this! for it is not even a large sum of money which we should pay to some men who are willing to save you and get you away from here. Besides, don't you see how cheap these informers are, and that not much money would be needed to silence them? And you have my money at your command, which is enough, I fancy; and moreover, if because you care for me you think you ought not to spend my money, there are foreigners here willing to spend theirs; and one of them, Simmias of Thebes, has brought for this especial purpose sufficient funds; and Cebes also and very many others are ready. So, as I say, do not give up saving yourself through fear of this. And do not be troubled by what you said in the court, that if you went away you would not know what to do with yourself. For in many other places, wherever you go, they will welcome you; and if you wish to go to Thessaly, I have friends there who will make much of you and will protect you, so that no one in Thessaly shall annoy you.

And besides, Socrates, it seems to me the thing you are undertaking to do is not even right—betraying yourself when you might save yourself. And you are eager to bring upon yourself just what your enemies would wish and just what those were eager for who wished to destroy you. And moreover, I think you are abandoning your children, too, for when you might bring them up and educate them, you are going to desert them and go away, and, so far as you are concerned, their fortunes in life will be whatever they happen to meet with, and they will probably meet with such treatment as generally comes to orphans in their destitution. No. Either one ought not to beget children, or one ought to stay by them and bring them up and educate them. But you seem to me to be choosing the laziest way; and you ought to choose as a good and brave man

would choose, you who have been saying all your life that you cared for virtue. So I am shamed both for you and for us, your friends, and I am afraid people will think that this whole affair of yours has been conducted with a sort of cowardice on our part—both the fact that the case came before the court, when it might have been avoided, and the way in which the trial itself was carried on, and finally they will think, as the crowning absurdity of the whole affair, that this opportunity has escaped us through some base cowardice on our part, since we did not save you, and you did not save yourself, though it was quite possible if we had been of any use whatever. Take care, Socrates, that these things be not disgraceful, as well as evil, both to you and to us. Just consider, or rather it is time not to consider any longer, but to have finished considering. And there is just one possible plan; for all this must be done in the coming night. And if we delay it can no longer be done. But I beg you, Socrates, do as I say and don't refuse.

S. My dear Crito, your eagerness is worth a great deal, if it should prove to be rightly directed; but otherwise, the greater it is, the more hard to bear. So we must examine the question whether we ought to do this or not; for I am not only now but always a man who follows nothing but the reasoning which on consideration seems to me best. And I cannot, now that this has happened to us, discard the arguments I used to advance, but they seem to me much the same as ever, and I revere and honour the same ones as before. And unless we can bring forward better ones in our present situation, be assured that I shall not give way to you, not even if the power of the multitude frighten us with even more terrors than at present, as children are frightened with goblins, threatening us with imprisonments and deaths and confiscations of property. Now how could we examine the matter most reasonably? By taking up first what you say about opinions and asking whether we were right when we always used to say that we ought to pay attention to some opinions and not to others? Or were we right before I was condemned to death, whereas it has now been made clear that we were

talking merely for the sake of argument and it was really mere play and nonsense? And I wish to investigate, Crito, in common with you, and see whether our former argument seems different to me under our present conditions, or the same, and whether we shall give it up or be guided by it. But it used to be said, I think, by those who thought they were speaking sensibly, just as I was saying now, that of the opinions held by men some ought to be highly esteemed and others not. In God's name, Crito, do you not think this is correct? For you, humanly speaking, are not involved in the necessity of dying tomorrow, and therefore present conditions would not lead your judgment astray. Now say, do you not think we were correct in saying that we ought not to esteem all the opinions of men, but some and not others, and not those of all men, but only of some? What do you think? Is not this true?

C. It is.

S. Then we ought to esteem the good opinions and not the bad ones?

C. Yes.

S. And the good ones are those of the wise and the bad ones those of the foolish?

C. Of course.

S. Come then, what used we to say about this? If a man is an athlete and makes that his business, does he pay attention to every man's praise and blame and opinion or to those of one man only who is a physician or a trainer?

C. To those of one man only.

S. Then he ought to fear the blame and welcome the praise of that one man and not of the multitude.

C. Obviously.

S. And he must act and exercise and eat and drink as the one man who is his director and who knows the business thinks best rather than as all the others think.

C. That is true.

S. Well then; if he disobeys the one man and disregards his opinion and his praise, but regards words of the many who have no special knowledge, will he not come to harm?

C. Of course he will.

S. And what is this harm? In what direction and upon what part of the one who disobeys does it act?

C. Evidently upon his body; for that is what it ruins.

S. Right. Then in other matters, not to enumerate them all, in questions of right and wrong and disgraceful and noble and good and bad, which we are now considering, ought we to follow and fear the opinion of the many or that of the one, if there is anyone who knows about them, whom we ought to revere and fear more than all the others? And if we do not follow him, we shall injure and cripple that which we used to say is benefited by the right and is ruined by the wrong. Or is there nothing in this?

C. I think it is true, Socrates.

S. Well then, if through yielding to the opinion of the ignorant we ruin that which is benefited by health and injured by disease, is life worth living for us when that is ruined? And that is the body, is it not?

C. Yes.

S. Then is life worth living when the body is worthless and ruined?

C. Certainly not.

S. But is it worth living when that is ruined which is injured by the wrong and improved by the right? Or do we think that part of us, whatever it is, which is concerned with right and wrong, is less important than the body?

C. By no means.

S. But more important?

C. Much more.

S. Then, most excellent friend, we must not consider at all what the many will say of us, but what he who knows about right and wrong, the one man,

and truth herself will say. And so you introduced the discussion wrongly in the first place, when you began by saying we ought to consider the opinion of the multitude about the right and the noble and the good and their opposites. But it might, of course, be said that the multitude can put us to death.

C. That is clear, too. It would be said, Socrates.

S. That is true. But, my friend, the argument we have just finished seems to me still much the same as before; and now see whether we still hold to this, or not, that it is not living, but living well which we ought to consider most important.

C. We do hold to it.

S. And that living well and living rightly are the same thing, do we hold to that, or not?

C. We do.

S. Then we agree that the question is whether it is right for me to try to escape from here without the permission of the Athenians, or not right. And if it appears to be right, let us try it, and if not, let us give it up. But the considerations you suggest, about spending money, and reputation, and bringing up my children, these are really, Crito, the reflections of those who lightly put men to death, and would bring them to life again, if they could, without any sense, I mean the multitude. But we, since our argument so constrains us, must consider only the question we just broached, whether we shall be doing right in giving money and thanks to these men who will help me to escape, and in escaping or aiding the escape ourselves, or shall in truth be doing wrong, if we do all these things. And if it appears that it is wrong for us to do them, it may be that we ought not to consider either whether we must die if we stay here and keep quiet or whether we must endure anything else whatsoever, but only the question of doing wrong.

C. I think what you say is right, Socrates; but think what we should do.

S. Let us, my good friend, investigate in common, and if you can contradict anything I say, do so, and I will yield to your arguments; but if you cannot, my

dear friend, stop at once saying the same thing to me over and over, that I ought to go away from here without the consent of the Athenians; for I am anxious to act in this matter with your approval, and not contrary to your wishes. Now see if the beginning of the investigation satisfies you, and try to reply to my questions to the best of your belief.

C. I will try.

S. Ought we in no way to do wrong intentionally, or should we do wrong in some ways but not in others? Or, as we often agreed in former times, is it never right or honourable to do wrong? Or have all those former conclusions of ours been overturned in these few days, and have we old men, seriously conversing with each other, failed all along to see that we were no better than children? Or is not what we used to say most certainly true, whether the world agree or not? And whether we must endure still more grievous sufferings than these, or lighter ones, is not wrongdoing inevitably an evil and a disgrace to the wrongdoer? Do we believe this or not?

C. We do.

S. Then we ought not to do wrong at all.

C. Why, no. . .

S. And we ought not even to requite wrong with wrong, as the world thinks, since we must not do wrong at all.

C. Apparently not.

S. Well, Crito, ought one to do evil or not?

C. Certainly not, Socrates.

S. Well, then, is it right to requite evil with evil, as the world says it is, or not right?

C. Not right, certainly.

S. For doing evil to people is the same thing as wronging them.

C. That is true.

S. Then we ought neither to requite wrong with wrong nor to do evil to anyone, no matter what he

may have done to us. And be careful, Crito, that you do not, in agreeing to this, agree to something you do not believe; for I know that there are few who believe or ever will believe this. Now those who believe this, and those who do not, have no common ground of discussion, but they must necessarily, in view of their opinions, despise one another. Do you therefore consider very carefully whether you agree and share in this opinion, and let us take as the starting point of our discussion the assumption that it is never right to do wrong or to requite wrong with wrong, or when we suffer evil to defend ourselves by doing evil in return. Or do you disagree and refuse your assent to this starting point? For I have long held this belief and I hold it yet, but if you have reached any other conclusion, speak and explain it to me. If you still hold to our former opinion, hear the next point.

C. I do hold to it and I agree with you; so go on.

S. Now the next thing I say, or rather ask, is this: "ought a man to do what he has agreed to do, provided it is right, or may he violate his agreements?"

C. He ought to do it.

S. Then consider whether, if we go away from here without the consent of the state, we are doing harm to the very ones to whom we least ought to do harm, or not, and whether we are abiding by what we agreed was right, or not.

C. I cannot answer your question, Socrates, for I do not understand.

S. Consider it in this way. If, as I was on the point of running away (or whatever it should be called), the laws and the commonwealth should come to me and ask, "Tell me, Socrates, what have you in mind to do? Are you not intending by this thing you are trying to do, to destroy us, the laws, and the entire state, so far as in you lies? Or do you think that state can exist and not be overturned, in which the decisions reached by the courts have no force but are made invalid and annulled by private persons?" What shall we say, Crito, in reply to this question and others of the same kind? For one might say many things,

especially if one were an orator, about the destruction of that law which provides that the decisions reached by the courts shall be valid. Or shall we say to them, "The state wronged me and did not judge the case rightly"? Shall we say that, or what?

C. That is what we shall say, by Zeus, Socrates.

S. What then if the laws should say, "Socrates, is this the agreement you made with us, or did you agree to abide by the verdicts pronounced by the state?" Now if I were surprised by what they said, perhaps they would continue, "Don't be surprised at what we say, Socrates, but answer, since you are in the habit of employing the method of question and answer. Come, what fault do you find with us and the state, that you are trying to destroy us? In the first place, did we not bring you forth? Is it not through us that your father married your mother and begat you? Now tell us, have you any fault to find with those of us who are the laws of marriage?"

"I find no fault," I should say. "Or with those that have to do with the nurture of the child after he is born and with his education which you, like others, received? Did those of us who are assigned to these matters not give good directions when we told your father to educate you in music and gymnastics?" "You did," I should say. "Well then, when you were born and nurtured and educated, could you say to begin with that you were not our offspring and our slave, you yourself and your ancestors? And if this is so, do you think right as between you and us rests on a basis of equality, so that whatever we undertake to do to you it is right for you to retaliate? There was no such equality of right between you and your father or your master, if you had one, so whatever treatment you received you might return it, answering them if you were reviled, or you think that it will be proper for you to act toward your country and the laws, so that if we undertake to destroy you, thinking it is right, you undertake in return to destroy us laws and your country, so far as you are able, and will say that in doing this you are doing right, you who really care for virtue? Or is your wisdom such that you do not

see that your country is more precious and more to be revered and is holier and in higher esteem among the gods and among men of understanding than your mother and your father and all your ancestors, and that you ought to show to her more reverence and obedience and humility when she is angry than to your father, and ought either to convince her by persuasion or to do whatever she commands, and to suffer, if she commands you to suffer, in silence, and if she orders you to be scourged or imprisoned or if she leads you to war to be wounded or slain, her will is to be done, and this is right, and you must not give way or draw back or leave your post, but in war and in court and everywhere, you must do whatever the state, your country, commands, or must show her by persuasion what is really right, but that it is impious to use violence against either your father or your mother, and much more impious to use it against your country?" What shall we reply to this, Crito, that the laws speak the truth, or not?

C. I think they do.

S. "Observe then, Socrates," perhaps the laws would say, "that if what we say is true, what you are now undertaking to do to us is not right. For we brought you into the world, nurtured you, and gave a share of all the good things we could to you and all the citizens. Yet we proclaim, by having offered the opportunity to any of the Athenians who wishes to avail himself of it, that anyone who is not pleased with us when he has become a man and has seen the administration of the city and us, the laws, may take his goods and go away wherever he likes. And none of us stands in the way or forbids any of you to take his goods and go away wherever he pleases, if we and the state do not please him, whether it be to an Athenian colony or to a foreign country where he will live as an alien. But we say that whoever of you stays here, seeing how we administer justice and how we govern the state in other respects, has thereby entered into an agreement with us to do what we command; and we say that he who does not obey does threefold wrong, because he disobeys us who are his parents, because he disobeys us who nurtured him,

and because after agreeing to obey us he neither obeys us nor convinces us that we are wrong, though we give him the opportunity and do not roughly order him to do what we command, but when we allow him a choice of two things, either to convince us of error or to do our bidding, he does neither of these things."

"We say that you, Socrates, will be exposed to these reproaches, if you do what you have in mind, and you not least of the Athenians but more than most others." If then I should say, "How so?" perhaps they might retort with justice that I had made this agreement with them more emphatically than most other Athenians. For they would say, "Socrates, we have strong evidence that we and the city pleased you; for you would never have stayed in it more than all other Athenians if you had not been better pleased with it than they; you never went out from the city to a festival, or anywhere else, except on military service, and you never made any other journey, as other people do, and you had no wish to know any other city or other laws, but you were contented with us and our city. So strongly did you prefer us and agree to live in accordance with us; and besides, you begat children in the city, showing that it pleased you. And moreover even at your trial you might have offered exile as your penalty, if you wished, and might have done with the state's consent what you are now undertaking to do without it. But you then put on airs and said you were not disturbed if you must die, and you preferred, as you said, death to exile. And now you are not ashamed to think of those words and you do not respect us, the laws, since you are trying to bring us to naught; and you are doing what the meanest slave would do, since you are trying to run away contrary to the compacts and agreements you made with us that you would live in accordance with us. First then, answer this question, whether we speak the truth or not when we say that you agreed, not in word, but by your acts, to live in accordance with us." What shall we say to this, Crito? Must we not agree that it is true?

C. We must, Socrates.

S. "Are you then," they would say, "not breaking your compacts and agreements with us, though you were not led into them by compulsion or fraud, and were not forced to make up your mind in a short time, but had seventy years, in which you could have gone away, if we did not please you and if you thought the agreements were unfair? But you preferred neither Lacedaemon nor Crete, which you are always saying are well governed, nor any other of the Greek states, or of the foreign ones, but you went away from this city less than the lame and the blind and the other cripples. So much more than the other Athenians were you satisfied with the city and evidently therefore with us, its laws; for who would be pleased with a city apart from its laws? And now will you not abide by your agreement? You will if you take our advice, Socrates; and you will not make yourself ridiculous by going away from the city.

"For consider. By transgressing in this way and committing these errors, what good will you do to yourself or any of your friends? For it is pretty clear that your friends also will be exposed to the risk of banishment and the loss of their homes in the city or of their property. And you yourself, if you go to one of the nearest cities, to Thebes or Megara—for both are well governed—will go as an enemy, Socrates, to their government, and all who care for their own cities will look askance at you, and will consider you a destroyer of the laws, and you will confirm the judges in their opinion, so that they will think their verdict was just. For he who is destroyer of the laws might certainly be regarded as a destroyer of young and thoughtless men. Will you then avoid the well-governed cities and the most civilized men? And if you do this will your life be worth living? Or will you go to them and have the face to carry on—what kind of conversation, Socrates? The same kind you carried on here, saying that virtue and justice and lawful things and the laws are the most precious things to men? And do you not think that the conduct of Socrates would seem most disgraceful? You cannot help thinking so. Or you will keep away from these places and go to Crito's friends in Thessaly; for there great disorder and lawlessness prevail, and perhaps

they would be amused to hear of the ludicrous way in which you ran away from prison by putting on a disguise, a peasant's leathern cloak or some of the other things in which runaways dress themselves up, and changing your appearance. But will no one say that you, an old man, who had probably but a short time yet to live, clung to life with such shameless greed that you transgressed the highest laws? Perhaps not, if you do not offend anyone; but if you do, Socrates, you will have to listen to many things that would be a disgrace to you. So you will live as an inferior and a slave to everyone. And what will you do except feast in Thessaly, as if you had gone to Thessaly to attend a banquet? What will become of our conversations about justice and virtue? But perhaps you wish to live for the sake of your children, that you may bring them up and educate them? How so? Will you take them to Thessaly to be brought up and educated, making exiles of them, that you may give them that blessing also? Or perhaps you will not do that, but if they are brought up here while you are living, will they be better brought up and educated if you are not with them than if you were dead? Oh yes! your friends will care for them. Will they care for them if you go away to Thessaly and not if you go away to the dwellings of the dead? If those who say they are your friends are of any use, we must believe they will care for them in both cases alike.

"Ah, Socrates, be guided by us who tended your infancy. Care neither for your children nor for life nor

for anything else more than for the right, that when you come to the home of the dead, you may have all these things to say in your own defence. For clearly if you do this thing it will not be better for you here, or more just or holier, no, nor for any of your friends, and neither will it be better when you reach that other abode. Now, however, you will go away wronged, if you do go away, not by us, the laws, but by men; but if you escape after so disgracefully requiting wrong with wrong and evil with evil, breaking your compacts and agreements with us, and injuring those whom you least ought to injure—yourself, your friends, your country and us—we shall be angry with you while you live, and there our brothers, the laws in Hades' realm, will not receive you graciously; for they will know that you tried, so far as in you lay, to destroy us. Do not let Crito persuade you to do what he says, but take our advice.

Be well assured, my dear friend, Crito, that this is what I seem to hear, as the frenzied dervishes of Cybele seem to hear the flutes, and this sound of these words re-echoes within me and prevents my hearing any other words. And be assured that, so far as I now believe, if you argue against these words you will speak in vain. Nevertheless, if you think you can accomplish anything, speak.

C. No, Socrates, I have nothing to say.

S. Then, Crito, let it be, and let us act in this way, since it is in this way that God leads us.

Study Questions

STOP.

Socrates refers to a "just agreement" that exists between citizens and the state (theorized by modern philosophers as the "social contract"). On what basis do individuals consent to this agreement, and what duties does it require of them as citizens? In his "Letter from Birmingham Jail," Martin Luther King, Jr., claims that Socrates was practising civil disobedience. Based on your reading of the *Crito*, do you agree or disagree?

When we think of democratic government, several slogans come to mind, such as "rule by

the people for the people," "one person, one vote," and "majority rules." In what ways does Plato criticize the ideas associated with these slogans?

- Given that Socrates' trial and execution followed the restoration of democratic government in Athens, what system of government might Plato have preferred instead? (Hint: he tells us we are to value the "good" opinions of "wise" people, i.e., those who understand "the just, beautiful, and good and their opposites.")

~~STORY~~**Martin Luther King, Jr.**

Martin Luther King, Jr. (1929–68) was born in Atlanta, Georgia. He had three degrees: a BA from Morehouse College, a BD from Crozer Theological Seminary, and a PhD from Boston University. King was pastor of a church in Montgomery, Alabama. He grew up during the time of segregation. After the Civil War, blacks were free but did not get equal rights. King very badly wanted to stop segregation. The Ku Klux Klan threatened his life, but the great Indian leader Mohandas Gandhi helped him understand not to use violence when fighting for civil rights. In "Letter from Birmingham Jail," King explains and defends this approach. King won the Nobel Peace Prize in 1964.—TG

Letter from Birmingham Jail¹

April 16, 1963

My Dear Fellow Clergymen:

While confined here in the Birmingham city jail, I came across your recent statement calling my present activities "unwise and untimely." Seldom do I pause to answer criticism of my work and ideas. If I sought to answer all the criticisms that cross my desk, my secretaries would have little time for anything other than such correspondence in the course of the day, and I would have no time for constructive work. But since I feel that you are men of genuine good will and that your criticisms are sincerely set forth, I want to try to answer your statement in what I hope will be patient and reasonable terms.

I think I should indicate why I am here in Birmingham, since you have been influenced by the view which argues against "outsiders coming in." I have the honour of serving as president of the Southern Christian Leadership Conference, an organization operating in every southern state, with headquarters in Atlanta, Georgia. We have some 85 affiliated organizations across the South, and one of them is the Alabama Christian Movement for Human Rights. Frequently we share staff, educational and financial resources with our affiliates. Several months ago the affiliate here in Birmingham asked us to be on call to engage in a non-violent direct-action program if such were deemed necessary. We readily consented, and when the hour came we lived up to our promise. So I, along with several members of my staff, am here because I was invited here. I am here because I have organizational ties here.

But more basically, I am in Birmingham because injustice is here. Just as the prophets of the eighth century BCE left their villages and carried their "thus saith the Lord" far beyond the boundaries of their home towns, and just as the Apostle Paul left his village of Tarsus and carried the gospel of Jesus Christ to the far corners of the Greco-Roman world, so am I compelled to carry the gospel of freedom beyond my own home town. Like Paul, I must constantly respond to the Macedonian call for aid.

Moreover, I am cognizant of the interrelatedness of all communities and states. I cannot sit idly by in Atlanta and not be concerned about what happens in Birmingham. **Injustice anywhere is a threat to justice everywhere.** We are caught in an inescapable network of mutuality, tied in a single garment of destiny. **Whatever affects one directly, affects all indirectly.** Never again can we afford to live with the narrow, provincial "outside agitator" idea. Anyone who lives inside the United States can never be considered an outsider anywhere within its bounds.

You deplore the demonstrations taking place in Birmingham. But your statement, I am sorry to say, fails to express a similar concern for the conditions that brought about the demonstrations. I am sure that none of you would want to rest content with the superficial kind of social analysis that deals merely with effects and does not grapple with underlying causes. It is unfortunate that demonstrations are taking place in Birmingham, but it is even more unfortunate that the city's white power structure left the Negro community with no alternative.

In any non-violent campaign there are four basic steps: collection of the facts to determine whether injustices exist; negotiation; self-purification; and direct action. We have gone through all these steps in Birmingham. There can be no gainsaying the fact that racial injustice engulfs this community. Birmingham is probably the most thoroughly segregated city in the United States. Its ugly record of brutality is widely known. Negroes have experienced grossly unjust treatment in the courts. There have been more unsolved bombings of Negro homes and churches in Birmingham than in any other city in the nation. These are the hard, brutal facts of the case. On the basis of these conditions, Negro leaders sought to negotiate with the city fathers. But the latter consistently refused to engage in good-faith negotiation.

Then, last September, came the opportunity to talk with leaders of Birmingham's economic community. In the course of the negotiations, certain promises were made by the merchants—for example, to remove the stores' humiliating racial signs. On the basis of these promises, the Reverend Fred Shuttlesworth and the leaders of the Alabama Christian Movement for Human Rights agreed to a moratorium on all demonstrations. As the weeks and months went by, we realized that we were the victims of a broken promise. A few signs, briefly removed, returned; the others remained.

As in so many past experiences, our hopes had been blasted, and the shadow of deep disappointment settled upon us. We had no alternative except to prepare for direct action, whereby we would present our very bodies as a means of laying our case before the conscience of the local and the national community. Mindful of the difficulties involved, we decided to undertake a process of self-purification. We began a series of workshops on non-violence, and we repeatedly asked ourselves: "Are you able to accept blows without retaliating?" "Are you able to endure the ordeal of jail?" We decided to schedule our direct-action program for the Easter season, realizing that except for Christmas, this is the main shopping period of the year. Knowing that a strong economic-withdrawal program would be the by-product of direct action, we felt that this would

be the best time to bring pressure to bear on the merchants for the needed change.

Then it occurred to us that Birmingham's mayoral election was coming up in March, and we speedily decided to postpone action until after election day. When we discovered that the Commissioner of Public Safety, Eugene "Bull" Connor, had piled up enough votes to be in the run-off, we decided again to postpone action until the day after the run-off so that the demonstrations could not be used to cloud the issues. Like many others, we waited to see Mr Connor defeated, and to this end we endured postponement after postponement. Having aided in this community need, we felt that our direct-action program could be delayed no longer.

You may well ask: "Why direct action? Why sit-ins, marches and so forth? Isn't negotiation a better path?" You are quite right in calling for negotiation. Indeed, this is the very purpose of direct action. Non-violent direct action seeks to create such a crisis and foster such a tension that a community which has constantly refused to negotiate is forced to confront the issue. It seeks so to dramatize the issue that it can no longer be ignored. My citing the creation of tension as part of the work of the non-violent-resister may sound rather shocking. But I must confess that I am not afraid of the word "tension." I have earnestly opposed violent tension, but there is a type of constructive, non-violent tension which is necessary for growth. Just as Socrates felt that it was necessary to create a tension in the mind so that individuals could rise from the bondage of myths and half-truths to the unfettered realm of creative analysis and objective appraisal, so must we see the need for non-violent gadflies to create the kind of tension in society that will help men rise from the dark depths of prejudice and racism to the majestic heights of understanding and brotherhood.

The purpose of our direct-action program is to create a situation so crisis-packed that it will inevitably open the door to negotiation. I therefore concur with you in your call for negotiation. Too long has our beloved Southland been bogged down in a tragic effort to live in monologue rather than dialogue.

One of the basic points in your statement is that the action that I and my associates have taken in Birmingham is untimely. Some have asked: "Why didn't you give the new city administration time to act?" The only answer that I can give to this query is that the new Birmingham administration must be prodded about as much as the outgoing one, before it will act. We are sadly mistaken if we feel that the election of Albert Boutwell as mayor will bring the millennium to Birmingham. While Mr. Boutwell is a much more gentle person than Mr. Connor, they are both segregationists, dedicated to maintenance of the status quo. I have hope that Mr. Boutwell will be reasonable enough to see the futility of massive resistance to desegregation. But he will not see this without pressure from devotees of civil rights. My friends, I must say to you that we have not made a single gain in civil rights without determined legal and non-violent pressure. Lamentably, it is an historical fact that privileged groups seldom give up their privileges voluntarily. Individuals may see the moral light and voluntarily give up their unjust posture; but, as Reinhold Niebuhr has reminded us, groups tend to be more immoral than individuals.

