In this *issue...*

Contents

[Editorial]

Welcome to the fourth issue of the Richmond Journal of Philosophy. The first paper looks at the question of free will. Galen Strawson sets out the central aspects of the enduring debate on the relationship between our will and a world apparently governed by deterministic laws. We then enter into a discussion of realism in mathematics by John Burgess, a topic which casts light on the very question of what it is to hold that something we refer to is real. Chris Hughes' paper examines the Christian doctrine of the Trinity: how can God be at once a single Being and three persons – the Father, Son and Holy Spirit? After God we turn to Nietzsche and Clare Carlisle's treatment of his attitude to the notion of truth. Next Paul Sheehy considers reduction and methodological individualism in the social sciences, while the topic of our final paper by Joaquim Siles i Borras is the notion of evil in Kant's moral philosophy. We thank all of our contributors.

The publication of this issue marks the first anniversary of the Journal. We should like to take this opportunity to extend our thanks to all of our contributors, subscribers and to Richmond upon Thames College. We hope that our subscribers have found the journal to be a source of serious and accessible philosophy.

Purpose of the Journal

The motivation for and ambition of the journal is to provide serious philosophy for students who are at an early stage in their philosophical studies. The style and content of the papers will be accessible to students who have yet to become hardened to the more technical and specialised journals of professional philosophy.

What do we mean by 'serious' philosophy? First, the content of the journal is not constrained by a remit to appeal to or reach the interested general public. Whilst the papers must speak to the needs of students who are relatively inexperienced in philosophy, they presuppose that their audience is actively engaged in philosophy. Second, the content is serious in its focus on the central areas of philosophy. The big or traditional questions of metaphysics, epistemology, and ethics will provide the journal's centre of gravity. The third way in which the philosophy is serious is through the scope, variety and depth of analysis that can be achieved by the accumulation of papers over time. Moreover, each paper is not simply an introduction to one of the main topics on A-level, IB or degree courses. Such papers will indeed have a role in the journal, but they will not be the only kind. Our contributors will be offering original papers based on their own research. The journal will be a forum for the kind of critical engagement and debate that characterise the practice of philosophy. The fourth way in which the philosophy is serious is in the contributors themselves. The vast bulk of the papers will be written by professional philosophers engaged in both research and teaching.

About the [Editorial] Board

Stephen Grant is a full-time lecturer in philosophy at Richmond upon Thames College. He has also taught at King's College London where he is completing his doctorate on the emotions. His main interests are in the emotions, ethics and political philosophy. He has published on the ontological argument.

Dr Paul Sheehy teaches philosophy at Richmond upon Thames College and King's College London. His main areas of interest are in metaphysics, political and moral philosophy and the philosophy of the social sciences. His doctoral thesis was on the ontological and moral status of social groups, and he has published papers on social groups, voting and explanation and realism.

Paul Sperring is head of the philosophy department at Richmond upon Thames College and an A-level examiner in philosophy. He completed his undergraduate and masters studies at Warwick University, studying both analytic and continental philosophy. He has recently been based in King's College London philosophy department as their Teacher Fellow.
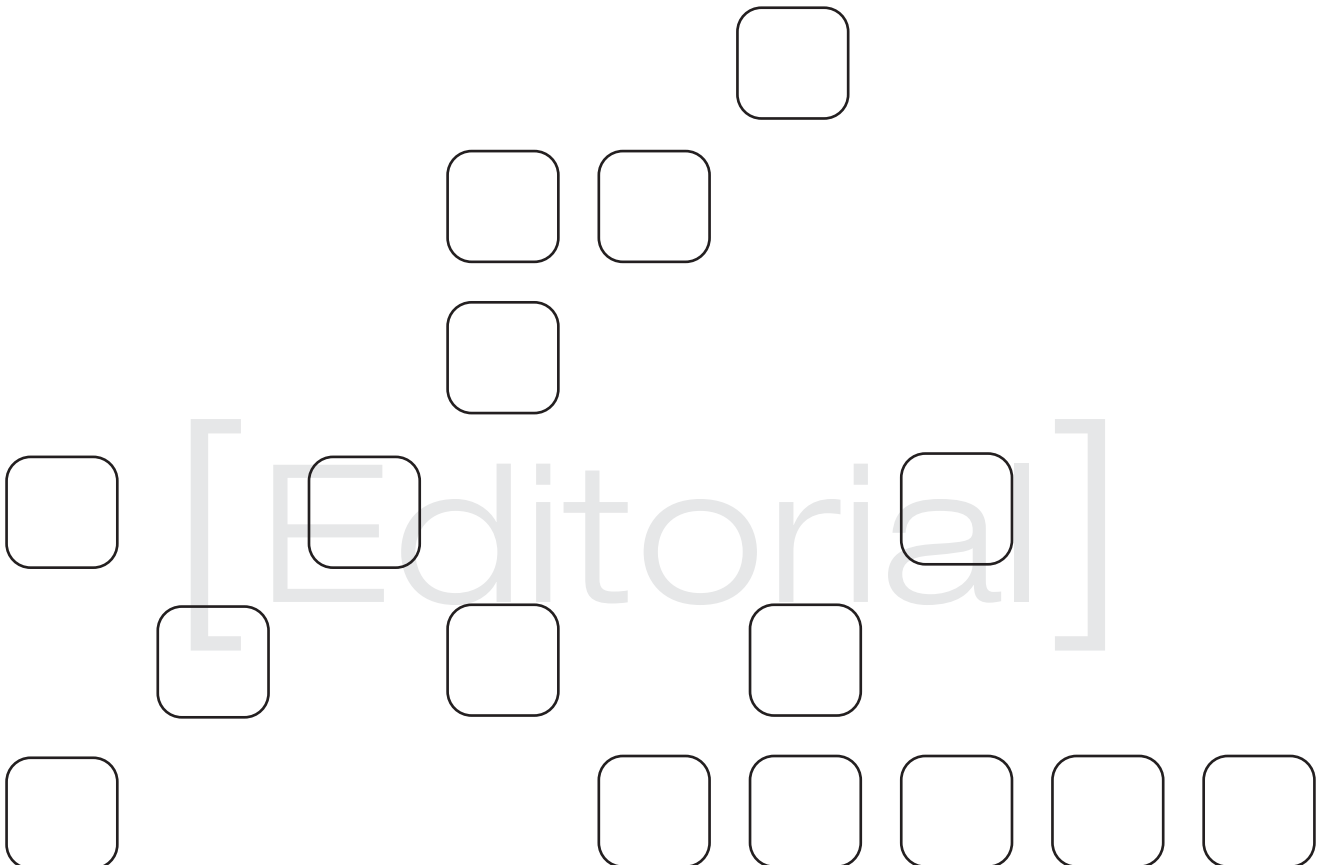
Galen Strawson

[Free Will] on

Are we free agents? Can we be morally responsible for what we do? Some philosophers answer Yes and Yes (we are fully free, and wholly morally responsible for what we do). Others answer Yes and No (certainly we are free agents – but we cannot be said to be *ultimately* responsible for what we do). A third group answers No and No (we are not free agents at all; *a fortiori* we cannot be morally responsible). A strange minority answers No and Yes (we can be morally responsible for what we do, even though we are not free agents!). This last view is rare, but it has a kind of existentialist panache, and appears to be embraced by Wintergreen in Joseph Heller's novel *Closing Time* (1994), as well as by some Protestants.

Who's right? Suppose tomorrow's a holiday, and that you're wondering what to do. You can climb a mountain or read Lao Tzu. You can restring your mandolin or go to the zoo. At the moment you're reading about free will. You're free to go on reading or stop now. You have started on this sentence, but you don't have to............finish it. Right now, as so often in life, you have a number of options. Nothing forces your hand. So surely you're entirely free to choose what to do, and responsible for what you do?

This is what the 'Compatibilists' think. They answer Yes and Yes. Their name derives from their claim that free will

is entirely compatible with determinism (the view that everything that happens in the universe is necessitated by what has already gone before, in such a way that nothing can happen otherwise than it does). Free will, they think, is just a matter of not being constrained or compelled in certain ways that have nothing to do with whether determinism is true or false. 'Consider yourself at this moment', they say. 'No one's holding a gun to your head. You're not being threatened or manhandled. You're not drugged, or in chains, or subject to a psychological compulsion like kleptomania, or a post-hypnotic command. So you're wholly free. This is what being a free agent is. It's wholly irrelevant that your character is determined, if indeed it is.

'And although things like guns and chains, threats to the life of your children, psychological obsessions, and so on, are standardly counted as constraints that can limit freedom and responsibility, there is another and more fundamental sense in which you are fully free in any situation in which you can choose or act in any way at all – in any situation in which you are not panicked, or literally forced to do what you do. Consider pilots of hijacked aeroplanes. They usually stay calm. They choose to comply with the hijackers' demands. They act responsibly, as we naturally say. They are able to do other than they do, but

they choose not to. They do what they most want to do, all things considered, in the circumstances in which they find themselves – and all circumstances limit one's options in some way. Some circumstances limit one's options much more drastically than others, but it doesn't follow that one isn't free to choose in those circumstances. Only literal compulsion, panic, or uncontrollable impulse really removes one's freedom to choose, and to (try to) do what one most wants to do, given one's character or personality. Even when one's finger is being forced down on the button, one can still act freely in resisting the pressure, in cursing one's oppressor, and in many other ways.'

So most of us are wholly free to choose and act throughout our waking lives, according to the Compatibilists. We're free to choose between the options we perceive to be open to us. One has options even when one is in chains, or falling through space. Even if one is completely paralysed, one is still free in so far as one is free to choose to think about one thing rather than another. There is, as Sartre observed, a sense in which we're condemned to freedom: we're not free not to be free.

One may well not be able to do everything one wants – one may want to fly unassisted, vapourize every gun in the United States by an act of thought, or house all those who sleep

on the streets of Calcutta by the end of the month - but few have supposed that free will is a matter of being able to do everything one wants. It's a possible view, but according to the Compatibilists free will is simply a matter of being unconstrained in such a way that one has genuine options and opportunities for action, and is able to choose between them according to what one wants or thinks best. It just doesn't matter if one's character, personality, preferences, and general motivational set are entirely determined by things for which one is in no way responsible - by one's genetic inheritance, upbringing, historical situation, chance encounters, and so on.

Even dogs count as free agents, on this view. So Compatibilists have to explain what distinguishes us from dogs - since we don't think that dogs are free in the way we are. Some say it's our capacity to act for reasons that we explicitly take to be moral reasons. Many say it's our capacity for explicitly self-conscious thought. Not because self-consciousness liberates anyone from determinism; if determinism is true, one is determined to have whatever self-conscious thoughts one has, whatever their complexity.

The idea is that self-consciousness makes it possible for one to be explicitly aware of oneself as facing choices and engaging in processes of reasoning about what to do, and thereby constitutes one as a radically free agent in a way unavailable to any unself-conscious agent. They think one's self-conscious deliberative presence in the situation of choice simply trumps the fact - if it is a fact - that one is in the final analysis wholly constituted as the sort of person one is by factors for which one isn't in any way ultimately responsible.

The Compatibilists, then, say Yes and Yes, and those who want to say this are well advised to follow them, for determinism is unfalsifiable, and may be true. (In the end, contemporary physics gives us no more reason to suppose that determinism is false than to suppose that it is true.)

Many, however, think that Compatibilism doesn't even touch the real problem of free will. For what is it to define freedom in such a way that it is compatible with determinism? It's to define it in such a way that an agent can be a free agent even if all its actions throughout its life are determined to happen as they do by events that have taken place before it is born, so that there is a clear sense in which it could not at any point in its life have done otherwise than it did. And this doesn't look like genuine free will or moral responsibility. How can one possibly be truly or ultimately morally responsible for what one does, if everything one does is ultimately a deterministic outcome of events for whose occurrence one is in no way responsible? This is the *Incompatibilists'* view.

The Incompatibilists divide into two groups: there are the *Libertarians*, on the one hand, and the *No-Freedom Theorists* or *Pessimists*, on the other. The Libertarians are up-beat. They say Yes and Yes, like the Compatibilists, but think the Compatibilists' account of freedom can be improved. They hold (1) that we do have free will, (2) that free will is not compatible with determinism, and (3) that determinism is therefore false. But then they face an extremely difficult task: they have to show how indeterminism (the falsity of determinism) can help with free will, and in particular with moral responsibility.
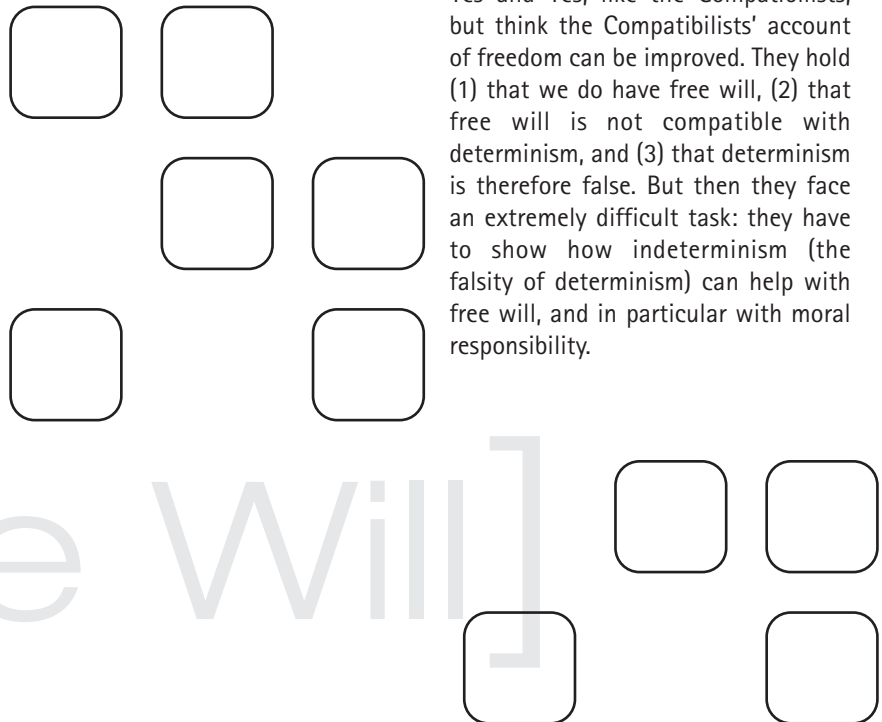
The Pessimists don't think this can be shown. They agree that free will is not compatible with determinism, but deny that indeterminism can help. They think that free will, of the sort that is necessary for genuine moral responsibility, is provably impossible. They say No and No.

They begin by granting what everyone must. They grant that there's a clear, important, compatibilist sense in which we can be free agents (we can be free, when unconstrained, to choose and to do what we want or think best, given how we are). But they insist that this isn't enough: it doesn't give us what we want in the way of free will. Nor does it give us what we believe we have. But (they continue) it is not as if the Compatibilists have missed something. The truth is that nothing can give us what we think we want, and ordinarily think we have. We cannot be morally responsible, in the absolute, buck-stopping way in which we often unreflectively think we are. We cannot have 'strong' free will of the kind that we would need to have, in order to be morally responsible in this way.

One way of setting out the Pessimists' argument is as follows: (1) When you act, you do what you do, in the situation in which you find yourself, because of the way you are.

But then (2) To be truly or ultimately morally responsible for what you do, you must be truly or ultimately responsible for the way you are, at least in certain crucial mental respects. (Obviously you don't have to be responsible for your height, age, sex, and so on.)

But (3) You can't be ultimately responsible for the way you are in any respect at all, so you can't be ultimately responsible for what you do.

For (4) To be ultimately responsible for the way you are, you must have somehow intentionally brought it about that you are the way you are.

And the problem is then this. Suppose (5) You have somehow intentionally brought it about that you are the way you now are, in certain mental respects: suppose you have brought it about that you have a certain mental nature Z, in such a way that you can be said to be ultimately responsible for Z.

For this to be true (6) You must already have had a certain mental nature Y, in the light of which you brought it about that you now have Z. (If you didn't already have a mental nature then you didn't have any intentions or preferences, and can't be responsible for the way you now are, even if you have changed.) But then (7) For it to be true that you are ultimately responsible for how you now are, you must be ultimately responsible for having had that nature, Y, in the light of which you brought it about that you now have Z.

So (8) You must have brought it about that you had Y.

But then (9) You must have existed already with a prior nature, X, in the light of which you brought it about that you had Y, in the light of which you brought it about that you now have Z.

And so on. Here, one is setting off on a potentially infinite regress. In order for one to be truly or ultimately responsible for how one is in such a way that one can be truly responsible for what one does, something impossible has to be true: there has to be, and cannot be, a starting point in the series of acts of bringing it about that one has a certain nature; a starting point that constitutes an act of ultimate self-origination.

There's a more concise way of putting the point: in order to be ultimately responsible, one would have to be *causa sui* - the ultimate cause or origin of oneself, or at least of some crucial part of one's mental nature. But nothing can be ultimately *causa sui* in any respect at all. Even if the property of being *causa sui* is allowed

to belong unintelligibly to God, it cannot plausibly be supposed to be possessed by ordinary finite human beings. 'The *causa sui* is the best self-contradiction that has been conceived so far', as Nietzsche remarked in *Beyond Good and Evil* in 1886:

> it is a sort of rape and perversion of logic. But the extravagant pride of man has managed to entangle itself profoundly and frightfully with just this nonsense. The desire for 'freedom of the will' in the superlative metaphysical sense, which still holds sway, unfortunately, in the minds of the half-educated; the desire to bear the entire and ultimate responsibility for one's actions oneself, and to absolve God, the world, ancestors, chance, and society involves nothing less than to be precisely this *causa sui* and, with more than Baron Muenchhausen's audacity, to pull oneself up into existence by the hair, out of the swamps of nothingness. . . .

In fact, nearly all of those who believe in strong free will do so without any conscious thought that it requires ultimate self-origination. But self-origination is the only thing that could actually ground the kind of strong free will that is regularly believed in.

The Pessimists' argument may seem contrived, but essentially the same argument can be given in a more natural form as follows. (A) One is the way one is, initially, as a result of heredity and early experience. (B) These are clearly things for which one cannot be held to be in any way responsible (this might not be true if there were reincarnation, but this would just shift the problem backwards). (C) One cannot at any later stage of one's life hope to accede to ultimate responsibility for the way one is by trying to change the way one already is as a result of heredity and experience. For one may well try to change oneself, but (D) both the particular way in which one is moved to try to change oneself, and the degree of one's success in one's attempt at change, will be determined by how one already is as a result of heredity and experience. And (E) any further changes that one can bring about only after one has brought about certain initial changes will in turn be determined, via the initial changes, by heredity and previous experience. (F) This may not be the whole story, for it may be that some changes in the way one is are traceable to the influence of indeterministic or random factors. But (G) it is absurd to suppose that indeterministic or random factors, for which one is *ex hypothesi* in no way responsible, can in themselves contribute to one's being truly or ultimately responsible for how one is.