We know through painful experience that freedom is never voluntarily given by the oppressor; it must be demanded by the oppressed. Frankly, I have yet to engage in a direct-action campaign that was "well timed" in the view of those who have not suffered unduly from the disease of segregation. For years now I have heard the word "Wait!" It rings in the ear of every Negro with piercing familiarity. This "Wait" has almost always meant "Never." We must come to see, with one of our distinguished jurists, that "justice too long delayed is justice denied."

We have waited for more than 340 years for our constitutional and God-given rights. The nations of Asia and Africa are moving with jet-like speed toward gaining political independence, but we still creep at horse-and-buggy pace toward gaining a cup of coffee at a lunch counter. Perhaps it is easy for those who have never felt the stinging darts of segregation to say, "Wait." But when you have seen vicious mobs lynch your mothers and fathers at will and drown your sisters and brothers at whim; when you have

seen hate-filled policemen curse, kick and even kill your black brothers and sisters; when you see the vast majority of your twenty million Negro brothers smothering in an airtight cage of poverty in the midst of an affluent society; when you suddenly find your tongue twisted and your speech stammering as you seek to explain to your six-year-old daughter why she can't go to the public amusement park that has just been advertised on television, and see tears welling up in her eyes when she is told that Funtown is closed to coloured children, and see ominous clouds of inferiority beginning to form in her little mental sky, and see her beginning to distort her personality by developing an unconscious bitterness toward white people; when you have to concoct an answer for a five-year-old son who is asking: "Daddy, why do white people treat coloured people so mean?"; when you take a cross-county drive and find it necessary to sleep night after night in the uncomfortable corners of your automobile because no motel will accept you; when you are humiliated day in and day out by nagging signs reading "white" and "coloured"; when your first name becomes "nigger," your middle name becomes "boy" (however old you are) and your last name becomes "John," and your wife and mother are never given the respected title "Mrs"; when you are harried by day and haunted by night by the fact that you are a Negro, living constantly at tiptoe stance, never quite knowing what to expect next, and are plagued with inner fears and outer resentments; when you are forever fighting a degenerating sense of "nobodiness"—then you will understand why we find it difficult to wait. There comes a time when the cup of endurance runs over, and men are no longer willing to be plunged into the abyss of despair. I hope, sirs, you can understand our legitimate and unavoidable impatience.

You express a great deal of anxiety over our willingness to break laws. This is certainly a legitimate concern. Since we so diligently urge people to obey the Supreme Court's decision of 1954 outlawing segregation in the public schools, at first glance it may seem rather paradoxical for us consciously to break laws. One may well ask: "How can you advocate breaking some laws and obeying

others?" The answer lies in the fact that there are two types of laws: just and unjust. I would be the first to advocate obeying just laws. One has not only a legal but a moral responsibility to obey just laws. Conversely, one has a moral responsibility to disobey unjust laws. I would agree with St Augustine that "an unjust law is no law at all."

Now, what is the difference between the two? How does one determine whether a law is just or unjust? A just law is a man-made code that squares with the moral law or the law of God. An unjust law is a code that is out of harmony with the moral law. To put it in the terms of St Thomas Aquinas: An unjust law is a human law that is not rooted in eternal law and natural law. Any law that uplifts human personality is just. Any law that degrades human personality is unjust. All segregation statutes are unjust because segregation distorts the soul and damages the personality. It gives the segregator a false sense of superiority and the segregated a false sense of inferiority. Segregation, to use the terminology of the Jewish philosopher Martin Buber, substitutes an "I-it" relationship for an "I-thou" relationship and ends up relegating persons to the status of things. Hence segregation is not only politically, economically and sociologically unsound, it is morally wrong and sinful. Paul Tillich has said that sin is separation. Is not segregation an existential expression of man's tragic separation, his awful estrangement, his terrible sinfulness? Thus it is that I can urge men to obey the 1954 decision of the Supreme Court, for it is morally right; and I can urge them to disobey segregation ordinances, for they are morally wrong.

Let us consider a more concrete example of just and unjust laws. An unjust law is a code that a numerical or power majority group compels a minority group to obey but does not make binding on itself. This is *difference made legal*. By the same token, a just law is a code that a majority compels a minority to follow and that it is willing to follow itself. This is *sameness made legal*.

Let me give another explanation. A law is unjust if it is inflicted on a minority that, as a result of being denied the right to vote, had no part in enacting or devising the law. Who can say that the legislature of

Alabama which set up that state's segregation laws was democratically elected? Throughout Alabama all sorts of devious methods are used to prevent Negroes from becoming registered voters, and there are some counties in which, even though Negroes constitute a majority of the population, not a single Negro is registered. Can any law enacted under such circumstances be considered democratically structured?

Sometimes a law is just on its face and unjust in its application. For instance, I have been arrested on a charge of parading without a permit. Now, there is nothing wrong in having an ordinance which requires a permit for a parade. But such an ordinance becomes unjust when it is used to maintain segregation and to deny citizens the First-Amendment privilege of peaceful assembly and protest.

I hope you are able to see the distinction I am trying to point out. In no sense do I advocate evading or defying the law, as would the rabid segregationist. That would lead to anarchy. One who breaks an unjust law must do so openly, lovingly, and with a willingness to accept the penalty. I submit that an individual who breaks a law that conscience tells him is unjust, and who willingly accepts the penalty of imprisonment in order to arouse the conscience of the community over its injustice, is in reality expressing the highest respect for law.

Of course, there is nothing new about this kind of civil disobedience. It was evidenced sublimely in the refusal of Shadrach, Meshach, and Abednego to obey the laws of Nebuchadnezzar, on the ground that a higher moral law was at stake. It was practised superbly by the early Christians, who were willing to face hungry lions and the excruciating pain of chopping blocks rather than submit to certain unjust laws of the Roman Empire. To a degree, academic freedom is a reality today because Socrates practiced civil disobedience. In our own nation, the Boston Tea Party represented a massive act of civil disobedience.

We should never forget that everything Adolf Hitler did in Germany was "legal" and everything the Hungarian freedom fighters did in Hungary was "illegal." It was "illegal" to aid and comfort a Jew

in Hitler's Germany. Even so, I am sure that, had I lived in Germany at the time, I would have aided and comforted my Jewish brothers. If today I lived in a Communist country where certain principles dear to the Christian faith are suppressed, I would openly advocate disobeying that country's anti-religious laws.

I must make two honest confessions to you, my Christian and Jewish brothers. First, I must confess that over the past few years I have been gravely disappointed with the white moderate. I have almost reached the regrettable conclusion that the Negro's great stumbling block in his stride toward freedom is not the **White Citizen's Counciler or the Ku Klux Klanner**, but the white moderate, who is more devoted to "order" than to justice; who prefers a negative peace which is the absence of tension to a positive peace which is the presence of justice; who constantly says: "I agree with you in the goal you seek, but I cannot agree with your methods of direct action"; who paternalistically believes he can set the timetable for another man's freedom; who lives by a mythical concept of time and who constantly advises the Negro to wait for a "more convenient season." Shallow understanding from people of good will is more frustrating than absolute misunderstanding from people of ill will. Lukewarm acceptance is much more bewildering than outright rejection.

I had hoped that the white moderate would understand that law and order exist for the purpose of establishing justice and that when they fail in this purpose they become the dangerously structured dams that block the flow of social progress. I had hoped that the white moderate would understand that the present tension in the South is a necessary phase of the transition from an obnoxious negative peace, in which the Negro passively accepted his unjust plight, to a substantive and positive peace, in which all men will respect the dignity and worth of human personality. Actually, we who engage in non-violent direct action are not the creators of tension. We merely bring to the surface the hidden tension that is already alive. We bring it out in the open, where it can be seen and dealt with. Like a boil that can never be cured so long as it is covered up but must be opened with all its ugliness to the

natural medicines of air and light, injustice must be exposed, with all the tension its exposure creates, to the light of human conscience and the air of national opinion before it can be cured.

In your statement you assert that our actions, even though peaceful, must be condemned because they precipitate violence. But is this a logical assertion? Isn't this like condemning a robbed man because his possession of money precipitated the evil act of robbery? Isn't this like condemning Socrates because his unswerving commitment to truth and his philosophical inquiries precipitated the act by the misguided populace in which they made him drink hemlock? Isn't this like condemning Jesus because his unique God-consciousness and never-ceasing devotion to God's will precipitated the evil act of crucifixion? We must come to see that, as the federal courts have consistently affirmed, it is wrong to urge an individual to cease his efforts to gain his basic constitutional rights because the quest may precipitate violence. Society must protect the robbed and punish the robber.

I had also hoped that the white moderate would reject the myth concerning time in relation to the struggle for freedom. I have just received a letter from a white brother in Texas. He writes: "All Christians know that the coloured people will receive equal rights eventually, but it is possible that you are in too great a religious hurry. It has taken Christianity almost two thousand years to accomplish what it has. The teachings of Christ take time to come to earth." Such an attitude stems from a tragic misconception of time, from the strangely irrational notion that there is something in the very flow of time that will inevitably cure all ills. Actually, time itself is neutral; it can be used either destructively or constructively. More and more I feel that the people of ill will have used time much more effectively than have the people of good will. We will have to repent in this generation not merely for the hateful words and actions of the bad people but for the appalling silence of the good people. Human progress never rolls in on wheels of inevitability; it comes through the tireless efforts of men willing to be co-workers with God, and without this hard work, time itself

becomes an ally of the forces of social stagnation. We must use time creatively, in the knowledge that the time is always ripe to do right. Now is the time to make real the promise of democracy and transform our pending national elegy into a creative psalm of brotherhood. Now is the time to lift our national policy from the quicksand of racial injustice to the solid rock of human dignity.

You speak of our activity in Birmingham as extreme. At first I was rather disappointed that fellow clergymen would see my non-violent efforts as those of an extremist. I began thinking about the fact that I stand in the middle of two opposing forces in the Negro community. One is a force of complacency, made up in part of Negroes who, as a result of long years of oppression, are so drained of self-respect and a sense of "somebodiness" that they have adjusted to segregation; and in part of a few middle-class Negroes who, because of a degree of academic and economic security and because in some ways they profit by segregation, have become insensitive to the problems of the masses. The other force is one of bitterness and hatred, and it comes perilously close to advocating violence. It is expressed in the various black nationalist groups that are springing up across the nation, the largest and best known being **Elijah Muhammad's** Muslim movement. Nourished by the Negro's frustration over the continued existence of racial discrimination, this movement is made up of people who have lost faith in America, who have absolutely repudiated Christianity, and who have concluded that the white man is an incorrigible "devil."

I have tried to stand between these two forces, saying that we need emulate neither the "do nothingism" of the complacent nor the hatred and despair of the black nationalist. For there is the more excellent way of love and non-violent protest. I am grateful to God that, through the influence of the Negro church, the way of non-violence became an integral part of our struggle.

If this philosophy had not emerged, by now many streets of the South would, I am convinced, be flowing with blood. And I am further convinced that if our white brothers dismiss as "rabble rousers"

and "outside agitators" those of us who employ non-violent direct action, and if they refuse to support our non-violent efforts, millions of Negroes will, out of frustration and despair, seek solace and security in black-nationalist ideologies—a development that would inevitably lead to a frightening racial nightmare.

Oppressed people cannot remain oppressed forever. The yearning for freedom eventually manifests itself, and that is what has happened to the American Negro. Something within has reminded him of his birthright of freedom, and something without has reminded him that it can be gained. Consciously or unconsciously, he has been caught up by the *Zeitgeist* [spirit of the time], and with his black brothers of Africa and his brown and yellow brothers of Asia, South America and the Caribbean, the United States Negro is moving with a sense of great urgency toward the promised land of racial justice. If one recognizes this vital urge that has engulfed the Negro community, one should readily understand why public demonstrations are taking place. The Negro has many pent-up resentments and latent frustrations, and he must release them. So let him march; let him make prayer pilgrimages to the city hall; let him go on freedom rides—and try to understand why he must do so. If his repressed emotions are not released in non-violent ways, they will seek expression through violence; this is not a threat but a fact of history. So I have not said to my people: "Get rid of your discontent." Rather, I have tried to say that this normal and healthy discontent can be channelled into the creative outlet of non-violent direct action. And now this approach is being termed extremist.

But though I was initially disappointed at being categorized as an extremist, as I continued to think about the matter I gradually gained a measure of satisfaction from the label. Was not Jesus an extremist for love: "Love your enemies, bless them that curse you, do good to them that hate you, and pray for them which despitefully use you, and persecute you." Was not Amos an extremist for justice: "Let justice roll down like waters and righteousness like an ever-flowing stream." Was not Paul an extremist

for the Christian gospel: "I bear in my body the marks of the Lord Jesus." Was not **Martin Luther** an extremist: "Here I stand; I cannot do otherwise, so help me God." And **John Bunyan**: "I will stay in jail to the end of my days before I make a butchery of my conscience." And **Abraham Lincoln**: "This nation cannot survive half slave and half free." And **Thomas Jefferson**: "We hold these truths to be self-evident, that all men are created equal . . ." So the question is not whether we will be extremists, but what kind of extremists we will be. Will we be extremists for hate or for love? Will we be extremists for the preservation of injustice or for the extension of justice? In that dramatic scene on Calvary's hill three men were crucified. We must never forget that all three were crucified for the same crime—the crime of extremism. Two were extremists for immorality, and thus fell below their environment. The other, Jesus Christ, was an extremist for love, truth and goodness, and thereby rose above his environment. Perhaps the South, the nation and the world are in dire need of creative extremists.

I had hoped that the white moderate would see this need. Perhaps I was too optimistic; perhaps I expected too much. I suppose I should have realized that few members of the oppressor race can understand the deep groans and passionate yearnings of the oppressed race, and still fewer have the vision to see that injustice must be rooted out by strong, persistent and determined action. I am thankful, however, that some of our white brothers in the South have grasped the meaning of this social revolution and committed themselves to it. They are still all too few in quantity, but they are big in quality. Some—such as Ralph McGill, Lillian Smith, Harry Golden, James McBride Dabbs, Ann Braden and Sarah Patton Boyle—have written about our struggle in eloquent and prophetic terms. Others have marched with us down nameless streets of the South. They have languished in filthy, roach-infested jails, suffering the abuse and brutality of policemen who view them as "dirty nigger-lovers." Unlike so many of their moderate brothers and sisters, they have recognized the urgency of the moment and sensed the need for powerful "action" antidotes to combat the disease of segregation.

Let me take note of my other major disappointment. I have been so greatly disappointed with the white church and its leadership. Of course, there are some notable exceptions. I am not unmindful of the fact that each of you has taken some significant stands on this issue. I commend you, Reverend Stallings, for your Christian stand on this past Sunday, in welcoming Negroes to your worship service on a non-segregated basis. I commend the Catholic leaders of this state for integrating Spring Hill College several years ago.

But despite these notable exceptions, I must honestly reiterate that I have been disappointed with the church. I do not say this as one of those negative critics who can always find something wrong with the church. I say this as a minister of the gospel, who loves the church; who was nurtured in its bosom; who has been sustained by its spiritual blessings and who will remain true to it as long as the cord of life shall lengthen.

When I was suddenly catapulted into the leadership of the bus protest in Montgomery, Alabama, a few years ago, I felt we would be supported by the white church. I felt that the white ministers, priests and rabbis of the South would be among our strongest allies. Instead, some have been outright opponents, refusing to understand the freedom movement and misrepresenting its leaders; all too many others have been more cautious than courageous and have remained silent behind the anaesthetizing security of stained-glass windows.

In spite of my shattered dreams, I came to Birmingham with the hope that the white religious leadership of this community would see the justice of our cause and, with deep moral concern, would serve as the channel through which our just grievances could reach the power structure. I had hoped that each of you would understand. But again I have been disappointed.

I have heard numerous southern religious leaders admonish their worshipers to comply with a desegregation decision because it is the law, but I have longed to hear white ministers declare: "Follow this decree because integration is morally right and because the Negro is your brother." In the midst of

blatant injustices inflicted upon the Negro, I have watched white churchmen stand on the sideline and mouth pious irrelevancies and sanctimonious trivialities. In the midst of a mighty struggle to rid our nation of racial and economic injustice, I have heard many ministers say: "Those are social issues, with which the gospel has no real concern." And I have watched many churches commit themselves to a completely otherworldly religion which makes a strange, un-Biblical distinction between body and soul, between the sacred and the secular.

I have travelled the length and breadth of Alabama, Mississippi and all the other southern states. On sweltering summer days and crisp autumn mornings I have looked at the South's beautiful churches with their lofty spires pointing heavenward. I have beheld the impressive outlines of her massive religious-education buildings. Over and over I have found myself asking: "What kind of people worship here? Who is their God? Where were their voices when the lips of Governor **Barnett** dripped with words of interposition and nullification? Where were they when Governor **Wallace** gave a clarion call for defiance and hatred? Where were their voices of support when bruised and weary Negro men and women decided to rise from the dark dungeons of complacency to the bright hills of creative protest?"

Yes, these questions are still in my mind. In deep disappointment I have wept over the laxity of the church. But be assured that my tears have been tears of love. There can be no deep disappointment where there is not deep love. Yes, I love the church. How could I do otherwise? I am in the rather unique position of being the son, the grandson and the great-grandson of preachers. Yes, I see the church as the body of Christ. But, oh! How we have blemished and scarred that body through social neglect and through fear of being non-conformists.

There was a time when the church was very powerful—in the time when the early Christians rejoiced at being deemed worthy to suffer for what they believed. In those days the church was not merely a thermometer that recorded the ideas and principles of popular opinion; it was a thermostat that transformed the mores of society. Whenever the

early Christians entered a town, the people in power became disturbed and immediately sought to convict the Christians for being "disturbers of the peace" and "outside agitators." But the Christians pressed on, in the conviction that they were "a colony of heaven," called to obey God rather than man. Small in number, they were big in commitment. They were too God-intoxicated to be "astronomically intimidated." By their effort and example they brought an end to such ancient evils as infanticide and gladiatorial contests.

Things are different now. So often the contemporary church is a weak, ineffectual voice with an uncertain sound. So often it is an archdefender of the status quo. Far from being disturbed by the presence of the church, the power structure of the average community is consoled by the church's silent—and often even vocal—sanction of things as they are.

But the judgment of God is upon the church as never before. If today's church does not recapture the sacrificial spirit of the early church, it will lose its authenticity, forfeit the loyalty of millions, and be dismissed as an irrelevant social club with no meaning for the twentieth century. Every day I meet young people whose disappointment with the church has turned into outright disgust.

Perhaps I have once again been too optimistic. Is organized religion too inextricably bound to the status quo to save our nation and the world? Perhaps I must turn my faith to the inner spiritual church, the church within the church, as the true *ekklesia* and the hope of the world. But again I am thankful to God that some noble souls from the ranks of organized religion have broken loose from the paralyzing chains of conformity and joined us as active partners in the struggle for freedom. They have left their secure congregations and walked the streets of Albany, Georgia, with us. They have gone down the highways of the South on tortuous rides for freedom. Yes, they have gone to jail with us. Some have been dismissed from their churches, have lost the support of their bishops and fellow ministers. But they have acted in the faith that right defeated is stronger than evil triumphant. Their witness has been the spiritual salt that has preserved

the true meaning of the gospel in these troubled times. They have carved a tunnel of hope through the dark mountain of disappointment.

I hope the church as a whole will meet the challenge of this decisive hour. But even if the church does not come to the aid of justice, I have no despair about the future. I have no fear about the outcome of our struggle in Birmingham, even if our motives are at present misunderstood. We will reach the goal of freedom in Birmingham and all over the nation, because the goal of America is freedom. Abused and scorned though we may be, our destiny is tied up with America's destiny. Before the pilgrims landed at Plymouth, we were here. Before the pen of Jefferson etched the majestic words of the Declaration of Independence across the pages of history, we were here. For more than two centuries our forebears laboured in this country without wages; they made cotton king; they built the homes of their masters while suffering gross injustice and shameful humiliation—and yet out of a bottomless vitality they continued to thrive and develop. If the inexpressible cruelties of slavery could not stop us, the opposition we now face will surely fail. We will win our freedom because the sacred heritage of our nation and the eternal will of God are embodied in our echoing demands.

Before closing I feel impelled to mention one other point in your statement that has troubled me profoundly. You warmly commended the Birmingham police force for keeping "order" and "preventing violence." I doubt that you would have so warmly commended the police force if you had seen its dogs sinking their teeth into unarmed, non-violent Negroes. I doubt that you would so quickly commend the policemen if you were to observe their ugly and inhumane treatment of Negroes here in the city jail; if you were to watch them push and curse old Negro women and young Negro girls; if you were to see them slap and kick old Negro men and young boys; if you were to observe them, as they did on two occasions, refuse to give us food because we wanted to sing our grace together. I cannot join you in your praise of the Birmingham police department.

It is true that the police have exercised a degree of discipline in handling the demonstrators. In

this sense they have conducted themselves rather "non-violently" in public. But for what purpose? To preserve the evil system of segregation. Over the past few years I have consistently preached that non-violence demands that the means we use must be as pure as the ends we seek. I have tried to make clear that it is wrong to use immoral means to attain moral ends. But now I must affirm that it is just as wrong, or perhaps even more so, to use moral means to preserve immoral ends. Perhaps Mr. Connor and his policemen have been rather non-violent in public, as was Chief Pritchett in Albany, Georgia, but they have used the moral means of non-violence to maintain the immoral end of racial injustice. As T.S. Eliot has said: "The last temptation is the greatest treason: To do the right deed for the wrong reason."

I wish you had commended the Negro sit-inners and demonstrators of Birmingham for their sublime courage, their willingness to suffer and their amazing discipline in the midst of great provocation. One day the South will recognize its real heroes. They will be the **James Merediths**, with the noble sense of purpose that enables them to face jeering and hostile mobs, and with the agonizing loneliness that characterizes the life of the pioneer. They will be old, oppressed, battered Negro women, symbolized in a 72-year-old woman in Montgomery, Alabama, who rose up with a sense of dignity and with her people decided not to ride segregated buses, and who responded with ungrammatical profundity to one who inquired about her weariness: "My feets is tired, but my soul is at rest." They will be the young high school and college students, the young ministers of the gospel and a host of their elders, courageously and non-violently sitting in at lunch counters and willingly going to jail for conscience' sake. One day children of God sat down at lunch counters, they were in reality standing up for what is best in the American dream and for the most sacred values in our Judaeo-Christian heritage, thereby bringing our nation back to those great wells of democracy which were dug deep by the founding fathers in their formulation of the Constitution and the Declaration of Independence.

Never before have I written so long a letter. I'm afraid it is much too long to take your precious time. I can assure you that it would have been much shorter if I had been writing from a comfortable desk, but what else can one do when he is alone in a narrow jail cell, other than write long letters, think long thoughts, and pray long prayers?

If I have said anything in this letter that overstates the truth and indicates an unreasonable impatience, I beg you to forgive me. If I have said anything that understates the truth and indicates my having a patience that allows me to settle for anything less than brotherhood, I beg God to forgive me.

I hope this letter finds you strong in the faith. I also hope that circumstances will soon make it possible for me to meet each of you, not as an integrationist or a civil-rights leader but as a fellow clergyman and a Christian brother. Let us all hope that the dark clouds of racial prejudice will soon pass away and the deep fog of misunderstanding will be lifted from our fear-drenched communities, and in some not too distant tomorrow the radiant stars of love and brotherhood will shine over our great nation with all their scintillating beauty.

Yours for the cause of Peace and Brotherhood,
Martin Luther King, Jr.

Note

1. Author's Note: This response to a published statement by eight fellow clergymen from Alabama [...] was composed under somewhat constricting circumstances. Begun on the margins of the newspaper in which the statement appeared while I was in jail, the letter was continued on

scraps of writing paper supplied by a friendly Negro trusty, and concluded on a pad my attorneys were eventually permitted to leave me. Although the text remains in substance unaltered, I have indulged in the author's prerogative of polishing it for publication.

Study Questions

1. What is King's response to the clergymen's portrayal of him as an "outsider" stirring up trouble in Birmingham? Do you agree or disagree?
2. Describe the four basic steps to be carried out in a non-violent campaign.
3. What is the Socratic method, and how is it used pedagogically? Why did King believe that the civil rights movement's use of non-violent civil disobedience played an analogous role in facilitating social and political change?
4. Provide an example of an act of non-violent civil disobedience that illustrates the apparent paradox that demonstrators consciously break some laws while urging people to uphold others. King resolves the paradox by saying that there are just and unjust laws. Critically evaluate the various ways in which King draws this distinction.

518

John Rawls

John Rawls (1921–2002) is considered the most influential political philosopher of the twentieth century. Born in Baltimore, Rawls obtained his BA at Princeton in 1943, served in the Pacific during the Second World War, and returned to Princeton to earn his PhD in 1950. After a Fulbright year spent at Oxford in 1952–3, Rawls taught at Cornell, MIT, and, for more than 30 years, Harvard. Rawls' *A Theory of Justice* (1971) critiques utilitarianism and draws on the social contract tradition to defend a theory of justice as fairness that accommodates tension between democratic ideals of freedom and equality. These ideas are introduced in this 1969 essay on civil disobedience.

The Justification of Civil Disobedience

1. Introduction

I should like to discuss briefly, and in an informal way, the grounds of civil disobedience in a constitutional democracy. Thus, I shall limit my remarks to the conditions under which we may, by civil disobedience, properly oppose legally established democratic authority; I am not concerned with the situation under other kinds of government nor, except incidentally, with other forms of resistance. My thought is that in a reasonably just (though of course not perfectly just) democratic regime, civil disobedience, when it is justified, is normally to be understood as a political action which addresses the sense of justice of the majority in order to urge reconsideration of the measures protested and to warn that in the firm opinion of the dissenters the conditions of social cooperation are not being honoured. This characterization of civil disobedience is intended to apply to dissent on fundamental questions of internal policy, a limitation which I shall follow to simplify our question.

2. The Social Contract Doctrine

[. . .] I believe that the appropriate conception [of justice], at least for an account of political obligation in a constitutional democracy, is that of the social contract theory from which so much of our political thought derives. If we are careful to

interpret it in a suitably general way, I hold that this doctrine provides a satisfactory basis for political theory, indeed even for ethical theory itself, but this is beyond our present concern.¹ The interpretation I suggest is the following: that the principles to which social arrangements must conform, and in particular the principles of justice, are those which free and rational men would agree to in an original position of equal liberty; and similarly, the principles which govern men's relations to institutions and define their natural duties and obligations are the principles to which they would consent when so situated. It should be noted straightaway that in this interpretation of the contract theory the principles of justice are understood as the outcome of a hypothetical agreement. They are principles which would be agreed to if the situation of the original position were to arise. There is no mention of an actual agreement, nor need such an agreement ever be made. Social arrangements are just or unjust according to whether they accord with the principles for assigning and securing fundamental rights and liberties which would be chosen in the original position. This position is, to be sure, the analytic analogue of the traditional notion of the state of nature, but it must not be mistaken for a historical occasion. Rather it is a hypothetical situation which embodies the basic ideas of the contract doctrine; the description of this situation enables us to work out which principles would be

adopted. I must now say something about these matters.

The contract doctrine has always supposed that the persons in the original position have equal powers and rights, that is, that they are symmetrically situated with respect to any arrangements for reaching agreement, and that coalitions and the like are excluded. But it is an essential element (which has not been sufficiently observed although it is implicit in Kant's version of the theory) that there are very strong restrictions on what the contracting parties are presumed to know. In particular, I interpret the theory to hold that the parties do not know their position in society, past, present, or future; nor do they know which institutions exist. Again, they do not know their own place in the distribution of natural talents and abilities, whether they are intelligent or strong, man or woman, and so on. Finally, they do not know their own particular interests and preferences or the system of ends which they wish to advance: they do not know their conception of the good. In all these respects the parties are confronted with a veil of ignorance which prevents any one from being able to take advantage of his good fortune or particular interests or from being disadvantaged by them. What the parties do know (or assume) is that Hume's circumstances of justice obtain: namely, that the bounty of nature is not so generous as to render cooperative schemes superfluous nor so harsh as to make them impossible. Moreover, they assume that the extent of their altruism is limited and that, in general, they do not take an interest in one another's interests. Thus, given the special features of the original position, each man tries to do the best he can for himself by insisting on principles calculated to protect and advance his system of ends whatever it turns out to be.

I believe that as a consequence of the peculiar nature of the original position there would be an agreement on the following two principles for assigning rights and duties and for regulating distributive shares as these are determined by the fundamental institutions of society: first, each person is to have an equal right to the most extensive liberty compatible with a like liberty for all; second,

social and economic inequalities (as defined by the institutional structure or fostered by it) are to be arranged so that they are both to everyone's advantage and attached to positions and offices open to all. In view of the content of these two principles and their application to the main institutions of society, and therefore to the social system as a whole, we may regard them as the two principles of justice. Basic social arrangements are just insofar as they conform to these principles, and we can, if we like, discuss questions of justice directly by reference to them. But a deeper understanding of the justification of civil disobedience requires, I think, an account of the derivation of these principles provided by the doctrine of the social contract. Part of our task is to show why this is so.

3. The Grounds of Compliance with an Unjust Law

If we assume that in the original position men would agree both to the principle of doing their part when they have accepted and plan to continue to accept the benefits of just institutions (the principle of fairness), and also to the principle of not preventing the establishment of just institutions and of upholding and complying with them when they do exist, then the contract doctrine easily accounts for our having to conform to just institutions. But how does it account for the fact that we are normally required to comply with unjust laws as well? The injustice of a law is not a sufficient ground for not complying with it any more than the legal validity of legislation is always sufficient to require obedience to it. Sometimes one hears these extremes asserted, but I think that we need not take them seriously.

An answer to our question can be given by elaborating the social contract theory in the following way. I interpret it to hold that one is to envisage a series of agreements as follows: first, men are to agree upon the principles of justice in the original position. Then they are to move to a constitutional convention in which they choose a constitution that satisfies the principles of justice already chosen.

5109

Finally they assume the role of a legislative body and, guided by the principles of justice, enact laws subject to the constraints and procedures of the just constitution. The decisions reached in any stage are binding in all subsequent stages. Now whereas in the original position the contracting parties have no knowledge of their society or of their own position in it, in both a constitutional convention and a legislature, they do know certain general facts about their institutions, for example, the statistics regarding employment and output required for fiscal and economic policy. But no one knows particular facts about his own social class or his place in the distribution of natural assets. On each occasion the contracting parties have the knowledge required to make their agreement rational from the appropriate point of view, but not so much as to make them prejudiced. They are unable to tailor principles and legislation to take advantage of their social or natural position; a veil of ignorance prevents their knowing what this position is. With this series of agreements in mind, we can characterize just laws and policies as those which would be enacted were this whole process correctly carried out.

In choosing a constitution the aim is to find among the just constitutions the one which is most likely, given the general facts about the society in question, to lead to just and effective legislation. The principles of justice provide a criterion for the laws desired; the problem is to find a set of political procedures that will give this outcome. I shall assume that, at least under the normal conditions of a modern state, the best constitution is some form of democratic regime affirming equal political liberty and using some sort of majority (or other plurality) rule. Thus it follows that on the contract theory a constitutional democracy of some sort is required by the principles of justice. At the same time it is essential to observe that the constitutional process is always a case of what we may call imperfect procedural justice: that is, there is no feasible political procedure which guarantees that the enacted legislation is just even though we have (let us suppose) a standard for just legislation. In simple cases, such as games of fair division,

there are procedures which always lead to the right outcome (assume that equal shares is fair and let the man who cuts the cake take the last piece). These situations are those of perfect procedural justice. In other cases it does not matter what the outcome is as long as the fair procedure is followed: fairness of the process is transferred to the result (fair gambling is an instance of this). These situations are those of pure procedural justice. The constitutional process, like a criminal trial, resembles neither of these; the result matters and we have a standard for it. The difficulty is that we cannot frame a procedure which guarantees that only just and effective legislation is enacted. Thus even under a just constitution unjust laws may be passed and unjust policies enforced. Some form of the majority principle is necessary but the majority may be mistaken, more or less wilfully, in what it legislates. In agreeing to a democratic constitution (as an instance of imperfect procedural justice) one accepts at the same time the principle of majority rule. Assuming that the constitution is just and that we have accepted and plan to continue to accept its benefits, we then have both an obligation and a natural duty (and in any case the duty) to comply with what the majority enacts even though it may be unjust. In this way we become bound to follow unjust laws, not always, of course, but provided the injustice does not exceed certain limits. We recognize that we must run the risk of suffering from the defects of one another's sense of justice; this burden we are prepared to carry as long as it is more or less evenly distributed or does not weigh too heavily. Justice binds us to a just constitution and to the unjust laws which may be enacted under it in precisely the same way that it binds us to any other social arrangement. Once we take the sequence of stages into account, there is nothing unusual in our being required to comply with unjust laws.

It should be observed that the majority principle has a secondary place as a rule of procedure which is perhaps the most efficient one under usual circumstances for working a democratic constitution. The basis for it rests essentially upon the principles of justice and therefore we may, when conditions allow, appeal to these principles against unjust

legislation. The justice of the constitution does not ensure the justice of laws enacted under it; and while we often have both an obligation and a duty to comply with what the majority legislates (as long as it does not exceed certain limits), there is, of course, no corresponding obligation or duty to regard what the majority enacts as itself just. The right to make law does not guarantee that the decision is rightly made; and while the citizen submits in his conduct to the judgment of democratic authority, he does not submit his judgment to it. And if in his judgment the enactments of the majority exceed certain bounds of injustice, the citizen may consider civil disobedience. For we are not required to accept the majority's acts unconditionally and to acquiesce in the denial of our and others' liberties; rather we submit our conduct to democratic authority to the extent necessary to share the burden of working a constitutional regime, distorted as it must inevitably be by men's lack of wisdom and the defects of their sense of justice.

4. The Place of Civil Disobedience in a Constitutional Democracy

We are now in a position to say a few things about civil disobedience. I shall understand it to be a public, non-violent, and conscientious act contrary to law usually done with the intent to bring about a change in the policies or laws of the government (Bedau). Civil disobedience is a political act in the sense that it is an act justified by moral principles which define a conception of civil society and the public good. It rests, then, on political conviction as opposed to a search for self or group interest; and in the case of a constitutional democracy, we may assume that this conviction involves the conception of justice (say that expressed by the contract doctrine) which underlies the constitution itself. That is, in a viable democratic regime there is a common conception of justice by reference to which its citizens regulate their political affairs and interpret the constitution. Civil disobedience is a public act which the dissenter believes to be justified by this conception of justice, and for this reason it may be understood

as addressing the sense of justice of the majority in order to urge reconsideration of the measures protested and to warn that, in the sincere opinion of the dissenters, the conditions of social cooperation are not being honoured. For the principles of justice express precisely such conditions, and their persistent and deliberate violation in regard to basic liberties over any extended period of time cuts the ties of community and invites either submission or forceful resistance. By engaging in civil disobedience a minority leads the majority to consider whether it wants to have its acts taken in this way, or whether, in view of the common sense of justice, it wishes to acknowledge the claims of the minority.

Civil disobedience is also civil in another sense. Not only is it the outcome of a sincere conviction based on principles which regulate civic life, but it is public and non-violent, that is, it is done in a situation where arrest and punishment are expected and accepted without resistance. In this way it manifests a respect for legal procedures. Civil disobedience expresses disobedience to law within the limits of fidelity to law, and this feature of it helps to establish in the eyes of the majority that it is indeed conscientious and sincere, that it really is meant to address their sense of justice (Fried). Being completely open about one's acts and being willing to accept the legal consequences of one's conduct is a bond given to make good one's sincerity, for that one's deeds are conscientious is not easy to demonstrate to another or even before oneself. No doubt it is possible to imagine a legal system in which conscientious belief that the law is unjust is accepted as a defence for noncompliance, and men of great honesty who are confident in one another might make such a system work. But as things are such a scheme would be unstable; we must pay a price in order to establish that we believe our actions have a moral basis in the convictions of the community.

The non-violent nature of civil disobedience refers to the fact that it is intended to address the sense of justice of the majority and as such it is a form of speech, an expression of conviction. To engage in violent acts likely to injure and to hurt is

incompatible with civil disobedience as a mode of address. Indeed, an interference with the basic rights of others tends to obscure the civilly disobedient quality of one's act. Civil disobedience is non-violent in the further sense that the legal penalty for one's action is accepted and that resistance is not (at least for the moment) contemplated. Non-violence in this sense is to be distinguished from non-violence as a religious or pacifist principle. While those engaging in civil disobedience have often held some such principle, there is no necessary connection between it and civil disobedience. For on the interpretation suggested, civil disobedience in a democratic society is best understood as an appeal to the principles of justice, the fundamental conditions of willing social cooperation among free men, which in the view of the community as a whole are expressed in the constitution and guide its interpretation. Being an appeal to the moral basis of public life, civil disobedience is a political and not primarily a religious act. It addresses itself to the common principles of justice which men can require one another to follow and not to the aspirations of love which they cannot. Moreover, by taking part in civilly disobedient acts one does not foreswear indefinitely the idea of forceful resistance; for if the appeal against injustice is repeatedly denied, then the majority has declared its intention to invite submission or resistance and the latter may conceivably be justified even in a democratic regime. We are not required to acquiesce in the crushing of fundamental liberties by democratic majorities which have shown themselves blind to the principles of justice upon which justification of the constitution depends.

5. The Justification of Civil Disobedience

So far we have said nothing about the justification of civil disobedience, that is, the conditions under which civil disobedience may be engaged in consistent with the principles of justice that support a democratic regime. Our task is to see how the

characterization of civil disobedience as addressed to the sense of justice of the majority (or to the citizens as a body) determines when such action is justified.

First of all, we may suppose that the normal political appeals to the majority have already been made in good faith and have been rejected, and that the standard means of redress have been tried. Thus, for example, existing political parties are indifferent to the claims of the minority, and attempts to repeal the laws protested have been met with further repression since legal institutions are in the control of the majority. While civil disobedience should be recognized, I think, as a form of political action within the limits of fidelity to the rule of law, at the same time it is a rather desperate act just within these limits, and therefore it should, in general, be undertaken as a last resort when standard democratic processes have failed. In this sense it is not a normal political action. When it is justified there has been a serious breakdown; not only is there grave injustice in the law but a refusal more or less deliberate to correct it.

Second, since civil disobedience is a political act addressed to the sense of justice of the majority, it should usually be limited to substantial and clear violations of justice and preferably to those which, if rectified, will establish a basis for doing away with remaining injustices. For this reason there is a presumption in favour of restricting civil disobedience to violations of the first principle of justice, the principle of equal liberty, and to barriers which contravene the second principle, the principle of open offices which protects equality of opportunity. It is not, of course, always easy to tell whether these principles are satisfied. But if we think of them as guaranteeing the fundamental equal political and civil liberties (including freedom of conscience and liberty of thought) and equality of opportunity, then it is often relatively clear whether their principles are being honoured. After all, the equal liberties are defined by the visible structure of the social institutions; they are to be incorporated into the recognized practice, if not the letter, of social arrangements. When minorities are denied the right to vote or to hold certain political offices, when

certain religious groups are repressed and others denied equality of opportunity in the economy, this is often obvious and there is no doubt that justice is not being given. However, the first part of the second principle which requires that inequalities be to everyone's advantage is a much more imprecise and controversial matter. Not only is there a problem of assigning it a determinate and precise sense, but even if we do so and agree on what it should be, there is often a wide variety of reasonable opinion as to whether the principle is satisfied. The reason for this is that the principle applies primarily to fundamental economic and social policies. The choice of these depends upon theoretical and speculative beliefs as well as upon a wealth of concrete information, and all of this mixed with judgment and plain hunch, not to mention in actual cases prejudice and self-interest. Thus unless the laws of taxation are clearly designed to attack a basic equal liberty, they should not be protested by civil disobedience; the appeal to justice is not sufficiently clear and its resolution is best left to the political process. But violations of the equal liberties that define the common status of citizenship are another matter. The deliberate denial of these more or less over any extended period of time in the face of normal political protest is, in general, an appropriate object of civil disobedience. We may think of the social system as divided roughly into two parts, one which incorporates the fundamental equal liberties (including equality of opportunity) and another which embodies social and economic policies properly aimed at promoting the advantage of everyone. As a rule civil disobedience is best limited to the former where the appeal to justice is not only more definite and precise, but where, if it is effective, it tends to correct the injustices in the latter.

Third, civil disobedience should be restricted to those cases where the dissenter is willing to affirm that everyone else similarly subjected to the same degree of injustice has the right to protest in a similar way. That is, we must be prepared to authorize others to dissent in similar situations and in the same way, and to accept the consequences of their doing so. Thus, we may hold, for example, that the widespread disposition to disobey civilly

clear violations of fundamental liberties more or less deliberate over an extended period of time would raise the degree of justice throughout society and would ensure men's self-esteem as well as their respect for one another. Indeed, I believe this to be true, though certainly it is partly a matter of conjecture. As the contract doctrine emphasizes, since the principles of justice are principles which we would agree to in an original position of equality when we do not know our social position and the like, the refusal to grant justice is either the denial of the other as an equal (as one in regard to whom we are prepared to constrain our actions by principles which we would consent to) or the manifestation of a willingness to take advantage of natural contingencies and social fortune at his expense. In either case, injustice invites submission or resistance; but submission arouses the contempt of the oppressor and confirms him in his intention. If straightaway, after a decent period of time to make reasonable political appeals in the normal way, men were in general to dissent by civil disobedience from infractions of the fundamental equal liberties, these liberties would, I believe, be more rather than less secure. Legitimate civil disobedience properly exercised is a stabilizing device in a constitutional regime, tending to make it more firmly just. [. . .]

The final condition, of a different nature, is the following. We have been considering when one has a right to engage in civil disobedience, and our conclusion is that one has this right should three conditions hold: when one is subject to injustice more or less deliberate over an extended period of time in the face of normal political protests; where the injustice is a clear violation of the liberties of equal citizenship; and provided that the general disposition to protest similarly in similar cases would have acceptable consequences. These conditions are not, I think, exhaustive, but they seem to cover the more obvious points; yet even when they are satisfied and one has the right to engage in civil disobedience, there is still the different question of whether one should exercise this right, that is, whether by doing so one is likely to further one's ends. Having established one's right

to protest, one is then free to consider these tactical questions. We may be acting within our rights but still foolishly if our action only serves to provoke the harsh retaliation of the majority; and it is likely to do so if the majority lacks a sense of justice, or if the action is poorly timed or not well designed to make the appeal to the sense of justice effective. It is easy to think of instances of this sort, and in each case these practical questions have to be faced. From the standpoint of the theory of political obligation we can only say that the exercise of the right should be rational and reasonably designed to advance the protester's aims, and that weighing tactical questions presupposes that one has already established one's right, since tactical advantages in themselves do not support it.

6. Conclusion: Several Objections Considered

In a reasonably affluent democratic society justice becomes the first virtue of institutions. Social arrangements irrespective of their efficiency must be reformed if they are significantly unjust. No increase in efficiency in the form of greater advantages for many justifies the loss of liberty of a few. That we believe this is shown by the fact that in a democracy the fundamental liberties of citizenship are not understood as the outcome of political bargaining, nor are they subject to the calculus of social interests. Rather these liberties are fixed points which serve to limit political transactions and which determine the scope of calculations of social advantage. It is this fundamental place of the equal liberties which makes their systematic violation over any extended period of time a proper object of civil disobedience. For to deny men these rights is to infringe the conditions of social cooperation among free and rational persons, a fact which is evident to the citizens of a constitutional regime since it follows from the principles of justice which underlie their institutions. The justification of civil disobedience rests on the priority of justice and the equal liberties which it guarantees.

It is natural to object to this view of civil disobedience that it relies too heavily upon the existence of a sense of justice. Some may hold that the feeling for justice is not a vital political force, and that what moves men are various other interests, the desire for wealth, power, prestige, and so on. Now this is a large question the answer to which is highly conjectural, and each tends to have his own opinion. But there are two remarks which may clarify what I have said: first, I have assumed that there is in a constitutional regime a common sense of justice the principles of which are recognized to support the constitution and to guide its interpretation. In any given situation particular men may be tempted to violate these principles, but the collective force in their behalf is usually effective since they are seen as the necessary terms of cooperation among free men; and presumably the citizens of a democracy (or sufficiently many of them) want to see justice done. Where these assumptions fail, the justifying conditions for civil disobedience (the first three) are not affected, but the rationality of engaging in it certainly is. In this case, unless the costs of repressing civil dissent injures the economic self-interest (or whatever) of the majority, protest may simply make the position of the minority worse. No doubt as a tactical matter civil disobedience is more effective when its appeal coincides with other interests, but a constitutional regime is not viable in the long run without an attachment to the principles of justice of the sort which we have assumed.

Then, further, there may be a misapprehension about the manner in which a sense of justice manifests itself. There is a tendency to think that it is shown by professions of the relevant principles together with actions of an altruistic nature requiring a considerable degree of self-sacrifice. But these conditions are obviously too strong, for the majority's sense of justice may show itself simply in its being unable to undertake the measures required to suppress the minority and to punish as the law requires the various acts of civil disobedience. The sense of justice undermines the will to uphold unjust institutions, and so a majority despite its superior power may give way. It is unprepared to

force the minority to be subject to injustice. Thus, although the majority's action is reluctant and grudging, the role of the sense of justice is nevertheless essential, for without it the majority would have been willing to enforce the law and to defend its position. Once we see the sense of justice as working in this negative way to make established injustices indefensible, then it is recognized as a central element of democratic politics.

Finally, it may be objected against this account that it does not settle the question of who is to say when the situation is such as to justify civil disobedience. And because it does not answer this question, it invites anarchy by encouraging every man to decide the matter for himself. Now the reply to this is that each man must indeed settle this question for himself, although he may, of course, decide wrongly. This is true on any theory of political duty and obligation, at least on any theory compatible with the principles of a democratic constitution. The citizen is responsible for what he does. If we usually think that we should comply with the law, this is because our political principles normally lead to this conclusion. There is a presumption in favour of compliance in the absence of good reasons to the contrary. But because each man is responsible and must decide for himself as best he can whether the circumstances justify civil disobedience, it does not follow that he may decide as he pleases. It is not by looking to our personal interests or to political allegiances narrowly construed, that we should make up our mind. The citizen must decide on the basis of the principles of justice that underlie and guide the interpretation of the constitution and in the light of his sincere conviction as to how these

principles should be applied in the circumstances. If he concludes that conditions obtain which justify civil disobedience and conducts himself accordingly, he has acted conscientiously and perhaps mistakenly, but not in any case at his convenience.

In a democratic society each man must act as he thinks the principles of political right require him to. We are to follow our understanding of these principles, and we cannot do otherwise. There can be no morally binding legal interpretation of these principles, not even by a supreme court or legislature. Nor is there any infallible procedure for determining what or who is right. In our system the Supreme Court, Congress, and the President often put forward rival interpretations of the Constitution. Although the Court has the final say in settling any particular case, it is not immune from powerful political influence that may change its reading of the law of the land. The Court presents its point of view by reason and argument; its conception of the Constitution must, if it is to endure, persuade men of its soundness. The final court of appeal is not the Court, or Congress, or the President, but the electorate as a whole. The civilly disobedient appeal in effect to this body. There is no danger of anarchy as long as there is a sufficient working agreement in men's conceptions of political justice and what it requires. That men can achieve such an understanding when the essential political liberties are maintained is the assumption implicit in democratic institutions. There is no way to avoid entirely the risk of divisive strife. But if legitimate civil disobedience seems to threaten civil peace, the responsibility falls not so much on those who protest as upon those whose abuse of authority and power justifies such opposition.

Note

1. By the social contract theory I have in mind the doctrine found in Locke, Rousseau, and Kant.

Works Cited

Bedau, H.A. "On Civil Disobedience." *Journal of Philosophy* 58 (1961): 653–65.

Fried, C. "Moral Causation." *Harvard Law Review* (1964): 1258–70.

Study Questions

1. Explain the roles played by Rawls' hypothetical devices of the "original position" and "veil of ignorance" in his theory of justice.
2. On what grounds does Rawls argue that we have a duty to comply with unjust laws?
3. What two senses of "civil" are implicated in civil disobedience, according to Rawls, and in what two ways is civil disobedience non-violent?
4. Rawls tells us that simply following constitutionally proper procedures in enacting a law does not make it just but also that simply finding a law to be unjust does not sanction disobeying it. Provide an example of an act of civil disobedience. Does it satisfy the four conditions Rawls identifies as providing justification for civil disobedience?

► Further Reading

- Heidegger, Martin. "The Origin of the Work of Art." *Poetry, Language, Thought*. Trans. Albert Hofstadter. New York: Harper & Row, 1971.
- Heyes, Cressida J. *Self-Transformations: Foucault, Ethics, and Normalized Bodies*. Oxford and New York: Oxford University Press, 2007.
- Kernohan, Andrew. *A Guide for the Godless: The Secular Path to Meaning*. Halifax: self-published, 2008; <http://myweb.dal.ca/kernohan/godless>.
- Lopes, Dominic McIver. *Sight and Sensibility: Evaluating Pictures*. Oxford: Oxford University Press, 2005.

Plato

Plato (429–347 BCE) composed his works as dialogues, not treatises, thus drawing on the literary tradition of poets and playwrights. Nevertheless, in *The Republic*, Plato deems the arts a threat to the ideal city because they inflame people's passions and promote illusion instead of truth.

Art as Imitation

From *The Republic*

Characters: Socrates, Glaucon

1. "And truly," I [Socrates] said, "many other considerations assure me that we were entirely right [in Book 3] in our organization of the state, and especially, I think, in the matter of poetry." "What about it?" he [Glaucon] said. "In refusing to admit at all so much of it as is imitative; for that it is certainly not to be received is, I think, still more plainly apparent now that we have distinguished the several parts of the soul." "What do you mean?" "Why, between ourselves—for you will not betray me to the tragic poets and all other imitators—that kind of art seems to be a corruption of the mind of all listeners who do not possess as an antidote a knowledge of its real nature." "What is your idea in saying this?" he said. "I must speak out," I said, "though a certain love and reverence for Homer that has possessed me from a boy would stay me from speaking. For he appears to have been the first teacher and beginner of all these beauties of tragedy. Yet all the same we must not honour a man above truth, but, as I say, speak our minds."