The claim, then, is not that people cannot change the way they are. They can, in certain respects (which tend to be exaggerated by North Americans and underestimated, perhaps, by members of other cultures). The claim is only that people cannot be supposed to change themselves in such a way as to be or become ultimately responsible for the way they are, and hence for their actions. One can put the point by saying that in the final analysis the way you are is, in every last detail, a matter of luck - good or bad.

Philosophers will ask what exactly this 'ultimate' responsibility is supposed to be. They will suggest that it doesn't really make sense, and try to move from there to the claim that it can't really be what we have in mind when we talk about moral responsibility. It is very clear to most people, however, and one dramatic way to characterize it is by reference to the story of heaven and hell: it is responsibility of such a kind that, if we have it, it makes sense to propose that it could be just to punish some of us with torment in hell and reward others with bliss in heaven. It makes sense because what we do is absolutely up to us. The words 'makes sense' are stressed, because one doesn't have to believe in the story of heaven and hell in order to understand the notion of ultimate responsibility that it is used to illustrate. Nor does one have to believe in it in order to believe in ultimate responsibility (many atheists have done so).
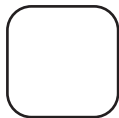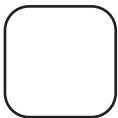
[Free Will]
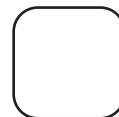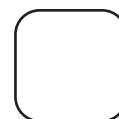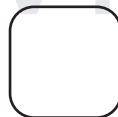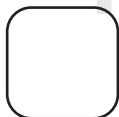
The story is useful, because it illustrates the kind of absolute or ultimate responsibility that many have supposed - and do suppose - themselves to have. (Another way to characterize it is to say that it exists if punishment and reward can be fair without having any pragmatic - or indeed aesthetic - justification.) But one doesn't have to appeal to it when describing the sorts of everyday situation that are primarily influential in giving rise to our belief in ultimate responsibility. Suppose you set off for a shop on the evening of a national holiday, intending to buy a cake with your last £10 note. Everything is closing down. There is one cake left; it costs £10. On the steps of the shop, someone is shaking an Oxfam tin. You stop, and it seems completely clear to you that it is entirely up to you what you do next: you are truly, radically free to choose, in such a way that you will be ultimately responsible for whatever you do choose. You can put the money in the tin, or go in and buy the cake, or just walk away. You are not only completely free to choose. You are not free not to choose.

Standing there, you may believe determinism is true: you may believe that in five minutes time you will be able to look back on the situation you are now in and say, of what you will by then have done, 'It was determined that I should do that'. But even if you do wholeheartedly believe this, it does not seem to touch your current sense of the absoluteness of your freedom and moral responsibility.

One diagnosis of this phenomenon is that one can't really believe that determinism is true in such situations, and also can't help thinking that its falsity might make freedom possible. But the feeling of ultimate responsibility seems to remain inescapable even if this is not so. Suppose one fully accepts the Pessimists' argument that no one can be *causa sui*, and that one has to be *causa sui* (in certain crucial mental respects) in order to be ultimately responsible for one's actions. This does not seem to have any impact on one's sense of one's radical freedom and responsibility, as one stands there, wondering what to do. One's radical

responsibility seems to stem simply from the fact that one is fully conscious of one's situation, and knows that one can choose, and believes that one action is morally better than the other. This seems to be immediately enough to confer full and ultimate responsibility. And yet it cannot really do so, according to the Pessimists. For whatever one actually does, one will do what one does because of the way one is, and the way one is is something for which one neither is nor can be responsible, however self-consciously aware of one's situation one is.

The Pessimists' argument is hard to stomach (even Hitler is let off the hook), and one challenge to it runs as follows. 'Look, the reason why one can be ultimately responsible for what one does is that one's self is, in some crucial sense, independent of one's general mental nature (character or motivational structure). Suppose one faces a difficult choice between A, doing one's moral duty, and B, following one's desires. You Pessimists describe this situation as follows. Given one's mental nature, you say, one responds in a certain way. One is swayed by reasons for and against both A and B. One tends towards A or B, and in the end one does one or the other, given one's mental nature, which is something for which one cannot be ultimately responsible. But this description of yours forgets the self - it forgets what one might call 'the agent-self'. As an agent-self, one is in some way independent of one's mental nature. One's mental nature inclines one to do one thing rather than another, but it does not thereby necessitate one to do one thing rather than the other (to use Leibniz's terms).
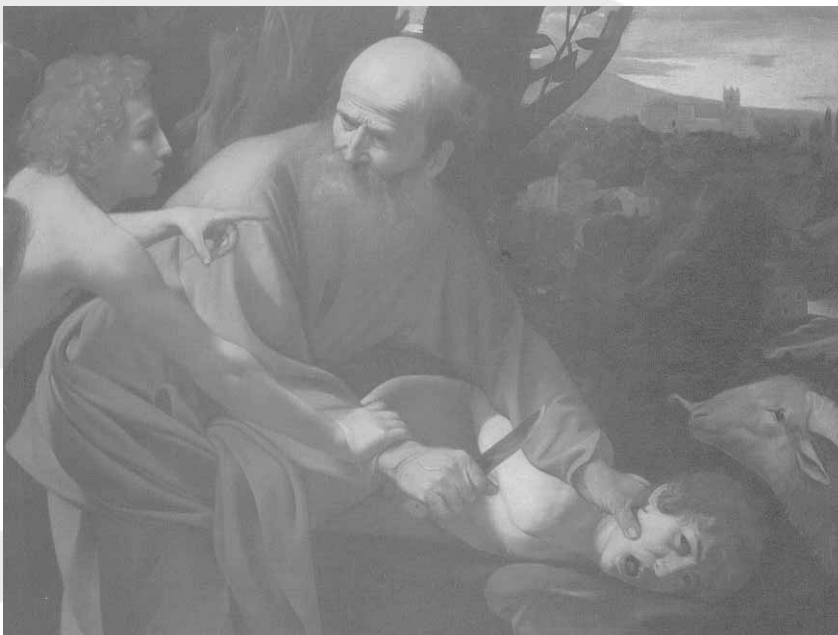
As an agent-self, one incorporates a power of free decision that is independent of all the particularities of one's mental nature in such a way that one can, after all, count as ultimately morally responsible in one's decisions and actions, even though one is not ultimately responsible for any aspect of one's mental nature.'

The Pessimists are unimpressed: 'Even if one grants the validity of this conception of the agent-self for the sake of argument', they say, 'it cannot help. For if the agent-self decides in the light of the agent's mental nature but is not determined by the agent's mental nature, the following question immediately arises: Why does the dear old agent-self decide as it does? The general answer is clear. Whatever it decides, it decides as it does because of the overall way it is, and this necessary truth returns us to where we started: somehow, the agent-self is going to have to get to be responsible for being the way it is, in order for its decisions to be a source of ultimate responsibility. But this is impossible: nothing can be *causa sui* in the required way. Whatever the nature of the agent-self, it is ultimately a matter of luck. Maybe the agent-self decides as it does partly or wholly because of the presence of indeterministic occurrences in the decision process. Maybe, maybe not. It makes no difference, for indeterministic occurrences can never contribute to ultimate moral responsibility.'

Some think they can avoid this by asserting that free will and moral responsibility are just a matter of being governed by reason - or by Reason with a dignifying capital 'R'. But being governed by Reason can't be the source of ultimate responsibility. It can't be a property that makes punishment ultimately just or fair for those who possess it, and unfair for those who don't. For to be morally responsible, on this view, is simply to possess one sort of motivational set among others. But if you do possess this motivational set, then you are simply lucky - if it is indeed a good thing - while those who lack it are unlucky.

This will be denied. It will be said, truly, that some people struggle to become more morally responsible, and make an enormous effort. Their moral responsibility is then not a matter of luck; it's their own hard-won achievement. The Pessimists' reply is immediate. 'Suppose you are someone who struggles to be morally responsible, and make an enormous effort. Well, that too is a matter of luck. You are lucky to be someone who has a character of a sort that disposes you to make that sort of effort. Someone who lacks a character of that sort is merely unlucky.'

In the end, luck swallows everything. This is one (admittedly contentious) way of putting the point that there can be no ultimate responsibility, given the natural, strong conception of responsibility that was characterized by reference to the story of heaven and hell. Relative to that conception, no punishment or reward is ever ultimately just or fair, however natural or useful or otherwise humanly appropriate it may be or seem.