"By all means," he said. "Listen, then, or rather, answer my question." "Ask it," he said. "Could you tell me in general what imitation is? For neither do I myself quite apprehend what it would be at." "It is likely, then," he said, "that I should apprehend!" "It would be nothing strange," said I, "since it often happens that the dimmer vision sees things in advance of the keener." "That is so," he said; "but in your presence I could not even be eager to try to state anything that appears to me, but do you yourself consider it." "Shall we, then, start the inquiry at this point by our customary procedure? We are in the habit, I take it, of positing a single idea or form in the case of the various multiplicities to which we give the same name. Do you not understand?" "I tipicity you please; for example, there are many couches and tables." "Of course." "But these utensils imply, I suppose, only two ideas or forms, one of a couch and one of a table." "Yes." "And are we not also in the habit of saying that the craftsman who produces

either of them fixes his eyes on the idea or form, and so makes in the one case the couches and in the other the tables that we use, and similarly of other things? For surely no craftsman makes the idea itself. How could he?" "By no means." "But now consider what name you would give to this craftsman." "What one?" "Him who makes all the things that all handicraftsmen severally produce." "A truly clever and wondrous man you tell of." "Ah, but wait, and you will say so indeed, for this same handicraftsman is not only able to make all implements, but he produces all plants and animals, including himself,¹ and thereto earth and heaven and the gods and all things in heaven and in Hades under the earth." "A most marvellous sophist," he said. "Are you incredulous?" said I. "Tell me, do you deny altogether the possibility of such a craftsman, or do you admit that in a sense there could be such a creator of all these things, and in another sense not? Or do you not perceive that you yourself would be able to make all these things in a way?" "And in what way, I ask you," he said. "There is no difficulty," said I, "but it is something that the craftsman can make everywhere and quickly. You could do it most quickly if you should choose to take a mirror and carry it about everywhere. You will speedily produce the sun and all the things in the sky, and speedily the earth and yourself and the other animals and implements and plants and all the objects of which we just now spoke." "Yes," he said, "the appearance of them, but not the reality and the truth." "Excellent," said I, "and you come to the aid of the argument opportunely. For I take it that the painter too belongs to this class of producers, does he not?" "Of course." "But you will say, I suppose, that his creations are not real and true. And yet, after a fashion, the painter too makes a couch, does he not?" "Yes," he said, "the appearance of one, he too."

2. "What of the cabinet-maker? Were you not just now saying that he does not make the idea or form which we say is the real couch, the couch in itself, but only some particular couch?" "Yes, I was." "Then if he does not make that which really is, he could not be said to make real being but something that resembles real being but is not that. But if anyone should say that being in the complete sense belongs to the work of the cabinet-maker or to that of any other handicraftsman,

it seems that he would say what is not true." "That would be the view," he said, "of those who are versed in this kind of reasoning."² "We must not be surprised, then, if this too is only a dim adumbration in comparison with reality." "No, we must not." "Shall we, then, use these very examples in our quest for the true nature of this imitator?" "If you please," he said. "We get, then, these three couches, one, that in nature which, I take it, we would say that God produces, or who else?" "No one, I think." "And then there was one which the carpenter made." "Yes," he said. "And one which the painter. Is not that so?" "So be it." "The painter, then, the cabinet-maker, and God, there are these three presiding over three kinds of couches." "Yes, three." "Now God, whether because he so willed or because some compulsion was laid upon him not to make more than one couch in nature, so wrought and created one only, the couch which really and in itself is. But two or more such were never created by God and never will come into being." "How so?" he said. "Because," said I, "if he should make only two, there would again appear one of which they both would possess the form or idea, and that would be the couch that really is in and of itself, and not the other two."³ "Right," he said. "God, then, I take it, knowing this and wishing to be the real author of the couch that has real being and not of some particular couch, nor yet a particular cabinet-maker, produced it in nature unique." "So it seems." "Shall we, then, call him its true and natural begetter, or something of the kind?" "That would certainly be right," he said, "Since it is by and in nature that he has made this and all other things." "And what of the carpenter? Shall we not call him the creator of a couch?" "Yes." "Shall we also say that the painter is the creator and maker of that sort of thing?" "By no means." "What will you say he is in relation to the couch?" "This," said he, "Seems to me the most reasonable designation for him, that he is the imitator of the thing which those others produce." "Very good," said I; "the producer of the product three removes from nature you call the imitator?" "By all means," he said. "This, then, will apply to the maker of tragedies also, if he is an imitator and is in his nature three removes from the king and the truth, as are all other imitators." "It would seem so." "We

are in agreement, then, about the imitator. But tell me now this about the painter. Do you think that what he tries to imitate is in each case that thing itself in nature or the works of the craftsmen?" "The works of the craftsmen," he said. "Is it the reality of them or the appearance? Define that further point." "What do you mean?" he said. "This: Does a couch differ from itself according as you view it from the side or the front or in any other way? Or does it differ not at all in fact though it appears different, and so of other things?" "That is the way of it," he said: "it appears other but differs not at all." "Consider, then, this very point. To which is painting directed in every case, to the imitation of reality as it is or of appearance as it appears? Is it an imitation of a phantasm or of the truth?" "Of a phantasm," he said. "Then the mimetic art is far removed from truth, and this, it seems, is the reason why it can produce everything, because it touches or lays hold of only a small part of the object and that a phantom; as, for example, a painter, we say, will paint us a cobbler, a carpenter, and other craftsmen, though he himself has no expertness in any of these arts, but nevertheless if he were a good painter, by exhibiting at a distance his picture of a carpenter he would deceive children and foolish men, and make them believe it to be a real carpenter." "Why not?" "But for all that, my friend, this, I take it, is what we ought to bear in mind in all such cases: When anyone reports to us of someone, that he has met a man who knows all the crafts and everything else that men severally know, and that there is nothing that he does not know more exactly than anybody else, our tacit rejoinder must be that he is a simple fellow, who apparently has met some magician or sleight-of-hand man and imitator and has been deceived by him into the belief that he is all-wise, because of his own inability to put to the proof and distinguish knowledge, ignorance, and imitation." "Most true," he said.

3. "Then," said I, "have we not next to scrutinize tragedy and its leader Homer, since some people tell us that these poets know all the arts and all things human pertaining to virtue and vice, and all things divine? For the good poet, if he is to poetize things rightly, must, they argue, create with knowledge or else be unable to create. So we must consider whether these critics have not fallen in with such

imitators and been deceived by them, so that looking upon their works they cannot perceive that these are three removes from reality, and easy to produce without knowledge of the truth. For it is phantoms, not realities, that they produce. Or is there something in their claim, and do good poets really know the things about which the multitude fancy they speak well?" "We certainly must examine the matter," he said. "Do you suppose, then, that if a man were able to produce both the exemplar and the semblance, he would be eager to abandon himself to the fashioning of phantoms and set this in the forefront of his life as the best thing he had?" "I do not." "But, I take it, if he had genuine knowledge of the things he imitates he would far rather devote himself to real things than to the imitation of them, and would endeavour to leave after him many noble deeds and works as memorials of himself, and would be more eager to be the theme of praise than the praiser." "I think so," he said; "for there is no parity in the honour and the gain." "Let us not, then, demand a reckoning from Homer or any other of the poets on other matters by asking them, if any one of them was a physician and not merely an imitator of a physician's talk, what men any poet, old or new, is reported to have restored to health as **Asclepius** did, or what disciples of the medical art he left after him as Asclepius did his descendants; and let us dismiss the other arts and not question them about them; but concerning the greatest and finest things of which Homer undertakes to speak, wars and generalship and the administration of cities and the education of men, it surely is fair to question him and ask, "Friend Homer, if you are not at the third remove from truth and reality in human excellence, being merely that creator of phantoms whom we defined as the imitator, but if you are even in the second place and were capable of knowing what pursuits make men better or worse in private or public life, tell us what city was better governed owing to you, even as Lacedaemon [Spartal was because of **Lycurgus**, and many other cities great and small because of other legislators. But what city credits you with having been a good legislator and having benefited them? Italy and Sicily say this of **Charondas** and we of **Solon**. But who says it of you?" Will he be able to name any?" "I think

not," said Glaucon; "at any rate none is mentioned even by the Homerids themselves." "Well, then, is there any tradition of a war in Homer's time that was well conducted by his command or counsel?" "None." "Well, then, as might be expected of a man wise in practical affairs, are many and ingenious inventions for the arts and business of life reported of Homer as they are of **Thales the Milesian** and **Anacharsis the Scythian?**" "Nothing whatever of the sort." "Well, then, if no public service is credited to him, is Homer reported while he lived to have been a guide in education to men who took pleasure in associating with him and transmitted to posterity a certain Homeric way of life just as **Pythagoras** was himself especially honoured for this, and his successors, even to this day, denominating a certain way of life the Pythagorean, are distinguished among their contemporaries?" "No, nothing of this sort either is reported; for **Creophylos**⁴, Socrates, the friend of Homer, would perhaps be even more ridiculous than his name as a representative of Homeric culture and education, if what is said about Homer is true. For the tradition is that Homer was completely neglected in his own lifetime by that friend of the flesh."

4. "Why, yes, that is the tradition," said I; "but do you suppose, Glaucon, that, if Homer had really been able to educate men and make them better and had possessed not the art of imitation but real knowledge, he would not have acquired many companions and been honoured and loved by them? But are we to believe that while **Protagoras of Abdera** and **Prodicus of Ceos** and many others are able by private teaching to impress upon their contemporaries the conviction that they will not be capable of governing their homes or the city unless they put them in charge of their education, and make themselves so beloved for this wisdom that their companions all but carry them about on their shoulders,

yet, forsooth, that Homer's contemporaries, if he had been able to help men to achieve excellence, would have suffered him or **Hesiod** to roam about rhapsodizing and would not have clung to them far rather than to their gold, and constrained them to dwell with them in their homes, or failing to persuade them, would themselves have escorted them whereversoever they went until they should have sufficiently imbibed their culture?" "What you say seems to me to be altogether true, Socrates," he said. "Shall we, then, lay it down that all the poetic tribe, beginning with Homer, are imitators of images of excellence and of the other things that they 'create,' and do not lay hold on truth? But, as we were just now saying, the painter will fashion, himself knowing nothing of the cobbler's art, what appears to be a cobbler to him and likewise to those who know nothing but judge only by forms and colours?" "Certainly." And similarly, I suppose, we shall say that the poet himself, knowing nothing but how to imitate, lays on with words and phrases the colours of the several arts in such fashion that others equally ignorant, who see things only through words, will deem his words most excellent, whether he speak in rhythm, metre and harmony about cobbling or generalship or anything whatever. So mighty is the spell that these adornments naturally exercise; though when they are stripped bare of their musical colouring and taken by themselves, I think you know what sort of a showing these sayings of the poets make. For you, I believe, have observed them." "I have," he said. "Do they not," said I, "resemble the faces of adolescents, young but not really beautiful, when the bloom of youth abandons them?" "By all means," he said. "Come, then," said I, "consider this point: The creator of the phantom, the imitator, we say, knows nothing of the reality but only the appearance. Is not that so?" "Yes."

Notes

1. Tr.: It is a tempting error to refer this to God, as I once did. [...] But the producer of everything, including himself, is the imitator generalized and then exemplified by the painter and the poet.

2. Tr.: An indirect reference to Plato and his school.
3. Tr.: The famous argument of the third man.
4. Tr.: "Of the beef-clan." The scholiast says he was a Chian and an epic poet.

Study Questions

1. Socrates tells us that the god, the carpenter, and the painter make three different kinds of couches (or beds in other translations). What are these? Why does Socrates characterize the painter's couch as "three removes from nature"?
2. Evaluate Socrates' argument that great works of literature cannot provide us with knowledge.

Plato

Arthur Danto

Arthur C. Danto is Johnsonian Professor Emeritus of Philosophy at Columbia University. Danto received a BA from Wayne State University, where he studied art and history, and an MA and PhD from Columbia. Danto's research interests include philosophy of art and philosophical psychology. Art critic at *The Nation* since 1984, Danto's *Encounters and Reflections: Art in the Historical Present* won the National Book Critics Circle Award in 1990. This influential 1964 essay critiques Plato's theory of art as imitation and lays the groundwork for subsequent institutional theories of art.

The Artworld

Hamlet: Do you see nothing there?

The Queen: Nothing at all; yet all that is I see.

—Shakespeare: *Hamlet*, Act III, Scene IV

1. Introduction

Hamlet and Socrates, though in praise and depreciation respectively, spoke of art as a mirror held up

to nature. As with many disagreements in attitude, this one has a factual basis. Socrates saw mirrors as but reflecting what we can already see; so art, insofar as mirrorlike, yields idle accurate duplications of the appearances of things, and is of no cognitive benefit whatever. Hamlet, more acutely, recognized a remarkable feature of reflecting surfaces, namely that they show us what we could not otherwise

perceive—our own face and form—and so art, insofar as it is mirrorlike, reveals us to ourselves, and is, even by Socratic criteria, of some cognitive utility after all. As a philosopher, however, I find Socrates' discussion defective on other, perhaps less profound grounds than these. If a mirror-image of *o* is indeed an imitation of *o*, then, if art is imitation, mirror-images are art. But in fact mirroring objects no more is art than returning weapons to a madman is justice; and reference to mirrorings would be just the sly sort of counterinstance we would expect Socrates to bring forward in rebuttal of the theory he instead uses them to illustrate. If that theory requires us to class *these* as art, it thereby shows its inadequacy: "is an imitation" will not do as a sufficient condition for "is art." Yet, perhaps because artists were engaged in imitation, in Socrates' time and after, the insufficiency of the theory was not noticed until the invention of photography. Once rejected as a sufficient condition, mimesis was quickly discarded as even a necessary one; and since the achievement of Kandinsky, mimetic features have been relegated to the periphery of critical concern, so much so that some works survive in spite of possessing those virtues, excellence in which was once celebrated as the essence of art, narrowly escaping demotion to mere illustrations.

It is, of course, indispensable in Socratic discussion that all participants be masters of the concept up for analysis, since the aim is to match a real defining expression to a term in active use, and the test for adequacy presumably consists in showing that the former analyzes and applies to all and only those things of which the latter is true. The popular disclaimer notwithstanding, then, Socrates' auditors purportedly knew what art was as well as what they liked; and a theory of art, regarded here as a real definition of "Art," is accordingly not to be of great use in helping men to recognize instances of its application. Their antecedent ability to do this is precisely what the adequacy of the theory is to be tested against, the problem being only to make explicit what they already know. It is our use of the term that the theory allegedly means to capture, but we are supposed able, in the words of a recent writer, "to separate

those objects which are works of art from those which are not, because . . . we know how correctly to use the word 'art' and to apply the phrase 'work of art.'" Theories, on this account, are somewhat like mirror-images on Socrates' account, showing forth what we already know, wordy reflections of the actual linguistic practice we are masters in.

But telling artworks from other things is not so simple a matter, even for native speakers, and these days one might not be aware he was on artistic terrain without an artistic theory to tell him so. And part of the reason for this lies in the fact that terrain is constituted artistic in virtue of artistic theories, so that one use of theories, in addition to helping us discriminate art from the rest, consists in making art possible. Glaucon and the others could hardly have known what was art and what not: otherwise they would never have been taken in by mirror-images.

2. Reality Theory vs. Imitation Theory of Art

Suppose one thinks of the discovery of a whole new class of artworks as something analogous to the discovery of a whole new class of facts anywhere, viz., as something for theoreticians to explain. In science, as elsewhere, we often accommodate new facts to old theories via auxiliary hypotheses, a pardonable enough conservatism when the theory in question is deemed too valuable to be jettisoned all at once. Now the Imitation Theory of Art (IT) is, if one but thinks it through, an exceedingly powerful theory, explaining a great many phenomena connected with the causation and evaluation of artworks, bringing a surprising unity into a complex domain. Moreover, it is a simple matter to shore it up against many purported counterinstances by such auxiliary hypotheses as that the artist who deviates from mimeticity is perverse, inept, or mad. Ineptitude, chicanery, or folly are, in fact, testable predication. Suppose, then, tests reveal that these hypotheses fail to hold, that the theory, now beyond repair, must be replaced. And a new theory is worked out, capturing what it can of the old theory's competence, together with the heretofore

recalcitrant facts. One might, thinking along these lines, represent certain episodes in the history of art as not dissimilar to certain episodes in the history of science, where a conceptual revolution is being effected and where refusal to countenance certain facts, while in part due to prejudice, inertia, and self-interest, is due also to the fact that a well-established, or at least widely credited, theory is being threatened in such a way that all coherence goes.

Some such episode transpired with the advent of post-impressionist paintings. In terms of the prevailing artistic theory (IT), it was impossible to accept these as art unless inept art: otherwise they could be discounted as hoaxes, self-advertisements, or the visual counterparts of madmen's ravings. So to get them accepted as art, on a footing with the *Transfiguration* (not to speak of a *Landseer stag*), required not so much a revolution in taste as a theoretical revision of rather considerable proportions, involving not only the artistic enfranchisement of these objects, but an emphasis upon newly significant features of accepted artworks, so that quite different accounts of their status as artworks would now have to be given. As a result of the new theory's acceptance, not only were post-impressionist paintings taken up as art, but numbers of objects (masks, weapons, etc.) were transferred from anthropological museums (and heterogeneous other places) to *musées des beaux arts*, though, as we would expect from the fact that a criterion for the acceptance of a new theory is that it account for whatever the older one did, nothing had to be transferred out of the *musée des beaux arts*—even if there were internal rearrangements as between storage rooms and exhibition space. Countless native speakers hung upon suburban mantelpieces innumerable replicas of paradigm cases for teaching the expression "work of art" that would have sent their Edwardian forebears into linguistic apoplexy.

To be sure, I distort by speaking of a theory: historically, there were several, all, interestingly enough, more or less defined in terms of the IT. Art-historical complexities must yield before the exigencies of logical exposition, and I shall speak as though there were one replacing theory, partially compensating

for historical falsity by choosing one which was actually enunciated. According to it, the artists in question were to be understood not as unsuccessfully imitating real forms but as successfully creating new ones, quite as real as the forms which the older art had been thought, in its best examples, to be creditably imitating. Art, after all, had long since been thought of as creative (*Vasari* says that God was the first artist), and the post-impressionists were to be explained as genuinely creative, aiming, in Roger Fry's words, "not at illusion but reality." This theory (RT) furnished a whole new mode of looking at painting, old and new. Indeed, one might almost interpret the crude drawing in *Van Gogh* and *Cézanne*, the dislocation of form from contour in *Rouault* and *Dufy*, the arbitrary use of colour planes in *Gauguin* and the *Fauves*, as so many ways of drawing attention to the fact that these were *non-imitations*, specifically intended not to deceive. Logically, this would be roughly like printing "Not Legal Tender" across a brilliantly counterfeited dollar bill, the resulting object (counterfeit *cum* inscription) rendered incapable of deceiving anyone. It is not an illusory dollar bill, but then, just because it is non-illusory it does not automatically become a real dollar bill either. It rather occupies a freshly opened area between real objects and real facsimiles of real objects: it is a non-facsimile, if one requires a word, and a new contribution to the world. Thus, Van Gogh's *Potato Eaters*, as a consequence of certain unmistakable distortions, turns out to be a non-facsimile of real-life potato eaters; and inasmuch as these are not facsimiles of potato eaters, Van Gogh's picture, as a non-imitation, had as much right to be called a real object as did its putative subjects. By means of this theory (RT), artworks re-entered the thick of things from which Socratic theory (IT) had sought to evict them: if no *more* real than what carpenters wrought, they were at least no *less* real. The Post-Impressionist won a victory in ontology.

It is in terms of RT that we must understand the artworks around us today. Thus Roy Lichtenstein paints comic-strip panels, though ten or twelve feet high. These are reasonably faithful projections onto a gigantesque scale of the homely frames from the

daily tabloid, but it is precisely the scale that counts.

A skilled engraver might incise *The Virgin and the Chancellor Rollin* on a pinhead, and it would be recognizable as such to the keen of sight, but an engraving of a Barnett Newman on a similar scale would be a blob, disappearing in the reduction. A photograph of a Lichtenstein is indiscernible from a photograph of a counterpart panel from *Steve Canyon*, but the photograph fails to capture the scale, and hence is inaccurate a reproduction as a black-and-white engraving of Botticelli, scale being essential here as colour there. Lichtensteins, then, are not imitations but new entities, as giant whelks [marine snails] would be. Jasper Johns, by contrast, paints objects with respect to which questions of scale are irrelevant. Yet his objects cannot be imitations, for they have the remarkable property that any intended copy of a member of this class of objects is automatically a member of the class itself, so that these objects are logically inimitable. Thus, a copy of a numeral just is that numeral: a painting of 3 is a 3 made of paint. Johns, in addition, paints targets, flags, and maps. Finally, in what I hope are not unwitting footnotes to Plato, two of our pioneers—Robert Rauschenberg and Claes Oldenburg—have made genuine beds.

Rauschenberg's bed hangs on a wall, and is streaked with some desultory housepaint. Oldenburg's bed is a rhomboid, narrower at one end than the other, with what one might speak of as a built-in perspective ideal for small bedrooms. As beds, these sell at singularly inflated prices, but one could sleep in either of them: Rauschenberg has expressed the fear that someone might just climb into his bed and fall asleep. Imagine, now, a certain Testadura ["hard head"]—a plain speaker and noted philistine—who is not aware that these are art, and who takes them to be really simple and pure. He attributes the paintstreaks on Rauschenberg's bed to the slovenliness of the owner, and the bias in the Oldenburg bed to the ineptitude of the builder or the whimsy, perhaps, of whoever had it "custom-made." These would be mistakes, but mistakes of rather an odd kind, and not terribly different from that made by the stunned birds who pecked the sham grapes of *Zeuxis*. They mistook art for reality, and so has Testadura. But it was meant to

be reality, according to RT. Can one have mistaken reality for reality? How shall we describe Testadura from being a mishapen bed? This is equivalent to asking what makes it art, and with this query we enter a domain of conceptual inquiry where native speakers are poor guides: they are lost themselves.

3. The Is of Artistic Identification

To mistake an artwork for a real object is no great feat when an artwork is the real object one mistakes it for. The problem is how to avoid such errors, or to remove them once they are made. The artwork is a bed, and not a bed-illusion; so there is nothing like the traumatic encounter against a flat surface that brought it home to the birds of *Zeuxis* that they had been duped. Except for the guard cautioning Testadura not to sleep on the artworks, he might never have discovered that this was an artwork and not a bed; and since, after all, one cannot discover that a bed is not a bed, how is Testadura to realize that he has made an error? A certain sort of explanation is required, for the error here is a curiously philosophical one, rather like, if we may assume as correct some well-known views of PF Strawson, mistaking a person for a material body when the truth is that a person is a material body in the sense that a whole class of predicates, sensibly applicable to material bodies, are sensibly, and by appeal to no different criteria, applicable to persons. So you cannot discover that a person is not a material body.

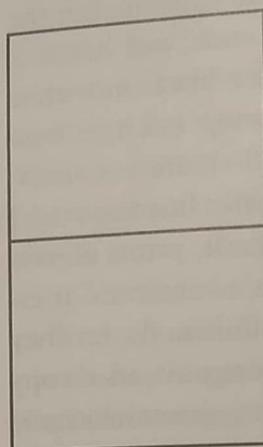
We begin by explaining, perhaps, that the paintstreaks are not to be explained away, that they are part of the object, so the object is not a mere bed with—as it happens—streaks of paint spilled over it, but a complex object fabricated out of a bed and some paintstreaks: a paint-bed. Similarly, a person is not a material body with—as it happens—some thoughts superadded, but is a complex entity made of the body and some conscious states, a conscious body. Persons, like artworks, must then be taken as up of a body and some thoughts, perhaps, and are in that sense primitive. Or, more accurately, the paintstreaks are not part of the real object—the bed—which are not part of the real part of the real object.

happens to be part of the artwork, but are, like the bed, part of the artwork as such. And this might be generalized into a rough characterization of artworks that happen to contain real objects as parts of themselves: not every part of an artwork *A* is part of a real object *R* when *R* is part of *A* and can, moreover, be detached from *A* and seen merely as *R*. The mistake thus far will have been to mistake *A* for part of itself, namely *R*, even though it would not be incorrect to say that *A* is *R*, that the artwork is a bed. It is the "is" which requires clarification here.

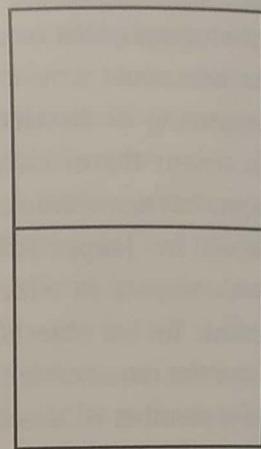
There is an *is* that figures prominently in statements concerning artworks which is not the *is* of either identity or predication; nor is it the *is* of existence, of identification, or some special *is* made up to serve a philosophic end. Nevertheless, it is in common usage, and is readily mastered by children. It is the sense of *is* in accordance with which a child, shown a circle and a triangle and asked which is him and which his sister, will point to the triangle saying "That is me"; or, in response to my question, the person next to me points to the man in purple and says "That one is Lear"; or in the gallery I point, for my companion's benefit, to a spot in the painting before us and say "That white dab is Icarus." We do not mean, in these instances, that whatever is pointed to stands for, or represents, what it is said to be, for the word "Icarus" stands for or represents Icarus: yet I would not in the same sense of *is* point to the word and say "That is Icarus." The sentence "That *a* is *b*" is perfectly compatible with "That *a* is not *b*" when the first employs this sense of *is* and the second employs some other, though *a* and *b* are used nonambiguously throughout. Often, indeed, the truth of the first requires the truth of the second. The first, in fact, is incompatible with "That *a* is not *b*" only when the *is* is used nonambiguously throughout. For want of a word I shall designate this the *is* of artistic identification; in each case in which it is used, the *a* stands for some specific physical property of, or physical part of, an object; and, finally, it is a necessary condition for something to be an artwork that some part or property of it be designable by the subject of a sentence that employs this special *is*. It is an *is*, incidentally, which has near-relatives in

marginal and mythical pronouncements. (Thus, one is Quetzalcoatl; those are the Pillars of Hercules.)

Let me illustrate. Two painters are asked to decorate the east and west walls of a science library with frescoes to be respectively called Newton's First Law and Newton's Third Law. These paintings, when finally unveiled, look, scale apart, as follows:



A



B

As objects I shall suppose the works to be indiscernible: a black, horizontal line on a white ground, equally large in each dimension and element. *B* explains his work as follows: a mass, pressing downward, is met by a mass pressing upward: the lower mass reacts equally and oppositely to the upper one. *A* explains his work as follows: the line through the space is the path of an isolated particle. The path goes from edge to edge, to give the sense of its going beyond. If it ended or began within the space, the line would be curved: and it is parallel to the top and bottom edges, for if it were closer to one than to another, there would have to be a force accounting for it, and this is inconsistent with its being the path of an isolated particle.

Much follows from these artistic identifications. To regard the middle line as an edge (mass meeting mass) imposes the need to identify the top and bottom half of the picture as rectangles, and as two distinct parts (not necessarily as two masses, for the line could be the edge of one mass jutting up—or down—into empty space). If it is an edge, we cannot thus take the entire area of the painting as a single space: it is rather composed of

two forms, or one form and a non-form. We could take the entire area as a single space only by taking the middle horizontal as a *line* which is not an edge. But this almost requires a three-dimensional identification of the whole picture: the area can be a flat surface which the line is *above* (*Jet-flight*), or *below* (*Submarine-path*), or *on* (*Line*), or *in* (*Fissure*), or *through* (*Newton's First Law*)—though in this last case the area is not a flat surface but a transparent cross section of absolute space. We could make all these prepositional qualifications clear by imagining perpendicular cross sections to the picture plane. Then, depending upon the applicable prepositional clause, the area is (artistically) interrupted or not by the horizontal element. If we take the line as *through* space, the edges of the picture are not really the edges of the space: the space goes beyond the picture if the line itself does; and we are in the same space as the line is. As *B*, the edges of the picture can be *part* of the picture in case the masses go right to the edges, so that the edges of the picture are *their* edges. In that case, the vertices of the picture would be the vertices of the masses, except that the masses have four vertices more than the picture itself does: here four vertices would be part of the art work which were not part of the real object. Again, the faces of the masses could be the face of the picture, and in looking at the picture, we are looking at these faces: but *space* has no face, and on the reading of *A* the work has to be read as faceless, and the face of the physical object would not be part of the artwork. Notice here how one artistic identification engenders another artistic identification, and how, consistently with a given identification, we are required to give others and precluded from still others: indeed, a given identification determines how many elements the work is to contain. These different identifications are incompatible with one another, or generally so, and each might be said to make a different artwork, even though each artwork contains the identical real object as part of itself—or at least parts of the identical real object as parts of itself. There are, of course, senseless identifications: no one could, I think, sensibly

read the middle horizontal as *Love's Labour's Lost* or *The Ascendancy of St Erasmus*. Finally, notice how acceptance of one identification rather than another is in effect to exchange one world for another. We could, indeed, enter a quiet poetic world by identifying the upper area with a clear and cloudless sky, reflected in the still surface of the water below, whiteness kept from whiteness only by the unreal boundary of the horizon.

And now Testadura, having hovered in the wings throughout this discussion, protests that *all he sees is paint*: a white painted oblong with a black line painted across it. And how right he really is: that is all he sees or that anybody can, we aesthetes included. So, if he asks us to show him what there is further to see, to demonstrate through pointing that this is an artwork (*Sea and Sky*), we cannot comply, for he has overlooked nothing (and it would be absurd to suppose he had, that there was something tiny we could point to and he, peering closely, say "So it is! A work of art after all!"). We cannot help him until he has mastered the *is of artistic identification* and so constitutes it a work of art. If he cannot achieve this, he will never look upon artworks: he will be like a child who sees sticks as sticks.

But what about pure abstractions, say something that looks just like *A* but is entitled No. 7? The 10th Street abstractionist blankly insists that there is nothing here but white paint and black, and none of our literary identifications need apply. What then distinguishes him from Testadura, whose philistine utterances are indiscernible from his? And how can it be an artwork for him and not for Testadura, when they agree that there is nothing that does not meet the eye? The answer, unpopular as it is likely to be to purists of every variety, lies in the fact that this artist has returned to the physicality of paint through an atmosphere compounded of artistic theories and the history of recent and remote painting, elements of which he is trying to refine out of his own work; and as a consequence of this his work belongs in this atmosphere and is part of this history. He has achieved abstraction through rejection of artistic identifications, returning to the real world from which such identifications remove us

(he thinks), somewhat in the mode of Ch'ing Yuan, who wrote:

Before I had studied Zen for thirty years, I saw mountains as mountains and waters as waters. When I arrived at a more intimate knowledge, I came to the point where I saw that mountains are not mountains, and waters are not waters. But now that I have got the very substance I am at rest. For it is just that I see mountains once again as mountains, and waters once again as waters.

His identification of what he has made is logically dependent upon the theories and history he rejects. The difference between his utterance and Testadura's "This is black paint and white paint and nothing more" lies in the fact that he is still using the *is* of artistic identification, so that his use of "That black paint is black paint" is not a tautology. Testadura is not at that stage. To see something as art requires something the eye cannot decry—an atmosphere of artistic theory, a knowledge of the history of art: an artworld.

4. Brillo Cartons as Artwork

Mr Andy Warhol, the Pop artist, displays facsimiles of Brillo cartons, piled high, in neat stacks, as in the stockroom of the supermarket. They happen to be of wood, painted to look like cardboard, and why not? To paraphrase the critic of the *Times*, if one may make the facsimile of a human being out of bronze, why not the facsimile of a Brillo carton out of plywood? The cost of these boxes happens to be 2×10^3 that of their homely counterparts in real life—a differential hardly ascribable to their advantage in durability. In fact the Brillo people might, at some slight increase in cost, make their boxes out of plywood without these becoming artworks, and Warhol might make *his* out of cardboard without their ceasing to be art. So we may forget questions of intrinsic value, and ask why the Brillo people cannot manufacture art and why Warhol cannot *but* make artworks. Well, his are made by hand, to be sure. Which is like an insane reversal of Picasso's strategy in pasting the

label from a bottle of Suze [French brand of bitters] onto a drawing, saying as it were that the academic artist, concerned with exact imitation, must always fall short of the real thing: so why not just use the real thing? The Pop artist laboriously reproduces machine-made objects by hand, e.g., painting the labels on coffee cans (one can hear the familiar commendation "Entirely made by hand" falling painfully out of the guide's vocabulary when confronted by these objects). But the difference cannot consist in craft: a man who carved pebbles out of stones and carefully constructed a work called *Gravel Pile* might invoke the labour theory of value to account for the price he demands; but the question is, What makes it art? And why need Warhol *make* these things anyway? Why not just scrawl his signature across one? Or crush one up and display it as *Crushed Brillo Box* ("A protest against mechanization . . .") or simply display a Brillo carton as *Uncrushed Brillo Box* ("A bold affirmation of the plastic authenticity of industrial . . .")? Is this man a kind of *Midas*, turning whatever he touches into the gold of pure art? And the whole world consisting of latent artworks waiting, like the bread and wine of reality, to be transfigured, through some dark mystery, into the indiscernible flesh and blood of the sacrament? Never mind that the Brillo box may not be good, much less great art. The impressive thing is that it is art at all. But if it is, why are not the indiscernible Brillo boxes that are in the stockroom? Or *has* the whole distinction between art and reality broken down?

Suppose a man collects objects (ready-mades), including a Brillo carton; we praise the exhibit for variety, ingenuity, what you will. Next he exhibits nothing but Brillo cartons, and we criticize it as dull, repetitive, self-plagiarizing—or (more profoundly) claim that he is obsessed by regularity and repetition, as in *Marienbad*. Or he piles them high, leaving a narrow path; we tread our way through the smooth opaque stacks and find it an unsettling experience, and write it up as the closing in of consumer products, confining us as prisoners: or we say he is a modern pyramid builder. True, we don't say these things about the stockboy. But then a stockroom is not an art gallery, and we cannot readily separate

the Brillo cartons from the gallery they are in, any more than we can separate the Rauschenberg bed from the paint upon it. Outside the gallery, they are pasteboard cartons. But then, scoured clean of paint, Rauschenberg's bed is a bed, just what it was before it was transformed into art. But then if we think this matter through, we discover that the artist has failed, really and of necessity, to produce a mere real object. He has produced an artwork, his use of real Brillo cartons being but an expansion of the resources available to artists, a contribution to *artists' materials*, as oil paint was, or *tuche*.

What in the end makes the difference between a Brillo box and a work of art consisting of a Brillo Box is a certain theory of art. It is the theory that takes it up into the world of art, and keeps it from collapsing into the real object which it is (in a sense of is other

than that of artistic identification). Of course, without the theory, one is unlikely to see it as art, and in order to see it as part of the artworld, one must have mastered a good deal of artistic theory as well as a considerable amount of the history of recent New York painting. It could not have been art fifty years ago. But then there could not have been, everything being equal, flight insurance in the Middle Ages, or **Etruscan** typewriter erasers. The world has to be ready for certain things, the artworld no less than the real one. It is the role of artistic theories, these days as always, to make the artworld, and art, possible. It would, I should think, never have occurred to the **painters of Lascaux** that they were producing *art* on those walls. Not unless there were neolithic aestheticians.

Study Questions

1. What is the "Imitation Theory of Art" or "IT"? Why, according to Danto, does it provide neither necessary nor sufficient conditions for defining art?
2. What is the "Real Theory of Art" or "RT"? Recalling Plato's distinction between the carpenter's and painter's couches, explain how the artworks of Rauschenberg and Oldenburg support RT over IT.
3. How does Danto use the distinction between "person" and "material body" to explain the "is of artistic identification" that constitutes Rauschenberg's bed as an artwork? How does Danto use the Newton frescoes to illustrate how the *is of artistic identification* constitutes identical paintings as different artworks?
4. Describe how Danto uses the concept "the artworld" to explain why a pure abstraction is an artwork even though it represents nothing and a Brillo box is an artwork in an art gallery but not a stock room.

Alfred Lessing

Alfred (Fred) Lessing was born in 1936, into a Jewish family in The Hague, Netherlands. Lessing survived the Holocaust as a "hidden child," spending most of the war separated from his family and forced to conceal his identity. After the war, the Lessings immigrated to the United States. Lessing has a BA from Carleton College, in Minnesota, a PhD from Yale University, and was a philosophy professor at Oakland University in Rochester, Michigan, when he published this influential article in 1965. Lessing's response to the question he poses in the article's title is that forgeries lack the originality that is expected for artistic works. Today, Lessing is a clinical psychologist in private practice in Birmingham, Michigan, and active as a lecturer about the Holocaust.

What Is Wrong with a Forgery?

1. *The Disciples and van Meegeren's Dilemma*

This paper attempts to answer the simple question: What is wrong with a forgery? It assumes, then, that something is wrong with a forgery. This is not an unreasonable assumption when one considers that the term *forgery* can be defined only in reference to a contrasting phenomenon which must somehow include the notion of genuineness or authenticity. When thus defined there can be little doubt that the concept of forgery is a normative one. It is clear, moreover, that it is a negative concept implying the absence or negation of value. The problem arises when we ask what kind of value we are speaking of. It appears to be generally assumed that in the case of artistic forgeries we are dealing with the absence or negation of aesthetic value. If this were so, a forgery would be an aesthetically inferior work of art. This, as I will try to show, is not the case. Pure aesthetics cannot explain forgery. [. . .]

Somehow critics have never understood this and have again and again allowed themselves to be forced into an embarrassing position upon the discovery of some forgery or other. Perhaps the classic, certainly the most celebrated, case in point was that of **Hans van Meegeren** who in 1945 disturbed the complacent tranquility of the world of art and art critics by confessing that he was the artist responsible for eight paintings of which six had been sold as legitimate

Vermeers and two as *de Hooghs*. It is not hard to imagine the discomfort felt by critics at that time, especially when we recall how thoroughly successful van Meegeren was in perpetrating his fraud. Here, for example, are some of the words with which the discovery of van Meegeren's *Christ and the Disciples at Emmaus* was announced to the world by Abraham Bredius, at that time probably the world's expert on Vermeer:

. . . we have here a—I am inclined to say—the masterpiece of Johannes Vermeer of Delft . . . quite different from all his other paintings and yet every inch a Vermeer. The subject is *Christ and the Disciples at Emmaus* and the colors are magnificent—and characteristic. . . .

In no other picture by the great Master of Delft do we find such sentiment, such a profound understanding of the Bible story—a sentiment so nobly human expressed through the medium of the highest art. . . .

The reproduction [accompanying Bredius's article] can only give a very inadequate idea of the splendid luminous effect of the rare combination of colors of this magnificent painting by one of the greatest artists of the Dutch School.

The picture referred to hung in the Boymans Museum in Rotterdam for seven years. During

that time thousands upon thousands admired and praised the painting. There was no doubt in anyone's mind that this was one of the greatest of Vermeer's paintings and, indeed, one of the most beautiful works of art in the world. It was undoubtedly this universal judgment of aesthetic excellence which accounts largely for the sensational effects of van Meegeren's confession in 1945.

It is of course embarrassing and irritating for an expert to make a mistake in his field. And it was, as it turned out, a mistake to identify the painting as a Vermeer. But it should be obvious from the words of Bredius just quoted that there is more involved here than a mere matter of misidentification. "The colors are magnificent," he writes. "The highest art . . . this magnificent painting . . . the masterpiece of Vermeer": this is more than identification. This clearly is aesthetic praise. And it is just the fact that the critics heaped such lavish praise on a picture which turned out to have been painted by a second-rate contemporary artist that made the van Meegeren case such a painful affair for them. To their way of thinking, which I am trying to show was not very logical, they were now apparently faced with the dilemma of either admitting that they had praised a worthless picture or continuing to do so.

This was of course precisely the trap that van Meegeren had laid for the critics. It was, in fact, the whole *raison d'être* of his perpetrating the fraud. He deliberately chose this extreme, perhaps pathological, way of exposing what he considered to be false aesthetic standards of art critics. In this respect his thinking was no more logical than that of the critics. His reasoning, at least about his first forgery *The Disciples*, was in effect as follows: Once my painting has been accepted and admired as a genuine Vermeer, I will confess publicly to the forgery and thus force the critics either to retract their earlier judgments of praise, thereby acknowledging their fallibility, or to recognize that I am as great an artist as Vermeer. The dilemma as stated contains a difficulty to which we shall come back later. What is important historically is that the critics accepted van Meegeren's dilemma as a genuine one (thereby

becoming the dupes of a logical forgery as well as an artistic one).

2. Forgeries Need Not Lack Aesthetic Value

The plain fact is that aesthetically it makes no difference whether a work of art is authentic or a forgery, and instead of being embarrassed at having praised a forgery, critics should have the courage of their convictions and take pride in having praised a work of beauty. Perhaps if critics did respond in this way we should be less inclined to think that so often their judgments are historical, biographical, economical, or sociological instead of aesthetic. For in a sense, of course, van Meegeren proved his point. Perhaps it is a point for which such radical proof was not even necessary. We all know very well that it is just the preponderance in the art world of non-aesthetic criteria such as fame of the artist and the age or cost of the canvas which is largely responsible for the existence of artistic forgeries in the first place. We all know that a few authentic pen and ink scratches by Picasso are far more valuable than a fine landscape by an unknown artist. If we were offered a choice between an inferior (but genuine) Degas sketch and a beautiful Jones or Smith or X, how many of us would choose the latter? In a museum that did not label its paintings, how many of us would not feel uneasy lest we condemn one of the greats or praise an unknown? But, it may be argued, all this we know. It is simply a fact and, moreover, probably an unavoidable, understandable—even a necessary—fact. Is this so serious or regrettable? The answer of course is that it is indeed serious and regrettable that the realm of art should be so infested with non-aesthetic standards of judgment that it is often impossible to distinguish artistic from economic value, taste or fashion from true artistic excellence, and good artists from clever businessmen.

This brings us to the point of our discussion so far. The matter of genuineness versus forgery is but yet another non-aesthetic standard of judgment. The fact that a work of art is a forgery is an

item of information about it on a level with such information as how old the artist was when he created it, the political situation in the time and place of its creation, the price it originally fetched, the kind of materials used in it, the stylistic influences discernible in it, the psychological state of the artist, his purpose in painting it, and so on. All such information belongs to areas of interest peripheral at best to the work of art as aesthetic object, areas such as biography, history of art, sociology, and psychology. It is not denied here that such areas of interest may be important and that their study may even help us to become better art appreciators. But it is denied that the information with which they deal is of the essence of the work of art or of the aesthetic experience which it engenders.

It would be merely foolish to assert that it is of no interest whatsoever to know that *The Disciples* is a forgery. But to the man who has never heard of either Vermeer or van Meegeren and who stands in front of *The Disciples* admiring it, it can make no difference whether he is told that it is a seventeenth century Vermeer or a twentieth century van Meegeren in the style of Vermeer. And when some deny this and argue vehemently that indeed it does make a great deal of difference, they are only admitting that *they* do know something about Vermeer and van Meegeren and the history of art and the value and reputation of certain masters. They are only admitting that *they* do not judge a work of art on purely aesthetic grounds, but also take into account when it was created, by whom, and how great a reputation it or its creator has. And instead of seeking justification in the fact that in truth it is difficult to make a pure, aesthetic judgment, unbiased by all our knowledge of the history and criticism of art, they generally confuse matters of aesthetics even more by rationalizing that it is the complexity of the aesthetic experience which accounts for the difference made by the knowledge that a work of art is a forgery. That the aesthetic experience is complex I do not deny. But it is not so complex that such items of information as the place and date of creation or the name of the creator of a work of art have to be considered. The fact that *The Disciples* is a forgery is

just that, a fact. It is a fact about the painting which stands entirely apart from it as an object for aesthetic contemplation. The knowledge of this fact can neither add anything to nor subtract anything from the aesthetic experience (as aesthetic), except insofar as preoccupation with it or disappointment on its account may in some degree prevent us from having an aesthetic experience at all. Whatever the reasons for the removal of *The Disciples* from the walls of the Boymans Museum in Rotterdam, they were assuredly not aesthetic.

3. Forgeries Need Not Be Morally or Legally Wrong

And yet, we can all sympathize, or at least understand, why *The Disciples* was removed. It was, after all, a forgery and even if we grant that it is not a matter of aesthetics, it still seems self-evident that forgery remains a normative term implying a defect or absence in its object. In short, we still need to answer our question: What is wrong with a forgery?

The most obvious answer to this question, after the aesthetic one, is that forgery is a moral or legal normative concept, and that thus it refers to an object which, if not necessarily aesthetically inferior, is always morally offensive. Specifically, the reason forgery is a moral offence, according to this view, is of course that it involves *deception*. Reasonable and commonsensical as this view seems at first, it does not, as I will try to show, answer our question adequately.

Now it cannot be denied, I think, that we do in fact often intend little more than this moral connotation when we speak of forgery. Just because forgery is a normative concept we implicitly condemn any instance of it because we generally assume that it involves the breaking of a legal or moral code. This assumption is, however, only sometimes correct. It is important to note this because historically by far the majority of artistic fakes or forgeries have not been legal forgeries. Most often they have been the result of simple mistakes, misunderstandings, and lack of

information about given works of art. We can, as a point of terminology, exclude all such instances from the category of forgery and restrict the term to only those cases involving deliberate deception. There is, after all, a whole class of forgeries, including simple copies, misattributions, composites, and works "in the manner of" some reputable artist, which represent deliberate frauds. In these cases of forgery, which are undoubtedly the most notorious and disconcerting, someone, e.g., artist or art dealer, has passed off a work of art as being something which it is not. The motive for doing so is almost always economic, but occasionally, as with van Meegeren, there is involved also a psychological motive of personal prestige or revenge. In any case, it seems clear that—if we leave out of consideration the factor of financial loss, which can of course be considerable, as again the van Meegeren case proved—such deliberate forgeries are condemned by us on moral grounds, that is, because they involve conscious deception.

Yet, as a final answer to our question as to what is wrong with a forgery, this definition fails. The reason is the following: Although to some extent it is true that passing *anything* off as *anything* that it is not constitutes deception and is thus an undesirable or morally repugnant act, the case of deception we have in mind when we define forgery in terms of it is that of passing off the inferior as the superior. Although, strictly speaking, passing off a genuine de Hoogh as a Vermeer is also an immoral act of deception, it is hard to think of it as a forgery at all, let alone a forgery in the same sense as passing off a van Meegeren as a Vermeer is. The reason is obviously that in the case of the de Hoogh a superior work is being passed off as a superior work (by another artist), while in the van Meegeren case an inferior work is passed off as a superior work.

What is needed, then, to make our moral definition of forgery more accurate is the specification "passing off the inferior as the superior." But it is just at this point that this common-sense definition of artistic forgery in moral terms breaks down. For we are now faced with the question of what is meant by superior and inferior in art. The moral definition

of forgery says in effect that a forgery is an inferior work passed off as a superior one. But what is meant here by inferior? We have already seen that the forgery is not necessarily *aesthetically* inferior. What, then, does it mean? Once again, what is wrong with a forgery?

The attempt to define forgery in moral terms fails because it inevitably already assumes that there exists a difference between genuine works of art and forgeries which makes passing off the latter as the former an offence against a moral or legal law. For only if such a difference does in fact exist can there be any rationale for the law. It is of course precisely this assumed real difference which we are trying to discover in this paper.

4. **Forgeries Lack the Originality Characteristic of the Creative Arts**

It seems to me that the offence felt to be involved in forgery is not so much against the spirit of beauty (aesthetics) or the spirit of the law (morality) as against the spirit of art. Somehow, a work such as *The Disciples* lacks artistic integrity. Even if it is beautiful and even if van Meegeren had not forged Vermeer's signature, there would still be something wrong with *The Disciples*. *What?* is still our question.

We may approach this problem by considering the following interesting point. The concept of forgery seems to be peculiarly inapplicable to the performing arts. It would be quite nonsensical to say, for example, that the man who played the **Bach suites** for unaccompanied cello and whom at the time we took to be **Pablo Casals** was in fact a forger. Similarly, we should want to argue that the term *forgery* was misused if we should read in the newspaper that **Margot Fonteyn**'s performance in *Swan Lake* last night was a forgery because as a matter of fact it was not Margot Fonteyn who danced last night but rather some unknown person whom everyone mistook for Margot Fonteyn. Again, it is difficult to see in what sense a performance of, say, *Oedipus Rex* or *Hamlet* could be termed a forgery.

Here, however, we must immediately clarify our point, for it is easily misunderstood. There is, of course, a sense in which a performance of *Hamlet* or *Swan Lake* or the Bach suites could be called a forgery. If, for example, someone gave a performance of *Hamlet* in which every gesture, every movement, every vocal interpretation had been copied or imitated from the performance of *Hamlet* by Laurence Olivier, we could, I suppose, call the former a forgery of the latter. But notice that in that case we are interpreting the art of acting not as a performing art but as a creative art. For what is meant is that Olivier's interpretation and performance of *Hamlet* is itself an original and creative work of art which can be forged. Similar comments would apply to Margot Fonteyn's *Swan Lake* and Casal's Bach suites and, in fact, to every performance.

My point is then that the concept of forgery applies only to the creative and not to the performing arts. It can be denied of course that there is any such ultimate distinction between creative and performing arts. But we shall still have to admit, I think, that the duality on which it is based—the duality of creativity or originality, on the one hand, and reproduction or technique, on the other—is real. We shall have to admit that originality and technique are two elements of all art; for not only can it be argued that a performance requires more than technique, namely originality, but also that the creation of a work of art requires more than originality, namely technique.

The truth of the matter is probably that both performances and works of art vary greatly and significantly in the degree to which they possess these elements. Their relative presence in works of art and performances makes, in fact, an interesting way of categorizing the latter. But it would be wrong to assert that these two elements are inseparable. I can assure the reader that a portrait painted by myself would be technically almost totally incompetent, and yet even I would not deny that it might be original. On the other hand, a really skilful copy of, for example, a Rembrandt drawing may be technically perfect and yet lack all originality. These two examples establish the two extreme cases of a kind

of continuum. The copy of Rembrandt is of course the forgery *par excellence*. My incompetent portrait is as far removed from being a forgery as any work can be. Somewhere in between lies the whole body of legitimate performances and works of art.

The implications of this long and devious argument are as follows. Forgery is a concept that can be made meaningful only by reference to the concept of originality, and hence only to art viewed as a *creative*, not as a reproductive or technical activity. The element of performance or technique in art cannot be an object for forgery because technique is not the kind of thing that can be forged. Technique is, as it were, public. One does or does not possess it or one acquires it or learns it. One may even pretend to have it. But one cannot forge it because in order to forge it one must already possess it, in which case there is no need to forge it. It is not Vermeer's technique in painting light which van Meegeren forged. That technique is public and may be had by anyone who is able and willing to learn it. It is rather Vermeer's discovery of this technique and his use of it, that is, Vermeer's originality, which is forged. The light, as well as the composition, the colour, and many other features, of course, were original with Vermeer. They are not original with van Meegeren. They are forged.

5. Concepts of Originality in Art

At this point our argument could conclude were it not for the fact that the case which we have used throughout as our chief example, *Christ and the Disciples at Emmaus*, is not in fact a skilful copy of a Vermeer but a novel painting in the style of Vermeer. This threatens our definition of forgery since this particular forgery (always assuming it is a forgery) obviously possesses originality in some sense of the word.

The problem of forgery, in other words, is a good deal more complex than might at first be supposed, and before we can rest content with our definition of forgery as works of art lacking originality, it must be shown that the concept of originality can indeed account for the meaning of forgery as an untrue or objectionable thing in all instances, including even such a bizarre case as that of van Meegeren's *Disciples*.

at Emmaus. It thus becomes important to examine the various possible meanings that the term *originality* may have in the context of art in order to determine in what sense *The Disciples* does and does not possess it, and hence in what sense it can meaningfully and justifiably be termed a forgery.

1) A work of art may be said to be original in the sense of being a particular object not identical with any other object. But this originality is trivial since it is a quality possessed by all things. *Particularity* or *self-identity* would be better names for it.

2) By originality in a work of art we may mean that it possesses a certain superficial individuality which serves to distinguish it from other works of art. Thus, for example, a certain subject matter in a particular arrangement painted in certain colours may serve to identify a painting and mark it as an original work of art in the sense that its subject matter is unique. Probably the term *individuality* specifies this quality more adequately than *originality*.