Galen Strawson
University of Reading

Notes

A version of this paper first appeared in *The Times Literary Supplement*

[Free Will]
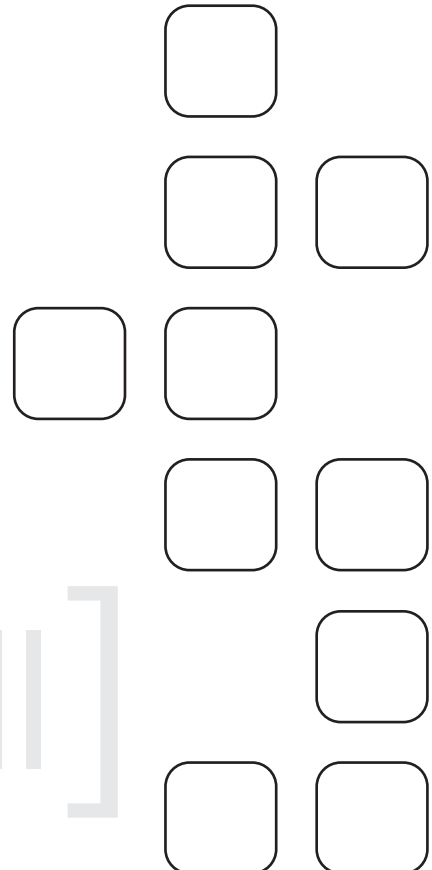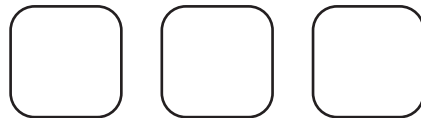
John Burgess

[Numbers]

and Ideas

I

Philosophy is a subject in which there is very little agreement. This is so almost by definition, for if it happens that in some area of philosophy inquirers begin to achieve stable agreement about some substantial range of issues, straight-away one ceases to think of that area as part of 'philosophy', and begins to call it something else. This happened with physics or 'natural philosophy' in the seventeenth century, and has happened with any number of other disciplines in the centuries since. Philosophy is left with whatever remains a matter of doubt and dispute.

Philosophy of mathematics, in particular, is an area where there are very profound disagreements. In this respect philosophy of mathematics is radically unlike mathematics itself, where there are today scarcely ever any controversies over the correctness of important results, once published in refereed journals. Some professional mathematicians are also amateur philosophers, and the best way for an observer to guess whether such persons are talking mathematics or philosophy on a given occasion is to look whether they are agreeing or disagreeing.

One major issue dividing philosophers of mathematics is that of the nature and existence of mathematical objects

and entities, such as *numbers*, by which I will always mean *positive integers* 1, 2, 3, and so on. The problem arises because, though it is common to contrast matter and mind as if the two exhausted the possibilities, numbers do not fit comfortably into either the material or the mental category.

Clearly numbers are not material bodies. The so-called numbers on the front of a house, marking its street address, may indeed be made of brass or wood or plastic. But these 'numbers' are not the numbers we speak of when we say that two is an even number, or that three is an odd number, or that both are prime numbers. Rather, they are *numerals*, or names of numbers.

Almost equally clearly, numbers are not mental in the way that, say, dreams or headaches are. They are not private to an individual. One does not speak of my number two and your number two, his number two and her number two, but simply of *the* number two. The individual, say a school child doing a simple sum, experiences the numbers as something external, about which he or she is *not* free to think whatever he or she wants.

But if numbers are not material bodies or private experiences, what (if anything) are they? Among professional academic philosophers the most commonly held views are two, for want of better terms called *realism* and *nominalism*.

Realism maintains that numbers exist, and are of a very different nature from human ideas: indeed, they differ quite as much from human ideas as they do from material bodies. They are *abstract* entities, to which it makes no sense to ascribe a position in space or date in time, and which are not causally active or acted upon. There is nowhere to go to look for a number, and you can't do anything to a number, any more than a number can do anything to you.

Nominalism maintains that numbers do not exist, and that theorems of mathematics asserting the existence of numbers are untrue, just like fairy tales asserting the existence of gnomes. To be sure, much of mathematics is applicable in science and everyday life in a way that fairy tales generally are not, but that, according to nominalists, only shows it is a *useful* fiction, not that it is non-fiction.

There are problems for both opposing philosophical views, and the problems of each are cited by the adherents of the other as reasons for embracing *it* instead. And formerly there were among philosophers also many who maintained a third view, *conceptualism* or *idealism,* according to which numbers exist, but only as shared human concepts or ideas.

The view has traditionally been popular among anthropologists and other social scientists, whose special subject matter is precisely the shared ideas of a culture. They point out that taking numbers to be such shared or *communal* ideas sufficiently explains why the school child doing a simple sum does not feel free to make up an answer at will. If numbers are ideas shared by a culture, no one member of that culture has the authority to change the rules of addition, any more than to change the rules of grammar of the culture's language.

The anthropological view has also found adherents among mathematics educators. Rather more surprisingly, the same view has won adherents among the minority of professional mathematicians who are also amateur philosophers.[1]

Conceptualist and idealist views, however, were subjected along with other nineteenth century views to a scathing critique by the late nineteenth century German mathematician and philosopher Gottlob Frege.[2] Largely as a result of that critique, the anthropological view today has virtually no adherents among professional academic philosophers. Its rejection is one of the rare cases of general agreement and consensus on an issue in philosophy.

Precisely because there is such general agreement, philosophers seldom stop to explain, in language more modern than Frege's, just what is wrong with the view that so many anthropologists, sociologists, psychologists, mathematics educators, and even mathematicians have found attractive. It is this task of explanation that I will be undertaking in the present essay, using an example of a kind that definitely would not have been used by Frege.

II

Let us begin by considering the proposition that *Bigfoot*, also known as the *Sasquatch* — a cousin of the *Abominable Snowman* or *Yeti* — exists in the realm of shared human ideas and concepts. Now certainly there is *something* in the neighbourhood that exists in the realm of shared human ideas and concepts, namely, the shared human idea or concept of Bigfoot. This is the idea of a large, hairy, humanoid creature inhabiting the wilder parts of the Pacific Northwest, from Northern California to British Columbia.

There are even people who claim to have sighted individual Bigfeet, and have formed ideas of these individuals, even to the point of giving them names like 'Harry' or 'Harriet'. The idea of an individual Bigfoot includes the traits that are common to all Bigfeet according to the general idea of Bigfoot, but also more specific elements: for instance, Harry is male and Harriet is female. These ideas of individual Bigfeet are less widely shared than the idea of the species, but we may suppose they are at least shared among members of the International Society for Cryptozoology, who take a special interest in such things.

The majority view among zoologists is that there do not, in fact, exist any large, hairy, humanoid creatures, and that the alleged sightings of Harry, Harriet, and other individual Bigfeet were either illusions or hoaxes. But I ask you to join me in assuming, just for the moment, that majority is wrong, and that creatures of the kind indicated, including Harry and Harriet, do exist. On this assumption, I will argue, two things should be clear.

The first is that Harry, Harriet, and other large, hairy, humanoid creatures inhabiting the wilder parts of the Pacific Northwest are very different sorts of things from shared human ideas and concepts, and in particular, are very different sorts of things from the ideas and concepts of Harry, of Harriet, and of Bigfoot in general. They differ in absolutely fundamental respects, for instance, in their location in space and time.

[Numbers]
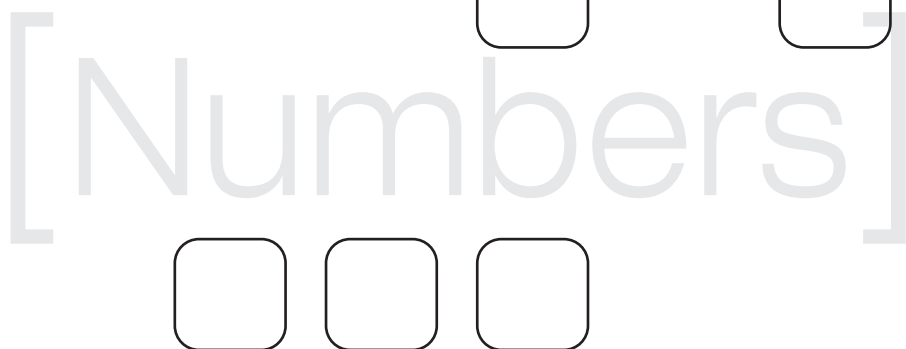
Let us consider space, for instance. (Similar considerations would apply to time.) It is not clear whether or where a shared human idea or concept should be thought of as located in space, but presumably if it is located anywhere, it is located where the human beings who share it are located. Thus if the International Society for Cryptozoology holds its annual convention on the banks of Loch Ness, the idea of Bigfoot in general, and the ideas of Harry and Harriet in particular, are located mainly in Scotland. Harry, Harriet, and the rest of their kind, however, are still located in Washington or Oregon or thereabouts. The creatures cannot be the ideas, because the two are located in different places.

The creatures differ from the ideas also in respect of how many of them there are. People have ideas of Harry, Harriet, and several more Bigfeet that have allegedly come into contact with human beings; but there are supposed to be, according to the minority view I have asked you to assume for the moment, more Bigfeet than just these: more individuals like Harry and Harriet than there are shared human ideas of individual Bigfeet. So again the creatures cannot be the ideas, since there are more of the former than of the latter.

A second point I hope will be clear is that it is the flesh-and-blood creatures, not the ideas, that are the Bigfeet. The term 'Bigfoot' refers to the inhabitants of the wilds of Washington and Oregon, not to the contents of minds or brains of the cryptozoologists assembled in Scotland. If we wish to refer to the latter, we must use some other expression than the word 'Bigfoot', such as the phrase 'the idea of Bigfoot'.

In short, on the minority view, according to which the flesh-and-blood creatures do exist, the following is the case: Bigfeet, being flesh-and-blood creatures, are not ideas, and are more numerous than the ideas of them and located in a different place from those ideas.

III

Are things any different on the majority view? It is when one assumes that there are no such flesh-and-blood creatures that some are tempted to say that the Bigfoot in general, or Harry and Harriet in particular, are human ideas. I think this temptation should be rejected.