It seems safe to assert that this quality of individuality is a necessary condition for any work of art to be called original in any significant sense. It is, however, not a necessary condition for a work to be called beautiful or to be the object of an aesthetic experience. A good reproduction or copy of a painting may be the object of aesthetic contemplation and yet it lacks all originality in the sense which we are here considering. Historically many forgeries are of this kind, i.e., more or less skilful copies of existing works of art. They may be described as being forgeries just because they lack this kind of originality and hence any other kind of originality as well. It is to be noticed that the quality which makes such a copy a forgery, i.e., its lack of individuality, is not a quality which exists in the work of art as such. It is a fact about the work of art which can be known only by placing the latter in the context of the history of art and observing whether any identical work predates it.

As we said above, it is not this kind of originality which is lacking in *The Disciples*.

3) By originality in art we may mean the kind of imaginative novelty or spontaneity which is a mark of every good work of art. It is the kind of originality which attaches to individual works of art and which

can be specified in formal or technical terms such as composition, balance, colour intensity, perspective, harmony, rhythm, tempo, texture, thyme, alliteration, suspense, character, plot, structure, choice of subject matter, and so on. Here again, however, in order for this quality to be meaningfully called originality, a reference must be made to a historical context in terms of which we are considering the particular work of art in question, e.g., this work of art is original because the artist has done something with the subject and its treatment which has never been done before, or this work is not original because many others just like it predate it.

In any case, *The Disciples* does, by common consent, possess this kind of originality and is therefore in this sense at least not a forgery.

4) The term *originality* is sometimes used to refer to the great artistic achievement of a specific work of art. Thus we might say that whereas nearly all of Milton's works are good and original in the sense of (c) above, *Paradise Lost* has a particularly profound originality possessed only by really superlative works of art. It is hard to state precisely what is meant by this use of the term *originality*. In justifying it we should probably point to the scope, profundity, daring, and novelty of the conception of the work of art in question as well as to the excellence of its execution. No doubt this kind of originality differs from that discussed under (c) above only in degree.

It is to be noted that it cannot be the lack of this kind of originality which defines a forgery since, almost by definition, it is a quality lacking in many—maybe the majority of—legitimate works of art. Moreover, judging from the critical commentary with which *The Disciples* was received at the time of its discovery—commentary unbiased by the knowledge that it was a forgery—it seems reasonable to infer that the kind of originality meant here is in fact one which *The Disciples* very likely possesses.

5) Finally, it would seem that by originality in art we can and often do mean the artistic novelty and achievement not of one particular work of art but of the totality of artistic productions of one man or even one school. Thus we may speak of the originality

of Vermeer or El Greco or Mozart or Dante or Impressionism or the Metaphysical Poets or even the Greeks or the Renaissance, always referring I presume, to the artistic accomplishments achieved and embodied in the works of art belonging to the particular man, movement, or period. In the case of Vermeer we may speak of the originality of the artists of design in the genre picture, the originality of his use of bright and pure colours, and of the originality of his treatment and execution of light.

It is to be noticed first of all that this meaning of originality, too, depends entirely on a historical context in which we are placing and considering the accomplishment of one man or one period. It would be meaningless to call Impressionism original, in the sense here considered, except in reference to the history of art which preceded it. Again, it is just because Vermeer's sense of pictorial design, his use of bright colours, and his mastery of the technique of painting light are not found in the history of art before him that we call these things original in Vermeer's work.

Originality, even in this more profound sense, or rather especially in this more profound sense, is a quality definable only in terms of the history of art. A second point of importance is that while originality as here considered is a quality which attaches to a whole corpus or style of works of art, it can be considered to exist in one particular work of art in the sense that that work of art is a typical example of the style or movement to which it belongs and therefore embodies the originality of that style or movement. Thus we may say that Vermeer's *A Painter in His Studio* is original because in this painting (as well as in several others, of course) we recognize those characteristics mentioned earlier (light, design, colour, etc.) which are so typical of Vermeer's work as a whole and which, when we consider the whole of Vermeer's work in the context of the history of art, allow us to ascribe originality to it.

Turning our attention once more to *The Disciples*, we are at last in a position to provide an adequate answer to our question as to the meaning of the term forgery when applied to a work of art such as *The Disciples*. We shall find, I think, that the fraudulent character of this painting is adequately

defined by stating that it lacks originality in the fifth and final sense which we have here considered. Whatever kinds of originality it can claim—and we have seen that it possesses all the kinds previously discussed—it is not original in the sense of being the product of a style, period, or technique which, when considered in its appropriate historical context, can be said to represent a significant achievement. It is just this fact which differentiates this painting from a genuine Vermeer! The latter, when considered in its historical context, i.e., the seventeenth century, possesses the qualities of artistic or creative novelty which justify us in calling it original. *The Disciples*, on the other hand, in its historical context, i.e., the twentieth century, is not original, since it presents nothing new or creative to the history of art even though, as we have emphasized earlier, it may well be as beautiful as the genuine Vermeer pictures.

It is to be noted that in this definition of forgery the phrase "appropriate historical context" refers to the date of production of the particular work of art in question, not the date which in the history of art is appropriate to its style or subject matter. In other words, what makes *The Disciples* a forgery is precisely the disparity or gap between its stylistically appropriate and its actual date of production. It is simply this disparity which we have in mind when we say that forgeries such as *The Disciples* lack integrity.

It is interesting at this point to recall van Meegeren's reasoning in perpetrating the Vermeer forgeries. "Either," he reasoned, "the critics must admit their fallibility or else acknowledge that I am as great an artist as Vermeer?" We can see now that this reasoning is not sound. For the notion of greatness involved in it depends on the same concept of historical originality which we have been considering. The only difference is that we are now thinking of it as an attribute of the artist rather than of the works of art. Van Meegeren's mistake was in depending that Vermeer's reputation as a great artist depended on his ability to paint beautiful pictures. If this were so, the dilemma which van Meegeren posed to the critics would have been a real one, for his picture is undeniably beautiful. But, in fact, Vermeer is a great artist only because he could paint beautifully

pictures. He is great for that reason plus something else. And that something else is precisely the fact of his originality, i.e., the fact that he painted certain pictures in a certain manner *at a certain time in the history and development of art*. Vermeer's art represents

a genuine creative achievement in the history of art. It is the work not merely of a master craftsman or technician, but of a creative genius as well. And it is for the latter rather than for the former reason that we call Vermeer great. [. . .]

Work Cited

Bredius, A. "A New Vermeer." *The Burlington Magazine* Nov 1937: 210–1.

Study Questions

1. Describe the dilemma van Meegeren assumed he was creating for the critics when he painted *Christ and the Disciples at Emmaus* in order to pass it off as a work by the seventeenth-century Dutch master Vermeer.
2. What makes a forgery wrong, according to Lessing, and why does this analysis show that van Meegeren did not present the critics with a dilemma after all?
3. What are the five senses of originality in art identified by Lessing? In which sense(s) is *The Disciples* original and in which sense(s) is it not?
4. Lessing emphasizes that aesthetic judgment is autonomous: what makes a painting beautiful is independent of non-aesthetic—e.g., biographical, historical, sociological, or psychological—considerations. What qualities do you think are implicated in aesthetic judgments of beauty?

James O. Young

James O. Young is a professor and chair of philosophy at the University of Victoria. He has a BA from Simon Fraser University, an MA from the University of Waterloo, and a PhD from Boston University. Young specializes in philosophy of language and philosophy of art. He has written three books: *Global Anti-Realism* (1995), *Art and Knowledge* (2001), and, excerpted here, *Cultural Appropriation and the Arts* (2008).

The Aesthetics of Cultural Appropriation

From *Cultural Appropriation and the Arts*

1. The Aesthetic Handicap Thesis

In this chapter [. . .] I will ask whether an artwork that employs appropriated content [. . .] will necessarily have (qua product of cultural appropriation) aesthetic flaws. Potentially both interpretive and creative artists suffer from a handicap. For example, both a Westerner who performs Beijing opera and a Chinese who paints landscapes in the style of Monet

could suffer from an aesthetic handicap. I will reach the conclusion that the works of artists who appropriate content [. . .] do not necessarily suffer from aesthetic flaws. On the contrary, artists who engage in cultural appropriation may produce works of considerable aesthetic merit. [. . .]

This chapter is divided into [five] sections. After this introductory section, I will consider the claim that works by outsiders have flaws that can

be detected without any information about who produced them. In the third section I will consider and accept the proposal that at least some of the aesthetic properties of an artwork depend, in part, on the cultural context in which it was produced. This is to admit that the fact that artists have engaged in content appropriation can affect the aesthetic properties of their works. An argument can be given for the conclusion that the effect is negative. In the fourth section I present arguments for the claim that works by outsiders who engage in content appropriation are flawed because they are inauthentic. In the fifth section these arguments are refuted. [. . .]

In assessing the aesthetic handicap thesis, we need to bear in mind that aesthetic properties come in two varieties. Nelson Goodman famously wrote that “the aesthetic properties of a picture include not only those found by looking at it but also those that determine how it is to be looked at” (111–12). Goodman limited his remarks to pictures, but presumably the same point applies to other sorts of artworks. Once we extend Goodman’s point to other classes of artworks, we would say that the first class of properties includes those that can be heard as well as those that can be seen. Presumably it also includes properties of literary works that are apprehended by intellectual experience. So the triteness or didacticism of a novel that readers can experience, without knowing anything about the circumstances of the work’s production, belongs to the first category. The second sort of property will include the categories to which artworks belong, and which influence the interpretation and evaluation of a work. In order to discover these properties of an artwork, audiences must do more than inspect the work. Audiences need to learn about the circumstances of the work’s production and, in particular, the context in which the work was produced.

Given Goodman’s distinction, the aesthetic handicap thesis can be interpreted in two ways. The first possibility is that, try as they may, when artists appropriate from another culture, they will produce works that have immediately observable aesthetic flaws. On this first interpretation of the thesis, one does not need to know who produced

the work for its flaws to be apparent. That is, the thesis is that when outsiders appropriate content, they will always employ it ineffectively. We will be able to tell, for example, just by listening to a blues performance by a non-African American that it has certain flaws. [. . .]

The alternative interpretation of the aesthetic handicap thesis states that the aesthetic flaws of works by cultural appropriators are properties of Goodman’s second sort. On this interpretation of the thesis, the fact that a work employs content from a culture other than the artist’s own, or represents a culture other than the artist’s, is an aesthetic property. Specifically, it is a negative aesthetic property that detracts from a work’s aesthetic value. Perhaps, for example, once a work is known to be by an outsider it will be experienced as (and actually be) insincere or unoriginal.

2. The Cultural Experience Argument

A common sort of argument tries to establish that works by outsiders will have observable aesthetic flaws. This argument starts from the premise that the ability to use a style successfully is linked to participation in a culture. On this view, artists cannot successfully employ a style unless they have had experiences available only to members of a culture. In other words, the experience of living as a member of a given culture is a necessary condition of being able to create successful works of the types developed by the culture. We may call this the *cultural experience argument* for the aesthetic handicap thesis.

This sort of argument is made about the appropriation of African-American music. A well-known advocate of this view is Amiri Baraka (the blues musician formerly known as LeRoi Jones). He has maintained that a musician cannot learn to produce the blues (or to produce the blues well) except via “the experience of a black man in America. The idea of a white blues singer seems an even more violent contradiction of terms than the idea of a middle-class blues singer. The materials of blues were not available to the white American” (148). (The debate about the

appropriation of the blues is often formulated with reference to race or ethnicity. References to race can be replaced by references to culture.) Crucial to this experience was the discrimination to which African Americans were subject. Baraka writes that "The blues was conceived by freedmen and ex-slaves . . . as an emotional confirmation of, and reaction to, the way in which most Negroes were still forced to exist in the United States" (142). Many non-African Americans do not simply lack the experience of being a member of African-American culture and experiencing discrimination. Given their skin pigmentation, they cannot have this experience. [. . .]

The conclusion of the cultural experience argument is that artists who lack the requisite cultural background are bound to produce works or performances of a poor aesthetic quality. According to Baraka, the result of the appropriation of the blues was commercialized and diluted music. [. . .] Since the blues originated in a specific cultural context, the cultural experience argument has an initial plausibility. Many performances of blues by musicians whose cultural background is non-African-American undeniably are aesthetically poor. They do not measure up to the standards established by pioneering African-American blues artists such as Ma Rainey, John Lee Hooker, Muddy Waters, and so on.

Despite its initial plausibility, the cultural experience argument faces telling objections. The first point to make is that even if artists cannot successfully master a style or genre developed in another culture, it does not follow that they are condemned to aesthetic failure when they engage in cultural appropriation. We need to distinguish between two ways in which artists can engage in content appropriation. The first may be called *non-innovative content appropriation*. We would have an example of such appropriation if an American performer were to attempt to enter into the tradition of Japanese epic ballad recitation by chanting (in Japanese) the *Tale of the Heike* while accompanying himself on a *biwa*. Artists engaged in non-innovative content appropriation are not creating a new category of artwork, but adding to a category that already exists. They attempt to succeed by the standards already established within the

culture from which they are appropriating. In this case, a non-Japanese performer is trying to master the techniques of an existing Japanese tradition.

Alternatively an artist might engage in *innovative content appropriation*. Artists who engage in this sort of appropriation appropriate a style or a motif from a culture but use it in a way that would not be found in the culture in which it originated. Picasso was engaged in innovative content appropriation (specifically: motif appropriation) when he borrowed ideas from African carvers in such paintings as *Les Demoiselles d'Avignon*. Although African carving influenced Picasso, he did not produce a work that belongs to any tradition of African carving. Similarly, Ravel borrowed from African-American culture in such jazz-influenced works as the Sonata for Violin and Cello (1922), the Sonata for Violin and Piano (1927), and the two piano concerti (the Piano Concerto for the Left Hand in D Major, composed for Paul Wittgenstein, and the Concerto in G Major) he composed between 1929 and 1931. Although these are not jazz compositions, we find in them innovative content appropriation. [. . .]

Equipped with the distinction between innovative and non-innovative content appropriation we are in a position to see the first problem with the cultural experience argument. At best it shows that artists who engage in non-innovative content appropriation are doomed to aesthetic failure. Perhaps the middle-class white college student who sings the blues, imitating the singers he has heard in recordings, will inevitably produce derivative and uninspired music. Perhaps my imaginary American performer will never be able to recite epic Japanese poetry as well as a Japanese. Even if this is so, the cultural experience argument provides us with no reason to believe that outsiders will not be able to make aesthetically successful innovative re-use of styles and motifs developed by other cultures. I take it that the examples of motif appropriation by Picasso and Ravel are sufficient to establish this point.

Even Baraka, who endorses the aesthetic handicap thesis, acknowledges this point. Middle-class whites are, on his view, unable to produce good music in an African-American style. Still, he admits that they are

able to appropriate from African-American culture and produce something valuable. Baraka concedes this in the contexts of a discussion of the trumpet playing of **Bix Beiderbecke**. According to Baraka, Beiderbecke's playing was "certainly an appropriation of black New Orleans brass style, most notably King Oliver's" (151). At the same time, Baraka recognizes that Beiderbecke was a successful musician. He "played 'white jazz' . . . music that is the product of attitudes expressive of a peculiar culture." Still, "the serious white musician" such as Beiderbecke was in a position to make creative re-use of what he had appropriated and he produced good music (154). [...]

The response just given to the cultural experience argument does not completely undermine the argument. It may still be successful against non-innovative appropriation. If so, this would be worrying since much cultural appropriation is likely to be non-innovative. Picasso, Ravel, and Beiderbecke were great artists and less talented and innovative artists will perform much cultural appropriation. A little reflection is sufficient to show that the cultural experience argument is unable even to show that non-innovative content appropriation will necessarily suffer from observable aesthetic flaws.

Very little evidence exists for the claim that mastery of an artistic style is linked to membership in a culture but vast amounts of evidence can be marshalled in support of the opposite claim. Many examples can be given of artists from diverse cultures who are successful practitioners of the same style. (At least, they are apparently successful. Consideration of the second sort of aesthetic property is still pending.) Consider *bel canto* singing, originally produced in the context of Italian culture. Singers with diverse cultural backgrounds have fully mastered the style: Kathleen Battle (an African-American) and Kiri Te Kanawa (half Maori) no less than Cecilia Bartoli. The greatest composer in the French baroque style was Jean-Baptiste Lully (born Giovanni Battista Lulli in Florence). Arguably the most proficient master of the Italian baroque style was *Il caro Sassone* ["the dear Saxon"], George Frederic Handel.

What is true of Italian musical styles is as true of those that originated in the context of

African-American culture. Ray Eldridge, the distinguished African-American jazz trumpeter, was an adherent of the cultural experience argument. He once bet Leonard Feather, the music critic, that he could reliably tell the difference between jazz performances by African Americans and non-African Americans. If anyone could tell the difference, Eldridge could. Widely regarded as the greatest trumpet soloist of his time, Eldridge's credentials were impeccable. Put to the test, he failed miserably. In blind listening situations, he misidentified the cultural origins of the performer more than half of the time. (Feather conducted a series of "blindfold tests" for *Down Beat* during his long tenure at that magazine.)

Let us return to our starting point: the blues. It is possible to name any number of outsiders who are good blues musicians by the standards established by insiders. Marcia Ball, John Hammond, James Harman, Charlie Musslewhite, Stevie Ray Vaughan, and Johnny Winter are all examples of outsiders (relative to African-American culture) who are successful blues musicians by the standards of insiders. Many of them are multiple winners of W. C. Handy Blues Awards. Presumably such a distinguished bluesman as Buddy Guy would notice the flaws in the blues of Eric Clapton, a born and bred Englishman. Yet Guy has said with reference to Clapton that "all I want to do is hear him play. Race, size, color, nothing matters when a guy's got it, and Eric's got it" (in Taylor 313). Guy was perfectly aware that Clapton is an outsider relative to African-American culture. [...]

The fact that there is no hard and fast link between culture and artistic success should not surprise us. The cultural experience argument only alleged that experience of living in a culture was a necessary condition of being able to create successful works in certain categories. It never said that such experience was a sufficient condition of being able to do so. Nor should it have. After all, many people belong to a given culture and yet are unable to produce successful works of art in the categories typical of their culture. (I, for example, am unable to produce a successful work of art in any of the genres characteristic of my culture.) Given that membership in a culture

is not a sufficient condition of being able to produce good works of a certain sort, we might well ask what else is necessary.

It is probably impossible to specify sufficient conditions for the successful production of works in any genre but it is fairly easy to identify a few necessary conditions. Aspiring artists must undergo a process of training. (Sometimes, an artist can be self-trained.) Artists must study successful works of the sort that they aspire to produce. They must become familiar with techniques and materials. Often they need to learn the significance of certain symbols, conventions, and so on. [...] Perhaps, most importantly, artists must repeatedly practise their craft. These are the really important preconditions of being a successful artist in any style. Notice that no mention is made of the cultural background of the aspiring artist. The required training can be undertaken by anyone. Being able to work in a given style is like learning a language and there is no reason why outsiders cannot learn this language every bit as well as insiders.

3. Aesthetic Properties and Cultural Context

In the previous section I assumed that if outsiders who engage in cultural appropriation labour under an aesthetic handicap, they will produce works with observable aesthetic flaws. I have argued that when outsiders engage in non-innovative appropriation they are successful when experience cannot reliably distinguish their works from those produced by good insider artists. Even if we cannot immediately determine which works are by outsiders, we cannot rule out the possibility that their works inevitably have aesthetic flaws. Perhaps they have flaws because of the second sort of aesthetic property Goodman identified. Perhaps facts about the origin of works by outsiders will enable us to see or interpret works by outsiders in new ways. Perhaps when we reinterpret works by outsiders, flaws will become apparent.

A long and distinguished tradition in aesthetics would have us reject Goodman's distinction. According to this empiricist tradition, nothing

is relevant to the aesthetic value of a work of art besides its observable properties. Facts about the intentions, beliefs, provenance, or (most importantly given present concerns) the cultural background of an artist have no relevance to the aesthetic properties of a work of art. Anyone who subscribes to this tradition will believe that much of what follows in this section is not only unnecessary but misguided. I am inclined, however, to think that Goodman's distinction cannot be disregarded.

The way in which information about the origin of an artwork can affect its aesthetic properties can be illustrated by a well-known literary thought experiment. **Jorge Luis Borges** famously compared the *Don Quixote* (1605) of Cervantes and the (imaginary) *Don Quixote* of Pierre Menard. In Part I, Chapter 9, of Cervantes' *Quixote* we find the following passage: "Truth, whose mother is history, who is the rival of time, depositor of deeds, witness of the past, example and lesson to the present, and warning to the future . . ." Menard in the early twentieth century, Borges imagines, writes instead: "Truth, whose mother is history, who is the rival of time, depositor of deeds, witness of the past, example and lesson to the present, and warning to the future . . ." Written in the seventeenth century, Cervantes' "enumeration is a mere rhetorical eulogy of history." For Menard, the contemporary of **William James**, the same passage is a statement of pragmatism: "Historical truth, for him, is not what took place; it is what we think took place" (102). This illustrates how information about the origin of a literary work can influence how it is interpreted and what properties it has.

The same point can be made in less fanciful terms. In the early eighteenth century, King George I described the work of the architect Sir Christopher Wren as "amusing, awful and artificial." He meant that it was amazing, awe-inspiring, and artistic. Obviously, the King's words do not have the same meaning today. The meanings of words depend on the contexts in which they are originally used. The meaning of words in a work of literature contributes to its aesthetic properties. Consequently, the aesthetic properties of a literary work depend on the work's original context. A similar point can be made

about the properties of musical and visual artworks. Certain harmonies will be dissonant if produced in the eighteenth century. The same harmonies will not be dissonant in a twentieth-century composition.

An even more general argument shows that the aesthetic properties of an artwork can depend on facts about its production. I have in mind the argument in Kendall Walton's classic paper, "Categories of Art." Walton argued that we can only determine what aesthetic properties an artwork has once we have determined the category to which it belongs. Indeed, a work only has aesthetic properties in the context of a category. One cannot, for example, say that a painting is gaudy *tout court*. First we must determine the category to which it belongs. Imagine that we are presented with a brightly coloured landscape. If it belongs to the category of Wu School landscape painting (exemplified by the works of Shen Zhou), it may very well be gaudy. If the painting belongs to the category of post-impressionist landscapes (say it is in the style of Van Gogh or Matisse), then probably it is not gaudy. In determining the category to which a work of art belongs we take into account a wide variety of factors. Among these factors will be knowledge about artists and their historical contexts. The culture of an artist is a large part of the historical context.

A few examples will illustrate how the cultural background of an artist can have an impact on the aesthetic properties of a work. Imagine that we have a certain number of artworks which are thought to have been produced by insiders. One is apparently a typical Mississippi Delta blues. Another is a novel about the *ante bellum* [pre-Civil War] American south, reputedly written by a freed slave. Imagine now that it turns out that the blues piece was not performed by an African-American insider, but is rather by a university student from Beijing. Or imagine that it turns out that a slave owner wrote the novel in question. [...] The cultural milieu of an artist is an important part of a work's original context. So it could well be that critics are justified in revising their assessment of the works in question. They may have been wrong when they judged the novel to be a sensitive and poignant indictment of slavery. Perhaps it is a sly parody of African-American manners.

Similarly, the blues performance will not have been a heartfelt expression of the plight of poor African-American sharecroppers. The judgment that it was is mistaken, even though no observable features of the performance would justify the conclusion that the judgment is mistaken.

4. Authenticity and Appropriation

Even if the aesthetic properties of an artwork depend on the cultural context in which it was produced, we still do not have an argument that shows that works produced by outsiders are necessarily subject to aesthetic flaws. No developed philosophical literature directly supports the aesthetic handicap thesis. There is, however, a literature that can be adapted to argue for the thesis. In introducing his distinction between the two types of aesthetic properties, Goodman was concerned about forgeries. Goodman, and many writers who have followed his lead, have held that forgeries and originals have different aesthetic properties. This could be the case even when a forgery and an original cannot be distinguished simply by experience of the works. A number of writers have suggested that forgeries and originals originate in different contexts and that this has an impact on the aesthetic properties of the works. In the case of forgeries, the impact is thought to be negative. One might similarly maintain that the cultural backgrounds of artists affect the aesthetic properties of artworks and affect them negatively.

Several philosophers have advanced arguments that open up this possibility. Colin Radford put the question in the following terms. He asked whether a critic could "be justified in feeling differently about a painting after he has learned . . . that it was not what he thought" (67). He has in mind the discovery that a painting is not, say, by Vermeer but by van Meegeren. (Famously, Han van Meegeren produced several paintings that he attributed to the great seventeenth-century Dutch master.) The same issue can arise in the context of cultural appropriation. Radford is only one example of a philosopher who has presented variations on what we may call the *authenticity argument*.

In Radford's version of the argument, we are asked to imagine that the discovery that some painting is a forgery is not accompanied by any discovery of observable aesthetic flaws. Imagine, for example, that some work is now known to be by van Meegeren, and not by Vermeer, but that everyone still agrees that the painting is very fine. Unlike other van Meegerens, it displays a profound knowledge of anatomy, the modelling is excellent, it contains no flaws in the perspective, the composition is balanced, and so on. The best critics agree that it is emotionally charged. There may still be grounds for revising the assessment of the painting given when it was thought to be by Vermeer. For example, when thought to be by Vermeer the work could be judged to be original. When known to be by van Meegeren, Radford suggests, the painting now looks derivative. It is not an expression of an original perspective, and in this sense, the work is not authentic.

Radford is not the only philosopher to argue along these lines. W.E. Kennick presents a related argument and explicitly extends it to cover the case of cultural appropriation. He writes that "I cannot, for example, even if I had the appropriate technical skills, paint a picture in the style of the thirteenth-century Chinese master Mu Ch'i (or Fa Ch'ang). The best I can do is *imitate* (ape, mimic) his style" (7). Kennick was primarily concerned about forgery. His point is, however, one that can be applied to acts of cultural appropriation. Kennick went on to observe that "The *Caravaggisti*, because they were in the suitable historical and geographical position to do so, adapted, *in parte*, the style of *Caravaggio*" (7). He might have added that the *Caravaggisti* were in the right cultural position to paint in the style of Caravaggio. [...] Since cultures have changed, no one today, Kennick leads us to believe, can authentically paint in the style of Mu Chi or Caravaggio. Anyone who tried would produce an inauthentic imitation instead of a bold, original work. [...]