Let me say straight-away that it would be pointless to object to someone expressing *disbelief* in Bigfoot by saying, 'Bigfoot exists only in the imagination of the credulous,' or something of the sort. Someone might well say this — I might well say it myself, for that matter, when not talking philosophy — and mean it only as a manner of speaking, as a way of saying, 'Bigfoot doesn't exist at all, though some credulous persons imagine that it does.' The proposition I want to consider, however, is that Bigfoot *literally* does exist, but only in the realm of shared human ideas and concepts, where, according to the anthropological view, numbers also have their being.

To indicate the reasons why I reject this proposition, suppose the population of some endangered forest or swamp species falls until there is only one left. So long as this one surviving flesh-and-blood or wood-and-sap organism lives, considerations of the kind already adduced in the case of Bigfoot indicate that *it* is the only member of the species, and *it* is

not an idea, from which it follows that the members of the species are not ideas.

Now suppose this last survivor also perishes. Are we now to say that the species still has members, but that the members of the species are now ideas? Should we say that the species has not become extinct but rather has undergone a *metamorphosis*, transcending its former carnal or xyline nature, and taken on a conceptual essence: that its members have cast aside their fleshly or wooden bodies, and are now made of whatever ideas are made of? Should we say that the species has undertaken a *migration*, abandoning the woods or marshes that were once its home, and occupying now instead a niche in the minds or brains of human subjects?

It seems to me about as plain as anything can be in philosophy — where admittedly things are never as plain as they are in some other disciplines — that this is *not* what we should say, and that the *correct* way to the describe the situation is by saying that creatures of this animal or plant species *simply no longer exist at all*, though of course human ideas about them do exist, and may perhaps continue to exist as long as the human species does.

Likewise in the case of Bigfoot. If the forest creature exists, then Bigfoot is that forest creature, and is something very different from an idea. If the forest creature does not exist, then Bigfoot is, so to speak, even *more* different from an idea: for in that case Bigfoot is *nothing*, while the idea is at least *something*, and what could be more different than something and nothing?

The case is the same, I maintain, with our shared human ideas and concept of number in general, and of individual
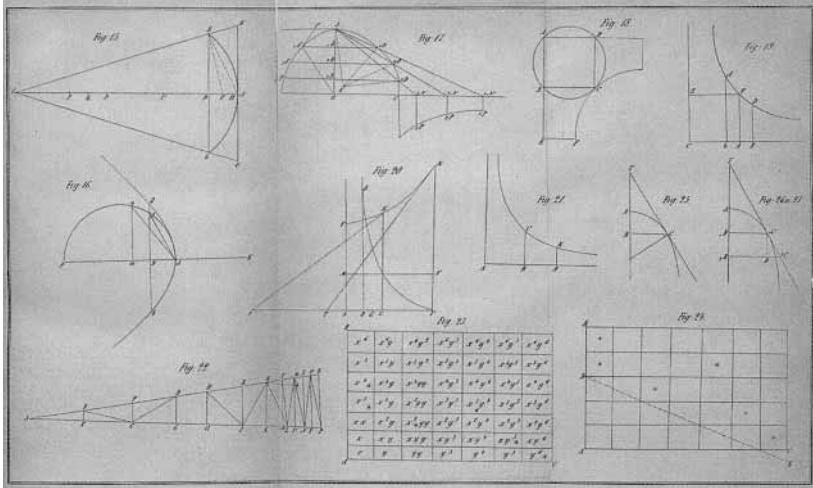
numbers such as one or two or three. (Again the individual ideas contain whatever is contained in the general idea, plus additional distinguishing elements. We no longer imagine, as did the Pythagoreans, that two is female and three is male, but, for instance, two is even and three is odd.)

These ideas are clear enough, I maintain, to indicate that one, two, three, and the other numbers, if they exist at all, do not have the same sort of spatial or temporal features as human ideas, and above all are more numerous than human ideas could possibly be.

Taking first issues of time and place, mathematics is used throughout science, and mathematical objects and entities are referred to in all its branches, including those like cosmology that deal with times and places very remote from any inhabited by human beings. Are we to say that a cosmologist's estimates of the relative numbers of heavy and light elements at a certain stage in the early evolution of the universe must be wrong, because there were no numbers at all back then, no human beings having yet evolved to create them? Surely not.

And then there is the matter of infinity. It is a crucial feature of the concept of the number system that it has infinitely many elements, that there are infinitely many numbers. But surely human beings have formed ideas or concepts of only finitely many of them. There simply are not enough human ideas and concepts for each number to be one. *Some* numbers at least must therefore either enjoy a mode of existence different from that of any human idea, as realists maintain, or else must simply fail to exist, as nominalists hold. And is it not preposterous to maintain that while one of the pair realism or nominalism gives the correct account of mathematical existence in the case of *some* numbers, conceptualism is correct *for the rest*? Surely the question of the existence and nature of numbers has a *uniform* answer, and if conceptualism fails in *any* case, then it must fail in *all*.

[Numbers]

IV

Such then, are some of the principal reasons why I and almost all professional philosophers of mathematics reject conceptualism, and consider the only real issue to be that between nominalism and realism. This last issue is far too large to be thrashed out here, but I do wish to say a word about it, and in particular about the character of the *realist* position, which very often tends to be misrepresented. Nominalists do not believe in numbers because they cannot *see* them (or see any visible effects *caused* by them), and tend to represent their opponents as claiming that they *can* see them.

According to an old story, Plato was once lecturing in his Academy on his Forms, and was speaking of the Forms of 'tableness' and 'cupness'. Diogenes the Cynic interrupted and said, 'O Plato, I see the table and the cup, but the tableness and the cupness I do not see.' To this Plato replied, 'Very naturally, Diogenes, since you have eyes, by which material things are perceived, but lack Intellect, by which the Forms are seen.'[3]

Nominalists tend to represent their opponents as Platonists, maintaining that if numbers do not emit electromagnetic radiation to which the eye is sensitive, then they must be emitting something else, perhaps *noetic rays*, which can be sensed by some other organ, perhaps the *pineal gland*. This, however, is a misrepresentation of realism. Or at least, I have never known a single realist who was in any meaningful sense a Platonist.

What is actually the case is that anti-nominalists take much more seriously than nominalists the thought that *mathematics* is a human creation, since mathematics is a body of theory expressed in language, and *language* is a human creation.

Now creating a language involves creating certain rules for its use. Among these is, I believe, a rule to the effect that tense and date are not to be applied to mathematical existence assertions. One can say 'There exist infinitely many prime numbers,' but to ask 'How many of them already existed in 1000 BCE, or during the Cenozoic Era?' is to commit a kind of grammatical solecism.

Nominalists say they are opposed to the view that numbers are 'eternal', existing 'outside of time'. But to say that numbers are 'eternal' is a misleadingly Platonistic way of putting the simple negative grammatical fact of the inapplicability of tense distinctions in mathematical contexts. That simple grammatical point is all the realist really believes about the 'timelessness' of number.

(By contrast with the case of the numbers themselves, it makes perfect sense to ask whether the *idea* or *concept* of prime number had emerged by 1000 BCE — the issue involved would be that of the interpretation of certain Babylonian tablets and Egyptian papyri — and it makes perfect sense to assert that it had *not* emerged in the Age of the Dinosaurs. This difference between the 'timeless' numbers proper and datable ideas of them was one of the points I was arguing in rejecting conceptualism.)

Likewise, there are certain rules or standards as to what counts as adequate or sufficient to establish or prove a mathematical existence theorem, and by these rules Euclid's Theorem on the existence of infinitely many prime numbers is as well-established as anything can be.

The *nominalists* assume that they have an understanding of what it would be for a mathematical object or entity to exist that is independent of ordinary mathematical standards of sufficient proof, by reference to which understanding they can *criticize* the ordinary mathematical standards. So-called realism is really just *skepticism* about the existence of any understanding of what 'existence' means in mathematics that is independent of ordinary mathematical standards for evaluating existence proofs. The nominalist denies the existence of numbers, while the realist denies that the nominalist understands what is meant by 'existence' as applied to numbers.

Thus the realists think the nominalists are confused. But realists and nominalists agree that the conceptualists are confused, and while I cannot hope to have convinced anyone by the foregoing very brief remarks that the realists are right as against the nominalists, I hope I have

convinced some of you that realists and nominalists are right in their common opposition to conceptualism.

Notes

1   The classical expression of the anthropological view is that of Leslie A. White, 'The locus of mathematical reality: an anthropological footnote', *Philosophy of Science* 14 (1947), 289-303, more readily available in Newman, J. R. (ed.) *The World of Mathematics* vol. 4 (New York: Simon & Schuster, 1956). For a recent endorsement by the mathematician Reuben Hersh, see his *What is Mathematics, Really?* (Oxford: Oxford University Press, 1997), a book that makes a professional philosopher's hair stand on end.

2   *Grundlagen der Arithmetik* (1884), translated by J. L. Austin as *The Foundations of Arithmetic*, (London: Blackwell, 1950), reprinted in paperback (New York: Harper & Row, 1960); the critical portions (the part of the book relevant to the present essay) are reprinted in Benacerraf, P. & Putnam, H. (eds.) *Philosophy of Mathematics: Selected Readings* 2nd ed. (Cambridge: Cambridge University Press, 1983).

3   See Diogenes Laertius, 'Diogenes [the Cynic]', in *Lives and Opinions of Eminent Philosophers*, translated by R. D. Hicks, Loeb Classical Library (Cambridge: Harvard University Press, 1925).

John P. Burgess
Princeton University

[Numbers]
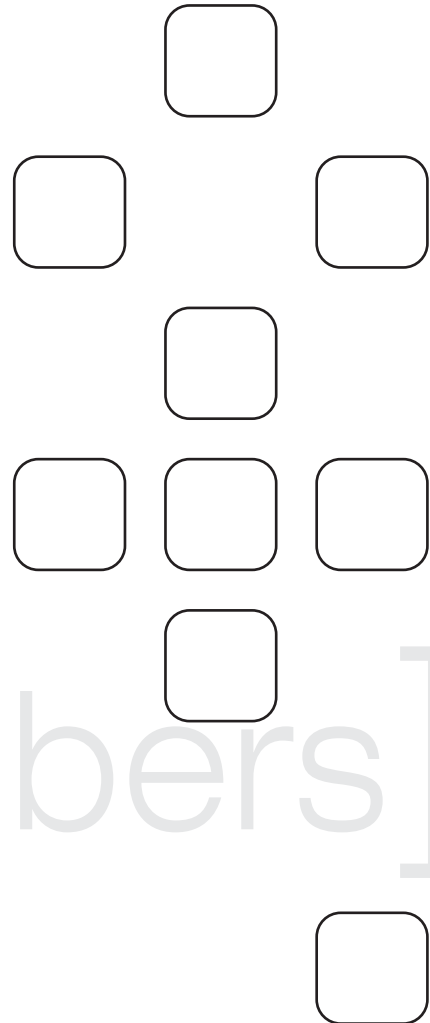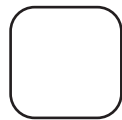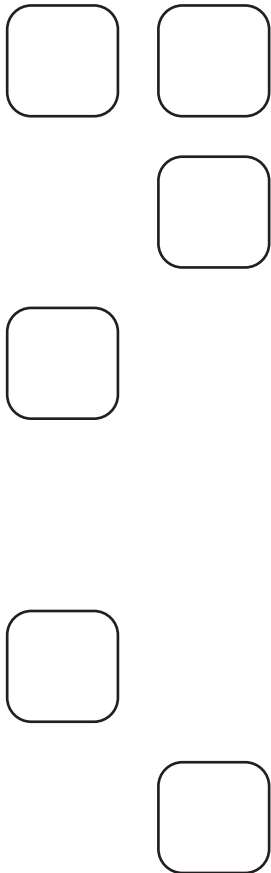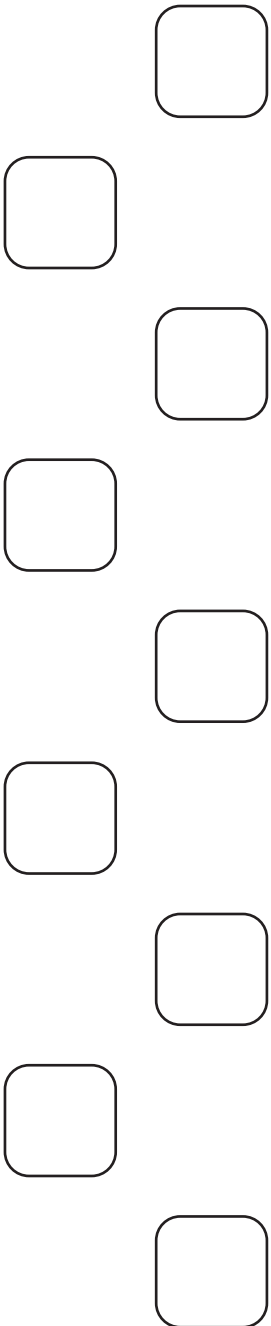
Chris Hughes

[Attribution]

Identity and

in the Trinity

I

Philosophers of religion often argue about whether certain religious doctrines (e.g. monotheism) are true. They also argue about whether certain religious doctrines are consistent (i.e. free from contradiction). Thus while some philosophers have argued that it is inconsistent to suppose that a perfectly good God created a world containing evil, others have argued (rightly, in my view) that there is no inconsistency in that supposition. One religious doctrine whose consistency has been a matter of long-standing dispute among philosophers of religion is the doctrine of the Trinity.

Why might someone think that the doctrine of the Trinity is *in*consistent? Well, the doctrine of the Trinity implies that:

(1) The Father, the Son, and the Holy Spirit are (three) different persons.

(2) The Father is God, the Son is God, and the Holy Spirit is God.

(3) There is just one God.

And, it might be thought, (1) – (3) are inconsistent.

Suppose someone says that the baker is Bob, and the butcher is Bob. If she has said something true, it seems, either the baker is (a) Bob, and the butcher is (a different) Bob, or the baker is (a) Bob, and the butcher is

(that same) Bob. In just the same way, it seems, if the Father is God, and the Son is God, then either the Father is (a) God, and the Son is (a different) God, or the Father is (a) God, and the Son is (that same) God. (Another, more idiomatic, way to say that the Father is (a) God and the Son is (that same) God is to say that the Father is God, and the Son is too).

If the Father is a God, and the Son is a different God, then there are two different Gods (the God the Father is, and the (different) God the Son is). And if there are two different Gods, then there isn't just one God (as (3) says). We may conclude that, if the Father is a God, and the Son is a different God, (1) – (3) cannot all be true.

If on the other hand, the Father is a God, and the Son is that same God, then there is no contradiction in the supposition that the Father is a God, and the Son is a God, and yet there is just one God (as (3) says). But if the Father is a God, and the Son is that same God, can it be true that the Father, the Son, and the Holy Spirit are three different persons (as (1) says)? Well, if the Father and the Son are different persons, and the Father is God, and the Son is that same God, then – it would seem to follow – the Father and the Son are different persons, but the same God. Is this possible?

Not a few philosophers have argued that if K and K^* are kinds of thing (person, God, human, hamster, or the like), there is a contradiction in supposing that different Ks are the same K^*.[1] Their argument has the following structure:

To say that Eric Blair and George Orwell are the same man is to say that Eric Blair is a man, and George Orwell is a man, and Eric Blair is the same as George Orwell. More generally, to say that x and y are the same K is to say that x is a K, and y is a K, and x is the same as y. In the same way, to say that Jerry Garcia and Phil Lesh are different men is to say that Jerry Garcia is a man, and Phil Lesh is a man, and Jerry Garcia is different from (not the same as) Phil Lesh. More generally, to say that x and y are different Ks is to say that x and y are Ks, and x is different from (not the same as) y.

Now suppose that x and y were different Ks, but the same K*. It would follow that (a) x and y are Ks, and x is different from (not the same as) y, and (b) x and y are K*s, and x is the same as y. But (a) and (b) are inconsistent, since (a) says that x is not the same as y, and (b) says that x is the same as y. So it cannot happen that x and y are different Ks, but the same K*.

Suppose this argument is sound. Then the Father and the Son could not be different persons, but the same God, since that would imply that both i) the Father and the Son are persons, and the Father is different from the Son, and ii) the Father and the Son are Gods, and the Father is the same as (not different from) the Son. This implies that *The Father and the Son are different persons* and *The Father is (a) God, and the Son is (the same) God* cannot both be true together. Thus if the Father is a God, and the Son is the same God, (1) – (3) cannot all be true.

Drawing the threads together, we have the following argument against the consistency of the doctrine of the Trinity:

Necessarily, if (2) is true, the Father is a God and the Son is a different God, or the Father is a God and the Son is the same God. If the Father is a God, and the Son is a different God, then (3) cannot be true. If the Father is a God, and the Son is the same God, then (1) cannot be true. So necessarily, if (2) is true, it cannot be that both (1) and (3) are true. Thus (1) – (3) are jointly inconsistent (in other w o r d s , cannot all be true together).