Martin A. Bestman reaches a similar conclusion. He states that no one today can authentically paint in the style of—he varies the example—*Masaccio*. His reason for stating this is that any present-day artist "stands in a different relationship to his craft than

those painters of a half millennium ago whose style he captures. He, after all, has to 'act to forget' the technical progress of his craft. Further, the content of his painting—its symbols, allusion, and interests—speak to a social world long gone" (116). Anyone who attempts to appropriate the style of painters from long ago is condemned, on this view, to produce works that are inauthentic.

Bestman is concerned in this passage with artists who employ styles from the past. He goes on to suggest that a similar inauthenticity will exist when an artist appropriates something from foreign contemporary cultures. He does not develop this point, but one can easily imagine how the argument would go. Suppose I were to sculpt a house pole in the style of the Songish or Salish cultures of Vancouver Island and suppose that I am a technically gifted sculptor. I could include the Thunderbird (eagle) and other totems without having the smallest idea about their mythological significance. As used by an insider, an image may have rich symbolic significance. It may be the insigne of a clan or of a deity. As used by an outsider, the same image is simply a strong graphic design. Since the outsider's work lacks the symbolic or cultural significance of works by insiders, there is a sense in which it is inauthentic. Moreover, Bestman would say that outsiders are also not employing the full range of skills they have as (let us suppose) academy-trained sculptors. This suggests that their work would be inauthentic in another way: they are not fully expressing themselves.

5. Authentic Appropriation

Each of the versions of the authenticity argument just presented leads to the conclusion that outsiders who appropriate content will produce works that are inauthentic. Before these arguments can be assessed we need to say a little more about the concept of authenticity. We need to be clearer about what authenticity is and why it is an aesthetic virtue. I will consider [three] sorts of authenticity. Not all of these sorts of authenticity are aesthetic virtues.

The first point to make is that we are not here concerned with authenticity in the sense of being

produced by insiders. In this sense, the authenticity of an artwork depends on its provenance. We may say that a work is authentic when it is in a style and genre of a given culture and insiders (relative to that culture) have produced it. Otherwise, it is inauthentic. Works that are authentic in this sense I will call *provenance authentic*.

Many people are concerned with provenance authenticity. Insider artists are justifiably upset when works by outsiders are represented as works by insiders. Apart from anything else, such fraud can harm the economic interests of insiders. Outsiders are also harmed. Collectors of, for example, Australian aboriginal art and the art of North American First Nations, often want to know that an artwork has been produced by an insider. Misrepresentation of the origins of arts and crafts is common. According to one estimate, aboriginal Australians paint only about half of the didgeridoos sold (www.didjshop.com/authenticity.html). Many fewer are cut by aboriginal craftsmen. Nevertheless, almost all of the instruments are sold as "authentic aboriginal didgeridoos." [. . .]

Here provenance authenticity is not at stake. For a start, we are concerned with works that are, by hypothesis, known to be by outsiders. More importantly, provenance authenticity is not, in itself, an aesthetic merit. A work could be provenance authentic but have a low aesthetic value. After all, every culture is home to poor artists. More importantly still, the suggestion that provenance inauthenticity is an aesthetic demerit begs the question. At issue is the question of whether works by outsiders necessarily have an aesthetic flaw qua works by outsiders. One cannot argue for the conclusion that they do by reiterating that such works are produced by outsiders.

We can identify a second sort of authenticity that plausibly is an aesthetic property. Authenticity, in this second sense of the word, is the property of being the product of an artist's individual genius. An inauthentic work is derivative or even imitative. Kennick and Radford seem to be concerned with this sort of authenticity. It is also akin to a sort of authenticity discussed by Peter Kivy. I will borrow Kivy's term and say that works that are an expression of an

artist's individual genius are characterized by *personal authenticity*. Those that are not are characterized by personal inauthenticity. Personal authenticity is an aesthetic merit. All things being equal, a work of art that is an original expression of an artist's genius is more valuable than a derivative one. The original artwork opens up new perspectives. It excites the imagination in new ways.

Now we need to ask the question of whether works by outsiders are necessarily characterized by personal inauthenticity. I do not see why this should be so. In order to see that appropriation can result in authentic works, we can begin by recalling the distinction between innovative and non-innovative content appropriation. There may be a reason to think that non-innovative content appropriation will not be personally authentic. It is in the nature of such works to be heavily indebted to other works. We have no similar reason to think that innovative content appropriation will be personally inauthentic. On the contrary, innovative content appropriation seems likely to result in personally authentic works.

Still, perhaps the artist who engages in non-innovative content appropriation necessarily produces personally inauthentic works. Certainly, anyone who engages in non-innovative content appropriation produces a work that is, to some extent, derivative. That said, every work of art, or at least every work of art that audiences can appreciate, is derivative to some extent. Unless a work fits into some category of artwork, audiences have no idea about how to appreciate it (Carroll 211ff). It does not follow from this that a work is not personally authentic. A debt to an existing tradition does not remove the possibility of personal authenticity. Any tradition provides scope for artists to innovate. Artworks can owe a great deal to previously existing works and still be personally authentic.

Certainly artists who borrowed liberally from others have produced some of the greatest art of all time. **Johann Sebastian Bach** was a conservative composer whose style closely approximates the style of his immediate predecessors. Or consider the music of **George Frederic Handel**. As long ago as the eighteenth century, Uvedale Price remarked that

"If ever there was a truly great and original genius in any art, Handel was that genius in music; and yet, what may seem no slight paradox, there never was a greater plagiary. He seized [that is, appropriated], without scruple or concealment, whatever suited his purpose" (573). And yet, in borrowing from others, Handel created great masterpieces. William Boyce remarked of Handel's borrowings that "He takes other men's pebbles and polishes them into diamonds." Handel was quite unashamed of his appropriation of material from other composers. When asked about a theme he had appropriated from Bononcini, Handel is said to have replied, "It's much too good for him; he did not know what to do with it." Artists such as Bach and Handel could, though they were greatly indebted to their predecessors, produce personally authentic compositions. Though they borrowed from others they also contributed a spark of individual genius. [. . .]

The examples of borrowing I have just given are not examples of cultural appropriation. (Though perhaps, in borrowing from Bononcini, Handel was engaged in cultural appropriation. [. . .]) One might grant that works that draw on other works can be personally authentic, but still think that outsiders are more likely to produce derivative work than insiders. I see no reason to believe that this is so. All artists working within a tradition depend on earlier artists, and in this sense, their work is derivative. This is as true of insiders as it is of outsiders. Indeed, in many artistic traditions, innovation is not valued. Consider, for example, the sculptures for the housing of spirits produced by the Kalabari of southern Nigeria. Close resemblance to previously existing sculptures is stated as a desideratum of new ones (Layton 7–8). Consequently, if anything, one might expect that the outsider, who may have the benefit of exposure to a wider range of artistic styles, may have an advantage over the insider. Broader artistic experience will make it easier to produce works characterized by personal authenticity.

Turn now to a consideration of a third sort of authenticity, one suggested by Bestman's remarks. [. . .] A work of art can be authentic in the sense that it is something to which the producing artist is fully

committed. To say that artists are committed to their work is to say that they understand the full significance of all of the elements of their work and that they embrace their work's perspectives on the world. Artists could be ignorant of the significance of symbols and other elements of their works. Alternatively, they could be aware of the significance of the elements of their works, but not embrace them. For example, I might not believe a myth I retell. In either case, the work would be inauthentic. This sort of inauthenticity is akin to the authenticity discussed by existentialists such as Heidegger and Sartre. Existentialists expect individuals to be fully committed to everything that they do. Inspired by the existentialists, I will identify another sort of authenticity, *existential authenticity*.

Before proceeding any further, we need to ask whether existential authenticity is an aesthetic merit. One can make a case on either side of this question. As we have already seen, a well-established aesthetic school instructs us to pay attention only to the perceptible qualities of artworks. According to this school, artists' grasp of the significance of the elements of their artworks is irrelevant to the interpretation and appreciation of their works. If this sort of approach to interpretation and appreciation is right, we still need to be concerned about existential authenticity. It may be the case that artists who are unfamiliar with the significance of symbols and imagery that they employ will not use them skilfully. If so, this will be apparent from an examination of the work.

Many people believe that artists' commitment to their work is more important than the previous paragraph suggests. A work of art (at least very often) presents a perspective on the world. It matters to audiences that artists are committed to the perspectives presented in their works. Consider for example two artists, one a Christian and the other an atheist, who both paint a Crucifixion. Imagine also that the paintings are visually indistinguishable. Audiences, particularly audiences of Christians, might very well care about whether an artist is committed to the Christian perspective on Christ's Passion. In all probability, if a church is going to commission an

altarpiece, it will commission the painting from the artist who is committed to the perspective presented in the painting. They may do so on the grounds that the Christian's altarpiece is existentially authentic while the atheist's painting is not. This may indicate that existential authenticity is an aesthetic virtue.

An example that is more explicitly a case of cultural appropriation seems to lead to the same conclusion. A musician who is a member of mainstream American culture could sing the blues without a full awareness of the significance of this style and its elements. This musician may use the words "mojo" and "hep" without any idea of their meanings within African-American culture. More generally, the musician might perform blues numbers without any idea of the cultural significance that the blues has for African-American culture. Such a singer would produce existentially inauthentic performances. An audience from African-American culture may well perceive this as an aesthetic flaw. [...]

This strikes me as a plausible position, but a contrary position also seems plausible. Consider, for example, performances of Bach's church cantatas. These cantatas had a very specific cultural significance. They were written for Lutheran church services and Bach was deeply committed to the idea that they presented a Christian perspective. (He marked each of them S.D.G.—*Sola Dei Gloria*.) Presumably only a Lutheran or, at any rate, a Christian can give an existentially authentic performance of these works. If existential authenticity is an aesthetic virtue, it follows that a performance not by a Christian has an aesthetic flaw. Nevertheless it seems credible that a performance by Chinese atheists or Israeli Jews could be an aesthetic triumph. The same point can be made about the works of outsiders who engage in cultural appropriation. An outsider may not be able to produce an existentially authentic work in the style of insiders, but perhaps it could be (like some atheists' performance of, say, Bach's "Ich habe genug") an aesthetic success.

The claim that (at least some) artworks and performances by outsiders must be existentially inauthentic (and deeply aesthetically flawed) and the claim that they can be aesthetically successful both seem

plausible. Perhaps the contradiction is only apparent. One could say that the aesthetic value of an artwork or a performance is relative to an audience. Perhaps relative to secular music lovers the atheists' performance of "Ich habe genug" is an aesthetic triumph. Perhaps, on the other hand, relative to the devout Lutheran audience the performance is an aesthetic failure because it is existentially inauthentic. One might posit a similar relativity in the context of cultural appropriation. The African-American audience might find a blues performance existentially inauthentic, and so badly aesthetically flawed. The audience of outsiders might be unconcerned by the same existential inauthenticity. Relative to them, the same blues performance [...] has high aesthetic value.

It would be better to resolve this issue without recourse to relativism. I am skeptical about the claim that existential authenticity is an aesthetic virtue. (At a certain point I want to say that the audiences of Christians and African-Americans that I have just imagined are unreasonably prejudiced against certain works.) In any case, even if we grant that existential authenticity is an aesthetic virtue, it does not follow that artists who engage in cultural appropriation cannot produce existentially authentic works and performances. Certainly, some outsiders will produce existentially inauthentic works. The middle-class college student who parrots blues lyrics without understanding their significance is just such an outsider but not all outsiders produce such inauthentic works. One way for outsiders to produce existentially authentic works is to engage in innovative content appropriation. Consider again Picasso's appropriation of motifs from West African carvers. Very likely Picasso was unaware of the significance these motifs had within their original cultural context. It does not follow from this that *Les Demoiselles d'Avignon* is not existentially authentic. Good artists can endow the elements of their artworks with a fresh significance. This is precisely what Picasso and other artists have done. These artists will not produce existentially authentic instances of works in the categories that insiders possess. But they will produce instances of works in new categories and there is no reason why these

cannot be existentially authentic and works of high aesthetic value. The insiders may judge that the work has a low aesthetic value because artists cannot produce an existentially authentic example of a work in the style of insiders. This judgment would be mistaken since it miscategorizes the work.

Still, even if this argument is accepted, it may seem that non-innovative content appropriation will be existentially inauthentic. To a large extent this conclusion can be avoided. An atheist who sings a Bach cantata can be fully committed to the work as an expression of themes of repentance, forgiveness, or longing, even if these themes are removed from their original theological contexts. Similarly, artists can use content appropriated from another culture to express perspectives to which they are fully committed. In both of these cases, we can say that artists have produced works or performances characterized by existential authenticity. Insiders may still regard the works as existentially inauthentic since

the outsider does not accept certain beliefs current in their culture. The response to this should be that this is irrelevant. The artist is fully committed to his work and this is sufficient to make it existentially authentic. Still, it must be admitted that outsiders cannot produce existentially authentic works or performances if they simply ape the artistic practices of members of another culture. Sometimes artists will ape others, and to the extent that existential inauthenticity is an aesthetic flaw, these artists will produce unsuccessful works of art. [. . .]

I conclude that in many senses of the word "authentic," outsiders can produce authentic works in styles appropriated from other cultures. Even if we accept, as I do, that the aesthetic properties of artworks depend on the cultural context in which they were produced, outsiders who engage in cultural appropriation do not labour under an aesthetic handicap that condemns them to produce seriously flawed works of art.

Works Cited

- Bestman, M.A. "Anachronistic Inauthenticity in Art." *Journal of Aesthetic Education* 18, 3 (1984): 115–18.
- Borges, J.L. "Pierre Menard, Author of the Quixote." *Borges: A Reader*. Eds. E. Rodriguez Monegal and A. Reid. New York: E.P. Dutton, 1981. 96–103.
- Carroll, N. "Art, Creativity and Tradition." *The Creation of Art: New Essays in Philosophical Aesthetics*. Cambridge: Cambridge University Press, 2003.
- Goodman, N. *Languages of Art*. Indianapolis, IN: Bobbs-Merrill, 1968.
- Jones, L. *Blues People*. New York: William Morrow, 1963.
- Kennick, W.E. "Art and Inauthenticity." *Journal of Aesthetics and Art Criticism* 44 (1985): 3–12.
- Kivy, P. *Authenticities: Philosophical Reflection on Musical Performance*. Ithaca, NY: Cornell University Press, 1995.
- Layton, R. *The Anthropology of Art*. 2nd ed. Cambridge: Cambridge University Press, 1991.
- Price, U. *On the Picturesque: With an Essay on the Origin of Taste*. Edinburgh: Caldwell, Lloyd, 1842.
- Radford, C. "Fakes." *Mind* 87 (1978): 66–76.
- Taylor, P.C. "So Black and Blue: Response to Rudinow." *Journal of Aesthetics and Art Criticism* 53 (1995): 313–17.
- Walton, K. "Categories of Art." *Philosophical Review* 79 (1970): 334–67.

Susan Wolf

Susan Wolf is Edna J. Koury Distinguished Professor at University of North Carolina at Chapel Hill. She has a BA from Yale and an MA and a PhD from Princeton. Before joining UNC in 2002, Wolf held positions at Harvard, Maryland, and Johns Hopkins. Wolf's research is in ethics and its links with philosophy of mind, philosophy of action, political philosophy, and aesthetics. Wolf was recently president of the eastern division of the American Philosophical Association. In this article, she offers what she considers to be an objective definition of what makes a life meaningful.

The Meanings of Lives

1. Introduction

This question, "What is the meaning of life?" was once taken to be a paradigm of philosophical inquiry. Perhaps, outside of the academy, it still is. In philosophy classrooms and academic journals, however, the question has nearly disappeared, and when the question is brought up, by a naive student, for example, or a prospective donor to the cause of a liberal arts education, it is apt to be greeted with uncomfortable embarrassment.

What is so wrong with the question? One answer is that it is extremely obscure, if not downright unintelligible. It is unclear what exactly the question is supposed to be asking. Talk of meaning in other contexts does not offer ready analogies for understanding the phrase "the meaning of life." When we ask the meaning of a word, for example, we want to know what the word stands for, what it represents. But life is not part of a language, or of any other sort of symbolic system. It is not clear how it could "stand for" anything, nor to whom. We sometimes use "meaning" in non-linguistic contexts: "Those dots mean measles." "Those footprints mean that someone was here since it rained." In these cases, talk of meaning seems to be equivalent to talk of evidence, but the contexts in which such claims are made tend to specify what hypotheses are in question, within relatively fixed bounds. To ask what life means, without a similarly specified context, leaves us at sea.

Still, when people do ask about the meaning of life, they are evidently expressing some concern or

other, and it would be disingenuous to insist that the rest of us haven't the faintest idea what that is. The question at least gestures toward a certain set of concerns with which most of us are at least somewhat familiar. Rather than dismiss a question with which many people have been passionately occupied as pure and simple nonsense, it seems more appropriate to try to interpret it and reformulate it in a way that can be more clearly and unambiguously understood. Though there may well be many things going on when people ask, "What is the meaning of life?" the most central among them seems to be a search to find a purpose or a point to human existence. It is a request to find out why we are here (that is, why we exist at all), with the hope that an answer to this question will also tell us something about what we should be doing with our lives.

If understanding the question in this way, however, makes the question intelligible, it might not give reason to reopen it as a live philosophical problem. Indeed, if some of professional philosophy's discomfort with discussion of the meaning of life comes from a desire to banish ambiguity and obscurity from the field, as much comes, I think, from the thought that the question, when made clearer, has already been answered, and that the answer is depressing. Specifically, if the question of the Meaning of Life is to be identified with the question of the purpose of life, then the standard view, at least among professional philosophers, would seem to be that it all depends on the existence of God. In other words, the

going opinion seems to be that if there is a God, then there is at least a chance that there is a purpose, and so a meaning to life. God may have created us for a reason, with a plan in mind. But to go any further along this branch of thinking is not in the purview of secular philosophers. If, on the other hand, there is no God, then there can be no meaning, in the sense of a point or a purpose to our existence. We are simply a product of physical processes—there are no reasons for our existence, just causes.

At the same time that talk of Life having a Meaning is banished from philosophy, however, the talk of lives being more or less *meaningful* seems to be on the rise. Newspapers, magazines, self-help manuals are filled with essays on how to find meaning in your life; sermons and therapies are built on the truism that happiness is not just a matter of material comfort, or sensual pleasure, but also of a deeper kind of fulfillment. Though philosophers to date have had relatively little to say about what gives meaning to individual lives, passing references can be found throughout the literature; it is generally acknowledged as an intelligible and appropriate thing to want in one's life. Indeed, it would be crass to think otherwise.

But how can individual lives have meaning if life as a whole has none? Are those of us who suspect there is no meaning to life deluding ourselves in continuing to talk about the possibility of finding meaning in life? (Are we being short-sighted, failing to see the implications of one part of our thought on another?) Alternatively, are these expressions mere homonyms, with no conceptual or logical connections between them? Are there simply two wholly unconnected topics here?

Many of you will be relieved to hear that I do not wish to revive the question of whether there is a meaning to life. I am inclined to accept the standard view that there is no plausible interpretation of that question that offers a positive answer in the absence of a fairly specific religious metaphysics. An understanding of meaningfulness in life, however, does seem to me to merit more philosophical attention than it has so far received, and I will have some things to say about it here. Here, too, I am

inclined to accept the standard view—or a part of the standard view—viz., that meaningfulness is an intelligible feature to be sought in a life, and that it is, at least sometimes, attainable but not everywhere assured. But what that feature is—what we are looking for—is controversial and unclear, and so the task of analyzing or interpreting that feature will take up a large portion of my remarks today. With an analysis proposed, I shall return to the question of how a positive view about the possibility of meaning in lives can fit with a negative or agnostic view about the meaning of life. The topics are not, I think, as unconnected as might at first seem necessary for their respectively optimistic and pessimistic answers to coexist. Though my discussion will offer nothing new in the way of an answer to the question of the meaning of life, therefore, it may offer a somewhat different perspective on that question's significance.

2. What Makes a Meaningful Life?

Let us begin, however, with the other question, that of understanding what it is to seek meaning in life. What do we want when we want a meaningful life? What is it that makes some lives meaningful, others less so?

If we focus on the agent's, or the subject's, perspective—on a person wanting meaning in her life, her feeling the need for more meaning—we might incline toward a subjective interpretation of the feature being sought. When a person self-consciously looks for something to give her life meaning, it signals a kind of unhappiness. One imagines, for example, the alienated housewife, whose life seems to her to be a series of endless chores. What she wants, it might appear, is something that she can find more subjectively rewarding.

This impression is reinforced if we consider references to "meaningful experiences." (The phrase might be applied, for example, to a certain kind of wedding or funeral.) The most salient feature of an event that is described as meaningful seems to be its "meaning a lot" to the participants. To say that a ceremony, or, for that matter, a job, is meaningful seems at the very least to include the idea that it

is emotionally satisfying. An absence of meaning is usually marked by a feeling of emptiness and dissatisfaction; in contrast, a meaningful life, or meaningful part of life, is necessarily at least somewhat rewarding or fulfilling. It is noteworthy, however, that meaningful experiences are not necessarily particularly happy. A trip to one's birthplace may well be meaningful; a visit to an amusement park is unlikely to be so.

If we step back, however, and ask ourselves, as observers, what lives strike us as especially meaningful, if we ask what sorts of lives exemplify meaningfulness, subjective criteria do not seem to be in the forefront. Who comes to mind? Perhaps, Gandhi, or Albert Schweitzer, or Mother Teresa; perhaps Einstein or Jonas Salk, Cézanne, or Manet, Beethoven, Charlie Parker, Tolstoy is an interesting case to which I shall return. Alternatively, we can look to our neighbours, our colleagues, our relatives—some of whom, it seems to me, live more meaningful lives than others. Some, indeed, of my acquaintances seem to me to live lives that are paradigms of meaning—right up there with the famous names on the earlier lists; while others (perhaps despite their modicum of fame) would score quite low on the meaningfulness scale. If those in the latter category feel a lack of meaning in their lives—well, they are right to feel it, and it is a step in the right direction that they notice that there is something about their lives that they should try to change.

What is it to live a meaningful life, then? What does meaningfulness in life amount to? It may be easier to make progress by focusing on what we want to avoid. In that spirit, let me offer some paradigms, not of meaningful, but of meaningless lives.

For me, the idea of a meaningless life is most clearly and effectively embodied in the image of a person who spends day after day, or night after night, in front of a television set, drinking beer and watching situation comedies. Not that I have anything against television or beer. Still the image, understood as an image of a person whose life is lived in hazy passivity, a life lived at a not unpleasant level of consciousness, but unconnected to anyone or anything, going nowhere, achieving nothing—is,

I submit, as strong an image of a meaningless life as there can be. Call this case The Blob.

If any life, any human life, is meaningless, the Blob's life is. But this doesn't mean that any meaningless life must be, in all important respects, like the Blob's. There are other paradigms that highlight by their absences other elements of meaningfulness. In contrast to the Blob's passivity, for example, we may imagine a life full of activity, but silly or decadent or useless activity. (And again, I have nothing against silly activity, but only against a life that is wholly occupied with it.) We may imagine, for example, one of the idle rich who flits about, fighting off boredom, moving from one amusement to another. She shops, she travels, she eats at expensive restaurants, she works out with her personal trainer.

Curiously, one might also take a very un-idle rich person to epitomize a meaningless life in a slightly different way. Consider, for example, the corporate executive who works twelve-hour, seven-day weeks, suffering great stress, for the sole purpose of the accumulation of personal wealth. Related to this perhaps is David Wiggins' example of the pig farmer who buys more land to grow more corn to feed more pigs to buy more land to grow more corn to feed more pigs.

These last three cases of the idle rich, the corporate executive and the pig farmer are in some ways very different, but they all share at least this feature: they can all be characterized as lives whose dominant activities seem pointless, useless, or empty. Classify these cases under the heading Useless.

A somewhat different and I think more controversial sort of case to consider involves someone who is engaged, even dedicated, to a project that is ultimately revealed as bankrupt, not because the person's values are shallow or misguided, but because the project fails. The person may go literally bankrupt: for example, a man may devote his life to creating and building up a company to hand over to his children, but the item his company manufactures is rendered obsolete by technology shortly before his planned retirement. Or consider a scientist whose life's work is rendered useless by the announcement of a medical breakthrough just weeks before his

own research would have yielded the same results. Perhaps more poignantly, imagine a woman whose life is centred around a relationship that turns out to be a fraud. Cases that fit this mould we may categorize under the heading Bankrupt. [...]

If the cases I have sketched capture our images of meaninglessness more or less accurately, they provide clues to what a positive case of a meaningful life must contain. In contrast to the Blob's passivity, a person who lives a meaningful life must be actively engaged. But, as the Useless cases teach us, it will not do to be engaged in just anything, for any reason or with any goal—one must be engaged in a project or projects that have some positive value, and in some way that is non-accidentally related to what gives them value. Finally, in order to avoid Bankruptcy, it seems necessary that one's activities be at least to some degree successful (though it may not be easy to determine what counts as the right kind or degree of success). Putting these criteria together, we get a proposal for what it is to live a meaningful life: *viz.*, a meaningful life is one that is actively and at least somewhat successfully engaged in a project (or projects) of positive value.