II

Some defenders of the doctrine of the Trinity would respond to the above argument by denying that 'x and y are the same K' 'splits up' into 'x and y are Ks, and x and y are the same', and denying that 'x and y are different Ks' 'splits up' into 'x and y are Ks, and x and y are different'.[2] If they are right, that would certainly block the argument for the inconsistency of the doctrine of the Trinity sketched above. If the relation, *is a different person from*, cannot be 'factored' into something one of whose factors is *is different from (is not the same as)*, and the relation, *is the same God as*, cannot be factored into something one of whose factors is *is the same as*, there is no longer any obvious reason to think that it is contradictory to suppose that the Father stands in the different-person-from relation to the Son, whilst standing in the same-God-as relation to the Son.
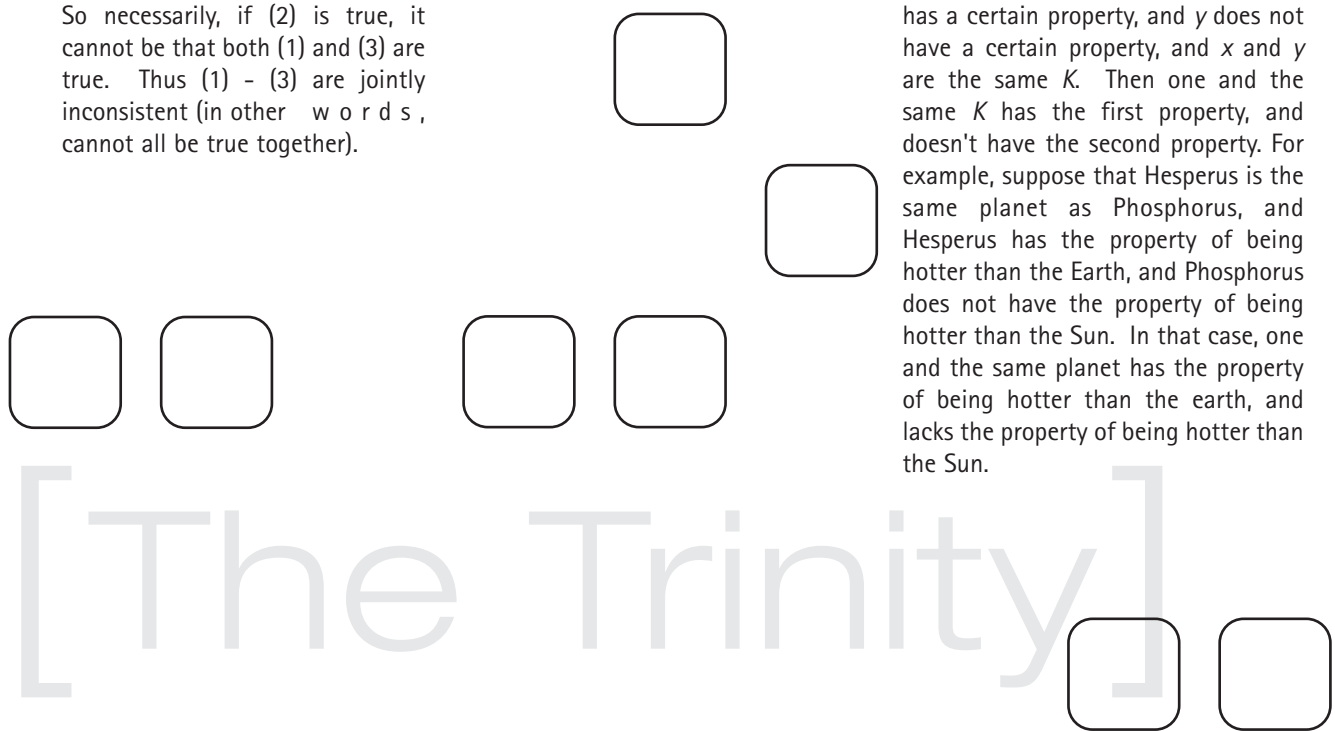
I think this sort of defence of the consistency of the doctrine of the Trinity has at least two drawbacks. First, there is something at least initially intuitive about the 'factorability' thesis, and the *onus probandi* (Ed. the burden of proof) is accordingly on those who deny it to explain why it is not true. (In fairness to those who deny it, they have attempted to do so; for reasons of space, I cannot go into why I judge their attempts unsuccessful). Second, there are compelling arguments to the effect that different Ks cannot be the same $K*$ which do not make explicit appeal to the factorability thesis. One (well-known) argument of this type relies on a *lemma* - an intermediate conclusion argued for on the way to arguing for the ultimate conclusion of the argument.[3] The lemma is:

(λ) (For any kind K) If x is the same K as y, then x and y are indiscernible (have e x a c t l y the same properties).

Why believe (λ)? Well, suppose that x has a certain property, and y does not have a certain property, and x and y are the same K. Then one and the same K has the first property, and doesn't have the second property. For example, suppose that Hesperus is the same planet as Phosphorus, and Hesperus has the property of being hotter than the Earth, and Phosphorus does not have the property of being hotter than the Sun. In that case, one and the same planet has the property of being hotter than the earth, and lacks the property of being hotter than the Sun.

[The Trinity]

Now suppose that x were the same K as y, and x and y were discernible with respect to some property P. That is, suppose that x were the same K as y, and x had the property P, and y lacked it. It would follow that one and the same K both had and did not have the property P. But one and the same K cannot both have and not have one and the same property. If there were a counterexample to (λ), one and the same K would both have and not have one and the same property. But that is a contradiction, and impossible; so there cannot be a counterexample to (λ).

Now suppose it were true that x and y were the same K. Could it also be true that x and y were different K^*s? Well, by our lemma (λ), if x and y are the same K, then x and y are indiscernible. If however, x and y were different K^*s, x and y would have to be discernible.

Why? Because if x and y were different K^*s, they would have to be discernible with respect to the property of *being the same K^* as y*. (If x is a different K^* from y, then x is not *the same K^** as y. Also, if x is a

different K^* from y, then y, like x, is a K^*, *and y surely* is the same K^* as y (since anything that is a K^* is the same K^* as itself). So if x and y were different K^*s, they would have to be discernible with respect to at least one property (to wit, *being the same K^* as y*).) But, by (λ), if x and y were the same Ks, they would have to be indiscernible with respect to all properties. From this it follows that if x and y are the same K, they can't be different K^*s. If x and y cannot be different K^*s without being discernible with respect to some property, and x and y cannot be the same K without being indiscernible with respect to all properties, then x and y cannot be the same K and different K^*s.

If all of this seems abstract, it may help to apply it to the theological case under consideration. Suppose that the Father and the Son are one and the same God. It follows, by (λ), that the Father and the Son are indiscernible with respect to all properties. If, however, the Father and the Son are different persons, then they are discernible with respect to at least one property - namely, *being the same person as the Father*. (If the Father and the Son are different persons, then the Father is a person, and the Father is the same person as the Father. But if the Father and the Son are different persons, then the Son is the not the same person as the Father. So if the Father and the Son are different persons, then the Father and the Son are discernible with respect to *being the same person as the Father*.)

So it doesn't look as though we can defend the consistency of the doctrine of the Trinity by saying that (a) the Father stands in the same-God-as relation to the Son, and (b) stands in the different-person-from relation to the Son. Given (the apparently unimpeachable) (λ), (a) implies that the Father and the Son have exactly the same properties, while (b) implies that the Father and the Son don't have exactly the same properties.

III

If Hesperus is Phosphorus, then Hesperus is the same planet as Phosphorus. If Albion is England, then Albion is the same country as England. On the basis of examples like these, it is tempting to conclude that, where t and t' are any terms,

If t is t', then t is the same something as t'.

i.e.

If t is t', then for some kind of thing K, t is the same K as t'.

But consider statements such as:

This statue is marble.

and

This penny is copper.

Assuming that this statue is marble, is the statue the same something as marble? It seems not. Why not? Well, a statue is a thing of a certain kind, and marble is a stuff of a certain kind. Since a statue is a thing of a certain (material) kind, a statue has a particular size, a particular shape, and so on. Since marble is a stuff of a certain kind, rather than a thing of a certain kind, marble does not have any particular size or shape. (The question, 'How large is that statue?' makes perfect sense; the question, 'How large is marble?' does not. Things *made of* marble have a particular size and shape; but marble itself does not).

Now suppose that for some kind K, this statue were the same K as marble. As we saw in the last section, if Hesperus and Phosphorus are the same planet, then one and the same planet has all

the properties that Hesperus has, and lacks all the properties that Phosphorus lacks. In just the same way, if for some kind *K*, this statue is the same *K* as marble, then one and the same *K* has all the properties that this statue has, and lacks all the properties that marble lacks. So if for some kind *K*, this statue is the same *K* as marble, then one and the same *K* both has a particular size (since every statue has some particular size), and lacks that particular size (since marble is not the kind of entity that has any particular size).   This is a contradiction. Since a contradiction follows from the supposition that this statue is the same *K* as marble, we may conclude that the statue isn't the same *K* as marble. This statue is marble, but it isn't 'the same anything' as marble.

How can this be? If *x* is not the same anything as *y*, how can *x* be *y*? Well, although a thing cannot be a (kind of) stuff, a thing can be *made of* a kind a stuff. A statue cannot be the kind of stuff, marble, but a statue can be made of the kind of stuff that is marble. A penny cannot be the metallic stuff (the metal) copper, but it can be made of the metal copper. And, it seems, a sentence such as

> This penny is copper.
>
> or
>
> This statue is marble.

can be understood (indeed, is naturally understood) as equivalent to:

> This penny is (made of) (the metal) copper.
>
> or
>
> This statue is (made of) (the stuff) marble.

Here is a slightly different case, not involving things, in which it can be true that *t* is *t'*, even though *t* is not the same anything as *t'*: graphite is carbon. But is graphite the same anything as carbon? I doubt it. Carbon is an element, but graphite is not (there is no such element as graphite). Also, suppose that graphite were the same element as carbon. It would then likewise be true that diamond is the same element as carbon. Now if *x* is the same *K* as *v*, and *y* is the same *K* as *w*, and if *x* stands in a certain relation to *y*, then *v* stands in that same relation to *w*.

So if graphite is the same element as carbon, and diamond is the same element as carbon, then: if graphite is softer than diamond, then carbon is softer than carbon. But graphite is softer than diamond, and carbon is not softer than carbon (since nothing is softer than itself). From this we may conclude that neither graphite nor diamond is the same element as carbon. And if graphite is not the same element as carbon, then graphite and carbon are discernible, since carbon certainly is the same element as carbon; whence (by (λ)) we may conclude that graphite isn't the same anything as carbon (isn't the same *K* as carbon, for any *K*).

Though graphite (or diamond) is not the same element as carbon, graphite (like diamond) is what chemists call an *allotrope* of carbon. And a sentence such as:
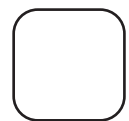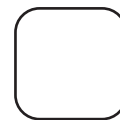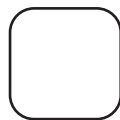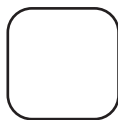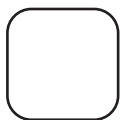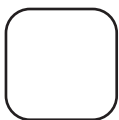
> Graphite is carbon.
>
> or
>
> Diamond is carbon.

can be understood (and is naturally understood) as equivalent to

> Graphite is (an allotrope of) carbon.
>
> or
>
> Diamond is (an allotrope of) carbon.