Several remarks are needed to qualify and refine this proposal. First, the use of the word "project" is not ideal: it is too suggestive of a finite, determinate task, something one takes on, and, if all goes well, completes. Among the things that come to mind as projects are certain kinds of hobbies or careers, or rather, specific tasks that fall within the sphere of such hobbies or careers: things that can be seen as accomplishments, like the producing of a proof or a poem or a pudding, the organizing of a union or a high school band. Although such activities are among the things that seem intuitively to contribute to the meaningfulness of people's lives, there are other forms of meaningfulness that are less directed, and less oriented to demonstrable achievement, and we should not let the use of the word "project" distort or deny the potential of these things to give meaningfulness to life. Relationships, in particular, seem at best awkwardly described as projects. Rarely does one deliberately take them on and, in some cases, one doesn't even have to work at them—one

may just have them and live, as it were, within them. Moreover, many of the activities that are naturally described as projects—coaching a school soccer team, planning a surprise party, reviewing an article for a journal—have the meaning they do for us only because of their place in the non-projectlike relationships in which we are enmeshed and with which we identify. In proposing that a meaningful life is a life actively engaged in projects, then, I mean to use "projects" in an unusually broad sense, to encompass not only goal-directed tasks but other sorts of ongoing activities and involvements as well.

Second, the suggestion that a meaningful life should be "actively engaged" in projects should be understood in a way that recognizes and embraces the connotations of "engagement." Although the idea that a meaningful life requires activity was introduced by contrast to the life of the ultra-passive Blob, we should note that meaning involves more than mere, literal activity. The alienated housewife, presumably, is active all the time—she buys groceries and fixes meals, cleans the house, does the laundry, chauffeurs the children from school to soccer to ballet, arranges doctors' appointment and babysitters. What makes her life insufficiently meaningful is that her heart, so to speak, isn't in these activities. She does not identify with what she is doing—she does not embrace her roles as wife, mother, and homemaker as expressive of who she is and wants to be. We may capture her alienated condition by saying that though she is active, she is not actively engaged. (She is, one might say, just going through the motions.) In characterizing a meaningful life, then, it is worth stressing that living such a life is not just a matter of having projects (broadly construed) and actively and somewhat successfully getting through them. The projects must engage the person whose life it is. Ideally, she would proudly and happily embrace them, as constituting at least part of what her life is about.

Finally, we must say more about the proposal's most blatantly problematic condition—*viz.*, that the project's engagement with which can contribute to a meaningful life must be projects "of positive value." The claim is that meaningful lives must be engaged in projects of positive value—but who is to

decide which projects have positive value, or even to guarantee that there is such a thing?

I would urge that we leave the phrase as unspecific as possible in all but one respect. We do not want to build a theory of positive value into our conception of meaningfulness. As a proposal that aims to capture what most people mean by a meaningful life, what we want is a concept that "tracks" whatever we think of as having positive value. This allows us to explain at least some divergent intuitions about meaningfulness in terms of divergent intuitions or beliefs about what has positive value, with the implication that if one is wrong about what has positive value, one will also be wrong about what contributes to a meaningful life. (Thus, a person who finds little to admire in sports—who finds ridiculous, for example, the sight of grown men trying to knock a little ball into a hole with a club, will find relatively little potential for meaning in the life of an avid golfer; a person who places little stock in esoteric intellectual pursuits will be puzzled by someone who strains to write, much less read, a lot of books on **supervenience**.)

The exception I would make to this otherwise maximally tolerant interpretation of the idea of positive value is that we exclude merely subjective value as a suitable interpretation of the phrase.

It will not do to allow that a meaningful life is a life involved in projects that seem to have positive value from the perspective of the one who lives it. Allowing this would have the effect of erasing the distinctiveness of our interest in meaningfulness; it would blur or remove the difference between an interest in living a meaningful life and an interest in living a life that feels or seems meaningful. That these interests are distinct, and that the former is not merely instrumental to the latter can be seen by reflecting on a certain way the wish or the need for meaning in one's life may make itself felt. What I have in mind is the possibility of a kind of epiphany, in which one wakes up—literally or figuratively—to the recognition that one's life to date has been meaningless. Such an experience would be nearly unintelligible if a lack of meaning were to be understood as a lack of a certain kind of subjective impression. One can hardly understand the idea of

waking up to the thought that one's life to date has seemed meaningless. To the contrary, it may be precisely because one did not realize the emptiness of one's projects or the shallowness of one's values until that moment that the experience I am imagining has the poignancy it does. It is the sort of experience that one might describe in terms of scales falling from one's eyes. And the yearning for meaningfulness, the impulse to do something about it, will not be satisfied (though it may be eliminated) by putting the scales back on, so to speak. If one suspects that the life one has been living is meaningless, one will not bring meaning to it by getting therapy or taking a pill that, without changing one's life in any other way, makes one believe that one's life has meaning.

To care that one's life is meaningful, then, is, according to my proposal, to care that one's life is actively and at least somewhat successfully engaged in projects (understanding this term broadly) that not just seem to have positive value, but that really do have it. To care that one's life be meaningful, in other words, is in part to care that what one does with one's life is, to pardon the expression, at least somewhat objectively good. We should be careful, however, not to equate objective goodness with moral goodness, at least not if we understand moral value as essentially involving benefitting or honouring humanity. The concern for meaning in one's life does not seem to be the same as the concern for moral worth, nor do our judgments about what sorts of lives are meaningful seem to track judgments of moral character or accomplishment.

To be sure, some of the paradigms of meaningful lives are lives of great moral virtue or accomplishment—I mentioned Gandhi and Mother Teresa, for example. Others, however, are not. Consider **Gauguin**, **Wittgenstein**, **Tchaikovsky**—morally unsavoury figures all, whose lives nonetheless seem chock full of meaning. If one thinks that even they deserve moral credit, for their achievements made the world a better place, consider instead Olympic athletes and world chess champions, whose accomplishments leave nothing behind but their world records. Even more important, consider the artists, scholars, musicians, athletes of our more ordinary sort. For us, too, the activities of

artistic creation and research, the development of our skills and our understanding of the world give meaning to our lives—but they do not give moral value to them.

It seems then that meaning in life may not be especially moral, and that indeed lives can be richly meaningful even if they are, on the whole, judged to be immoral. Conversely, that one's life is at least moderately moral, that it is lived, as it were, above reproach, is no assurance of its being moderately meaningful. The alienated housewife, for example, may be in no way subject to moral criticism. (And it is debatable whether even the Blob deserves specifically moral censure.)

That people do want meaning in their lives, I take it, is an observable, empirical fact. We have already noted the evidence of self-help manuals, and therapy groups. What I have offered so far is an analysis of what that desire or concern amounts to. I want now to turn to the question of whether the desire is one that it is good that people have, whether, that is, there is some positive reason why they *should* want this.

3. Why Seek a Meaningful Life?

At a minimum, we may acknowledge that it is at least not bad to want meaning in one's life. There is, after all, no harm in it. Since people do want this, and since there are no moral objections to it, we should recognize the concern for meaning as a legitimate concern, at least in the weak sense that people should be allowed to pursue it. Indeed, insofar as meaningfulness in one's life is a significant factor in a life's overall well-being, we should do more than merely allow its pursuit: we should positively try to increase opportunities for people to live lives of meaning.

Most of us, however, seem to have a stronger positive attitude toward the value of meaningfulness than this minimum concession admits. We do not think it is merely all right for people to want meaning in their lives—as it is all right for people to like country music, or to take an interest in figure skating. We think people positively ought to care that their lives be meaningful. It is disturbing, or at least

regrettable, to find someone who doesn't care about this. Yet this positive assessment ought to strike us, at least initially, as somewhat mysterious. What is the good, after all, of living a meaningful life, and to whom?

Since a meaningful life is not necessarily a *morally* better life than a meaningless one (the Olympic athlete may do no more good nor harm than the idly rich socialite), it is not necessarily better *for the world* that people try to live or even succeed in living meaningful lives. Neither is a meaningful life assured of being an especially happy one, however. Many of the things that give meaning to our lives (relationships to loved ones, aspirations to achieve) make us vulnerable to pain, disappointment and stress. From the inside, the Blob's hazy passivity may be preferable to the experience of the tortured artist or political crusader. By conventional standards, therefore, it is not clear that caring about or even succeeding in living a meaningful life is better *for the person herself*.

Yet, as I have already mentioned, those of us who do care that our lives be meaningful tend to think that it is a positively good thing that we do. We not only want to live meaningful lives, we want to want this—we approve of this desire, and think it is better for others if they have this desire, too. If, for example, you see a person you care about conducting her life in a way that you find devoid of worth—she is addicted to drugs, perhaps, or just to television, or she is overly enthusiastic in her career as a corporate lawyer—you are apt to encourage her to change, or at least hope that she will find a new direction on her own. Your most prominent worry may well be that she is heading for a fall. You fear that at some point she will wake up to the fact that she has been wasting or misdirecting her life, a point that may come too late for easy remedy and will, in any case, involve a lot of pain and self-criticism. But the fear that she will wake up to the fact that she has been wasting her life (and have difficulty turning her life around) may not be as terrible as the fear that she won't wake up to it. If you came to feel secure that no painful moment of awakening would ever come because your friend (or sister or daughter) simply

does not care whether her life is meaningful, you might well think that this situation is not better but worse. We seem to think there is something regrettable about a person living a meaningless life, even if the person herself does not mind that she is. We seem to think she *should* want meaning in her life, even if she doesn't realize it.

What, though, is the status of this "should," the nature or source of the regret? The mystery that I earlier suggested we should feel about our value in meaningfulness is reflected in the uneasy location of this judgment. If my own reaction to the woman who doesn't care whether her life is meaningful is typical, the thought that she should, or ought to care is closer to a prudential judgment than it is to a moral one. (If there is a moral objection to a person who lives a meaningless life and is content with that, it is not, in my opinion, a very strong one. The Blob, after all, is not hurting anyone, nor is the idle rich jet-setter. She may, for example, give money to environmental causes to offset the damage she is doing in her SUV, and write generous cheques to Oxfam and UNICEF on a regular basis.) The thought that it is too bad if a person does not live a meaningful life (even if she doesn't mind) seems rather to be the thought that it is too bad *for her*.

The closest analogue to this thought in the history of ethics of which I am aware is Aristotle's conception of *eudaimonia*. His conception of the virtuous life as the happiest life is offered as a conclusion of an enlightened self-interest. According to standard conceptions of self-interest, however (either **hedonistic** or preference-based), it is not obvious why this should be so, and, unfortunately, Aristotle himself does not address the question explicitly. Rather, he seems to think that if you do not just see that the virtuous life, in which one aims for and achieves what is "fine," is a better, more desirable life for yourself, that just shows that you were not well brought up, and in that case, there is no point trying to educate you.

Our question, the question of whether and what kind of reason there is for a person to strive for a meaningful life, is not quite the same as the question of whether and what kind of reason there is to

aspire to virtue,—though, when one is careful to interpret "virtue" in the broad and not specifically moral way that Aristotle uses the term, it is closer than it might seem. Still, as I say, Aristotle does not really address the question, and so, though I take my line of thought to be Aristotelian in spirit, a scholarly study of Aristotle's texts is not likely to be an efficient way of finding an answer to the question ourselves.

What reason is there, then, if any, for a person to want to live a meaningful life? I have said that we seem to think it would be better for her, that it is, at least roughly, in her self-interest. At the same time, the thought that she should care about meaning seems to depend on claims from outside herself. Even if there are no desires latent in her psychology which meaningfulness would satisfy, we seem to think, there is reason why she should have such desires. She seems to be making some kind of mistake.

If my analysis of what is involved in living a meaningful life is right, then the question of why one should care about living a meaningful life is equivalent to the question of why one should care that one's life be actively and somewhat successfully engaged in projects of positive value. The source of perplexity seems, in particular, to be about the reason to care that one's projects be positively valuable. As long as you are engaged by your activities, and they make you happy, why should one care that one's activities be objectively worthwhile?

The answer, I believe, is that to devote one's life entirely to activities whose value is merely subjective, to devote oneself to activities whose sole justification is that it is good for you, is, in a sense I shall try to explain, practically **solipsistic**. It flies in the face of one's status as, if you will, a tiny speck in a vast universe, a universe with countless perspectives of equal status with one's own, from which one's life might be assessed. Living a life that is engaged with and so at least partially focused on projects whose value has a non-subjective source is a way of acknowledging one's non-privileged position. It harmonizes, in a way that a purely egocentric life does not, with the fact that one is not the centre of the universe.

The basic idea is this: The recognition of one's place in the universe, of one's smallness, one might say, or

one's insignificance, and of the independent existence of the universe in which one is a part involves, among other things, the recognition of "the mereness" of one's subjective point of view. To think of one's place in the universe is to recognize the possibility of a perspective, of infinitely many perspectives, really, from which one's life is merely gratuitous; it is to recognize the possibility of a perspective, or rather of infinitely many perspectives, that are indifferent to whether one exists at all, and so to whether one is happy or sad, satisfied or unsatisfied, fulfilled or unfulfilled.

In the face of this recognition, a life that is directed solely to its subject's own fulfillment, or, to its mere survival or towards the pursuit of goals that are grounded in nothing but the subject's own psychology, appears either solipsistic or silly.

A person who lives a largely egocentric life—who devotes, in other words, lots of energy and attention and care toward himself, who occupies himself more specifically with satisfying and gratifying himself, expresses and reveals a belief that his happiness matters. Even if it doesn't express the view that his happiness matters objectively, it at least expresses the idea that it matters to him. To be solely devoted to his own gratification, then, would express and reveal the fact that his happiness is *all* that matters, at least all that matters to him. If, however, one accepts a framework that recognizes distinctions in non-subjective value, (and if one believes, as seems only reasonable, that what has non-subjective value has no special concentration in or connection to oneself) this attitude seems hard to justify.

To accept that framework is, after all, to accept the view that some things are better than others. To me, it makes sense partially to understand this literally: Some things, it seems to me, are better than others: people, for example, are better than rocks or mosquitoes, and a Vermeer painting is better than the scraps on my compost heap. What is essential, though, is that accepting a framework that recognizes distinctions in non-subjective value involves seeing the world as value-filled, as containing with it distinctions of better and worse, of more and less worthwhile, if not of better and worse objects per se, then of better and worse features of the world, or activities, or

opportunities to be realized. Against this background, a life solely devoted to one's own gratification or to the satisfaction of one's whims seems gratuitous and hard to defend. For, as I have said, to live such a life expresses the view that one's happiness is all that matters, at least to oneself. But why should this be the only thing that matters, when there is so much else worth caring about? [...]

I am suggesting that we can have a reason to do something or to care about something that is grounded not in our own psychologies, nor specifically in our own desires, but in a fact about the world. The fact in question in this case is the fact that we are, each of us, specks in a vast and value-filled universe, and that as such we have no privileged position as a source of or possessor of objective value. To devote oneself wholly to one's own satisfaction seems to me to fly in the face of this truth, to act "as if" one is the only thing that matters, or perhaps, more, that one's own psychology is the only source of (determining) what matters. By focusing one's attention and one's energies at least in part on things, activities, aspects of the world that have value independent of you, you implicitly acknowledge your place and your status in the world. Your behaviour, and your practical stance, is thus more in accord with the facts.

Admittedly, this is not the sort of reason that one must accept on pain of inconsistency or any other failure of logic. Just as a person may simply not care whether her life is meaningful, so she may also simply not care whether her life is in accord with, or harmonizes with the facts. (It is one thing to say we should live in accord with the facts of physics, geography, and the other sciences. Living in accordance with these facts has evident instrumental value—it helps us get around in the world. But living in a way that practically acknowledges, or harmonizes with the fact that we are tiny specks in a value-filled world will not make our lives go better that way.) Such a person cannot be accused in any strict sense of irrationality. Like non-instrumental reasons to be moral, the reason to care about living a worthwhile life is not one that narrow rationality requires one to accept. At the same time, it seems appropriate to

characterize my suggestion [...] as one that appeals to reason in a broader sense. For my suggestion is that an interest in living a meaningful life is an appropriate response to a fundamental truth, and that failure to have such a concern constitutes a failure to acknowledge that truth.

4. Meaning of Life

As we have already seen, the truth to which I am proposing a meaningful life provides a response is the truth that we are, each of us, tiny specks in a vast and value-filled universe. [...] It opposes what children and many adults may have a tendency to assume—namely, that they are the centre of the universe, either the possessor or the source of all value. [...] My proposal is that an appropriate response to our status as specks in a vast universe is a concern and aspiration to have one's life wrapped up with projects of positive value.

Perhaps, however, I have not made it clear why this is an appropriate response. The question may seem especially pressing because the thought that we are tiny specks in a vast universe, and the sense that it calls for or demands a response has, in the past, tended to move philosophers in a different direction. Specifically, the thought that we are tiny specks in a vast universe was in the past closely associated with that murky and ponderous question to which I referred at the beginning of my [paper]—the question of The Meaning of Life. The thought that we are tiny specks in a vast universe has indeed often evoked that question, and, to those who either do not believe in or do not want to rest their answers in the existence of a benevolent God, it has more or less immediately seemed also to indicate an answer. Considering their answer to the question of the Meaning of Life and contrasting it with my response to the fact of our smallness, may clarify the substance of my proposal.

The train of thought I have in mind is one that has, with variations, been expressed by many distinguished philosophers, including Camus [and] Tolstoy [...]. For them, the recognition of our place in the universe—our smallness, or our speckness, if

you will—seems to warrant the conclusion not only that there is no meaning to life as such but also that each individual life is necessarily absurd.

On the view of these philosophers, a life can be meaningful only if it can mean something to someone, and not just to *someone*, but to someone other than oneself and indeed someone of more intrinsic or ultimate value than oneself. Of course, anyone can live in such a way as to make her life meaningful to *someone* other than herself. She can maintain her relationship with parents and siblings, establish friendships with neighbours and colleagues. She can fall in love. If all else fails, she can have a child who will love her, or two children, or six. She can open up an entire clinic for God's sake. But if a life that is devoted solely to yourself, a life that is good to no one other than yourself lacks meaning, these philosophers not implausibly think, so will a life that is devoted to any other poor creature, for he or she will have no more objective importance than you have, and so will be no more fit a stopping place by which to ground the claim of meaningfulness than you. Nor, according to this train of thought, will it help to expand your circle, to be of use or to have an effect on a larger segment of humankind. If each life is individually lacking in meaning, then the collective is meaningless as well. If each life has but an infinitesimal amount of value, then although one's meaning will increase in proportion to one's effect, the total quantity of meaning relative to the cosmos will remain so small as to make the effort pathetic.

From the perspective of these philosophers, if there is no God, then human life, each human life, must be objectively meaningless, because if there is no God, there is no appropriate being *for whom* we could have meaning.

From this perspective, my suggestion that the living of a worthwhile life constitutes a response to a recognition of our place in the universe might seem ridiculously nearsighted, as if, having acknowledged the mereness of my own subjectivity, I then failed to acknowledge the equal mereness of the subjectivity of others. But I think this misunderstands the point in my proposal of living a life that realizes non-subjective value, a misunderstanding that derives from

too narrow a view about what an appropriate and satisfactory response to the fact of our place in the universe must be.

The philosophers I have been speaking about—we can call them the pessimists—take the fundamental lesson to be learned from the contemplation of our place in the universe to be that we are cosmically insignificant, a fact that clashes with our desire to be very significant indeed. If God existed, such philosophers might note, we would have a chance at being significant. For God himself, is presumably very significant and so we could be significant by being or by making ourselves significant to Him. In the absence of a God, however, it appears that we can only be significant to each other, to beings, that is, as pathetically small as ourselves. We want to be important, but we cannot be important, and so our lives are absurd.

The pessimists are right about the futility of trying to make ourselves important. Insofar as contemplation of the cosmos makes us aware of our smallness, whether as individuals or as a species, we simply must accept it and come to terms with it. Some people do undoubtedly get very upset, even despondent when they start to think about their cosmic insignificance. They want to be important, to have an impact on the world, to make a mark that will last forever. When they realize that they cannot achieve this, they are very disappointed. The only advice one can give to such people is: Get Over It.

Rather than fight the fact of our insignificance, however, and of the mereness of our subjectivity, my proposal is that we live in a way that acknowledges the fact, or, at any rate, that harmonizes with it. Living in a way that is significantly focused on, engaged with, and concerned to promote or realize value whose source comes from outside of oneself, does seem to harmonize with this, whereas living purely egocentrically does not. Living lives that attain or realize some non-subjective value may not make us meaningful, much less important, to anyone other than ourselves, but it will give us something to say, to think, in response to the recognition of perspectives that we ourselves imaginatively adopt that are indifferent to our existence and to our well-being.

At the beginning of this paper, I raised the question of how the meaning of life—or the absence of such meaning—was related to the meaningfulness of particular lives. As I might have put it, does it really make sense to think that there can be meaningful lives in a meaningless world? In light of this discussion, we can see how the answer to that question might be “yes” while still holding on to the idea that the similar wording of the two phrases is not merely coincidental.

If I am right about what is involved in living a meaningful life—if, that is, living a meaningful life is a matter of at least partly successful engagement in projects of positive value—then the possibility of living meaningful lives despite the absence of an overall meaning to life can be seen to depend on the fact that distinctions of value (that is, of objective value) do not rely on the existence of God or of any overarching purpose to the human race as a whole. Whether or not God exists, the fact remains that some objects, activities and ideas are better than others. Whether or not God exists, some ways of living are more worthwhile than others. Some activities are a waste of time.

People are sometimes tempted to think that if God doesn't exist, then nothing matters. They are tempted to think that if we will all die, and eventually all traces of our existence will fade from all consciousness, there is no point to doing anything; nothing makes any difference. Tolstoy evidently thought this sometimes, and gave eloquent voice to that view. But the reasoning is ridiculous. If one activity is worthwhile and another is a waste, then one has reason to prefer the former, even if there is no God to look down on us and approve. More generally, we seem to have reason to engage ourselves with projects of value whether God exists and gives life a purpose or not.

Putting things this way, however, fails to explain why we use the language of meaning to describe lives engaged in activities of worth. Putting things this way there seems to be no connection at all between the question of whether there is a meaning to life and the question of whether individual lives can be meaningful. I believe, however, that there is a connection, that shows itself, or perhaps that consists

in the fact that the wish for both kinds of meaning are evoked by the same thought, and that, perhaps, either kind of meaning would be an appropriate and satisfying response to that thought. The thought in question is the thought (the true thought) that we are tiny specks in a vast universe. It is a thought that is apt to be upsetting when it first hits you—at least in part because, looking back from that position, it may seem that one had until then lived “as if” something opposite were true. One had lived perhaps until then as if one were the centre of the universe, the sole possessor or source of all value. One had all along assumed one had a special and very important place in the world, and now one’s assumption is undermined. One can see how, in this context, one might wish for a meaning to life. For if there were a meaning—a purpose, that is, to human existence that can be presumed to be of great importance, then, by playing a role, by contributing to that purpose, one can recover some of the significance one thought one’s life had. Like the pessimistic philosophers I talked about a few minutes ago, I doubt that that path is open to us. But there seems another way one can respond to the thought, or to the recognition of our relatively insignificant place in the universe, that is more promising, and that can, and sometimes does, provide a different kind of comfort. If one lived one’s life, prior to the recognition of our smallness, as if one was the centre of the universe, the appropriate response to that recognition is simply *to stop living that way*. If one turns one’s attention to other parts of the universe—even to other specks like oneself—in a way that appreciates and engages with the values or valuable objects that come from outside oneself, then one corrects one’s practical stance. If, in addition, one is partly successful in producing, preserving, or promoting value—if one does some good, or realizes value, then one has something to say, or to think in

response to the worry that one’s life has no point.

Only if some suggestion like mine is right can we make sense of the intuitions about meaningfulness to which I called attention in the earlier part of this paper. According to those intuitions the difference between a meaningful and a meaningless life is not a difference between a life that does a lot of good, and a life that does a little. (Nor is it a difference between a life that makes a big splash and one that, so to speak, sprays only a few drops.) It is rather a difference between a life that does good or is good or realizes value and a life that is essentially a waste. According to these intuitions, there is as sharp a contrast between the Blob and a life devoted to the care of a single needy individual as there is between the Blob and someone who manages to change the world for the better on a grand scale. Indeed, there may be an equally sharp contrast between the Blob and the monk of a contemplative order whose existence confers no benefit or change on anyone else’s life at all. Ironically, along this dimension, Tolstoy fares exceptionally well.

Thus it seems to me that even if there is no meaning to life, even if, that is, life as a whole has no purpose, no direction, no point, that is no reason to doubt the possibility of finding and making meaning in life—that is no reason, in other words, to doubt the possibility of people living meaningful lives. In coming to terms with our place and our status in the universe, it is natural and appropriate that people should want to explore the possibility of both types of meaning. Even if philosophers have nothing new or encouraging to say about the possibility of meaning of the first sort, there may be some point to elaborating the different meanings of the idea of finding meaning in life, and in pointing out the different forms that coming to terms with the human condition can take.

Work Cited

- Wiggins, D. “Truth, Invention, and the Meaning of Life.” *Proceedings of the British Academy* 62 (1976): 331–78.

Study Questions

1. Describe how Wolf begins with three paradigms of meaningless lives as a method of arriving at her definition of what constitutes a meaningful life. Based on Wolf's definition, assess the meaningfulness of your own life.
2. Provide an example of a life that is meaningful in a subjective sense. Provide an example of a life that is meaningful in an objective sense. Evaluate Wolf's argument for claiming that we should prefer lives that are objectively, not merely subjectively, meaningful.
3. Wolf cautions us not to equate meaningful lives with moral or happy lives. Provide examples of someone living a meaningful but not moral life, a meaningful but not happy life, a moral but not meaningful life, and a happy but not meaningful life.
4. Why do the philosophers Wolf calls "pessimists" believe that the question "What is the meaning of life?" and the question "What makes a meaningful life?" depend on god's existence? Wolf finds these questions not as closely connected as other philosophers assume, but not unconnected either. Explain why.

Bernard Williams

Bernard Williams (1929–2003) is considered one of the most influential British philosophers of the twentieth century. Williams studied classics at Balliol College, Oxford, and spent his National Service flying Spitfires in Canada. Williams' career was divided between Oxford, Cambridge, and Berkeley. Williams is best known for his work on the philosophy of action and theories of well-being.