Once we see that *'t* is *t"* has a reading on which it is equivalent to '*t* is (made of) (the stuff) *t"*, we can see that there is no problem about the consistency of the following triad:

(4) This penny, that penny, and that other penny are three different pennies.

(5) This penny is copper, that penny is copper, and that other penny is copper.

(6) There is just one copper.

(We take (5) to be equivalent to 'This penny is (made of) (the element) copper, and that penny is (made of) (the element) copper, and that other penny is (made of) (the element) copper' and we take (6) to be equivalent to 'There is just one (element) copper.') There is no reason that three different things can't all 'be' one and the same (kind of) stuff, if 'be' means 'be made of'.

Similarly, there is no problem about the consistency of this triad:

(7) Graphite and Diamond are different kinds of stuff.

(8) Graphite is carbon, and Diamond is carbon.

(9) There is just one carbon.

((7) is true because Graphite and Diamond are different allotropes, and allotropes are kinds of stuff; we take (8) to be equivalent to 'Graphite is (an allotrope of) (the element) carbon, and Diamond is (an allotrope of) (the element) carbon, and (9) to be equivalent to 'There is just one (element) copper'.) Again, there is no reason that three different (allotropic) stuffs cannot be the same (elementary) stuff, if 'be' means 'be an allotrope of'.

To return to the doctrine of the Trinity: that doctrine is traditionally understood as involving the claim that in God there is a unity of substance (or essence or nature), and a trinity of persons. Thus according to the *Quicumque Vult*, 'the Catholic Faith is this: that we worship one God in Trinity, and Trinity in Unity, neither confounding the Persons, nor dividing the Substance.' There are three divine persons, but only one divine nature or essence or substance, just as there is only one God. Thus the Father and the Son are 'one in nature' or 'one in substance' ('consubstantial').

Suppose that there are three divine persons of one and the same divine essence or nature or substance. Suppose further that the divine nature or essence or substance just is God. (That God just is His nature or essence was a supposition made by almost all mediaeval philosophical theologians: that supposition is reflected in the fact that, although 'man' and 'humanity' are not synonyms, 'God' and '(the) Deity' *are* usually understood as synonyms).

In that case, it seems it could be true that:

(1) The Father, the Son, and the Holy Spirit are three different persons.

(2) The Father is God, the Son is God, and the Holy Spirit is God.

(3) There is just one God.

– as long as (2) is understood as equivalent to:

(2*) The Father is (a person of) (the nature or substance) God, The Son is (a person of) (the nature or substance) God, and the Holy Spirit is (a person of) (the nature or substance) God.

and (3) is understood as equivalent to:

(3*) There is just one (nature or substance) God.

Just as three different statues could all be (statues of) the same (kind of) stuff – marble, and two different allotropes could both be (allotropes) of the same element - carbon, so (it seems) three different divine persons could all be (persons of) the same nature or substance - God.

To forestall a possible misunderstanding: the point here is not that God is a kind of (material) stuff, or an element (any more than the divine persons are material things or allotropes of (material) stuffs). It is instead that there is no evident inconsistency in supposing that the Father and the Son are different, but the God the Father is and the God the Son is are the same God, as long as 'the God the Father is' and 'the God the Son is' are understood as equivalent respectively to 'the substance God that the Father is a person of', and 'the substance God that the Son is a person of' (Compare: this penny and that penny are different, but the kind of stuff this penny is (made of) and the kind of stuff that penny is (made of) are the same kind of stuff)

We can now see where the argument for the inconsistency of the doctrine of the Trinity sketched in section III breaks down. If 'The Father is God' is equivalent to 'The Father is a person of the substance God', we cannot move from

(i) The Father is God, and the Son is too.

to

(ii) The Father and the Son are the same God.

and thence to

(iii) The Father is God, and the Son is God, and the Father is the same as the Son.

and

(iv) The Father is not a different person from the Son.

Arguing that way will be no better than moving from:

(i') Graphite is carbon, and Diamond is too.

to

(ii') Graphite and Diamond are the same element--carbon.

and thence to

(iii') Graphite is carbon, and Diamond is carbon, and Graphite is the same as Diamond.

and

(iv') Graphite is not a different allotrope from Diamond.

This last argument is a bad one, because if we understand *Graphite and Diamond are the same same element - carbon* as equivalent to *Graphite is Carbon, and Diamond is too*, where 'is' means 'is an allotrope of the element', we may not move from *Graphite and Diamond are the same element - carbon* to *Graphite and Diamond are the element carbon, and Graphite and Diamond are the same.* In just the same way, if we understand *The Father and the Son are the same God* as equivalent to *The Father is God, and the Son is too*, where 'is' means 'is a person of the substance', we cannot move from *The Father and the Son are the same God* to *The Father and the Son are Gods, and the Father and the Son are the same.*[4]
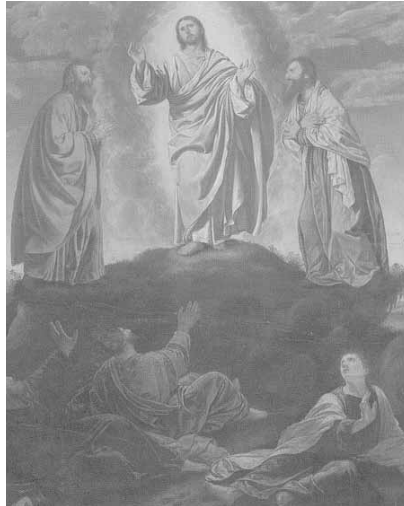
IV

This is, I think, a promising way of meeting one sort of challenge to the consistency of the doctrine of the Trinity. But other worries might be raised, one of which I shall set out briefly.

Consider properties such as omnipotence, omniscience, omnibenevolence, and so on (which I shall call *omni-properties*). If there are three divine persons and one God, it seems, there are four possibilities:

(a) The omni-properties are properties of neither the divine persons nor God.

(b) The omni-properties are properties of (some or all of) the divine persons, but are not properties of God.

(c) The omni-properties are properties of God, but are not properties of (some or all of) the divine persons.

(d) The omni-properties are properties of both the divine persons and God.

I take it that any orthodox account of the Trinity will rule out (a).

There are at least two difficulties about endorsing (b). First, if we deny that God has the property of being omnipotent, it sounds as though we are denying that God is omnipotent. (How can we (truly) speak of 'God Almighty', and if God doesn't actually have the property of almightiness (that is, omnipotence)?

[The Trinity]
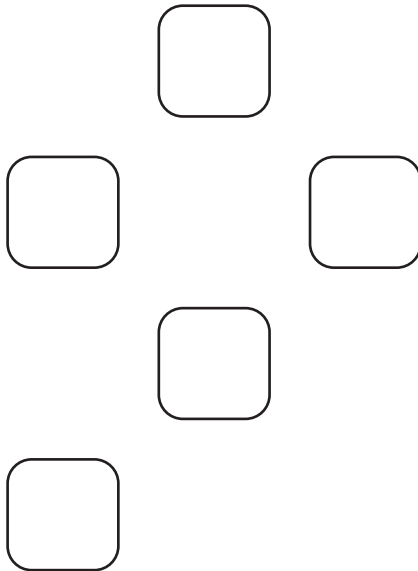
Second, if we say that each one of three different divine persons has all the omni-properties, there is a question about whether we are preserving the spirit as well as the letter of monotheism. Someone might well offer 'omnipotent, omniscient, being' as a definition of 'God'. So, we could ask, if there really are three different omnipotent, omniscient beings, why aren't there three Gods? Moreover (b) implies that the being that we are calling 'God' lacks the omni-properties. If the divine persons have the omni-properties, and the thing we are calling 'God' lacks them, why does the thing we are calling 'God' deserve that name better than the things we are calling 'divine persons'?

Suppose on the other hand, we accept (c). We now have no problem about how we can speak of, say, God almighty, since, if (c) is true, God has the property of almightiness or omnipotence. But accepting (c) raises new difficulties.

First, it seems that nothing could be omnipotent, omniscient, and the like, unless it is a person. Now if God is a person, is He the same person as the Father, and the same person as the Son, and the same person as the Holy Spirit? No, because if x is the same person as y and x is the same person as z, then y is the same person as z. As an instance of this general principle, if God is the same person as the Father, and God is the same person as the Son, then the Father is the same person as the Son. But the Father is not the same person as the Son, since the Father, the Son, and the Holy Spirit are three different persons (cf. (1)).

So if God is a person, is He the same person as the Father, and a different person from the Son? Surely, the doctrine of the Trinity rules out that

God is the same person as *this* person of the Trinity, but a different person from *that* person of the Trinity.

If, however, God is a person, but He's not the same person as all the persons of the Trinity, and He's not the same person as some but not all the persons of the Trinity, then God must be a different person from all three divine persons. So now it looks as though in the Godhead there are *four* different divine persons rather than three, which is surely contrary to the doctrine of the Trinity.

If we endorse (d), we face the same question about why there aren't four persons in the Godhead. There is also a question about why we haven't in effect got quadritheism. If the Father, the Son, and the Holy Spirit, and God, are all persons with a full complement of omni-properties, why aren't there four Gods?

V

The first challenge to the consistency of the doctrine we considered had this form:

for any kinds K, K^*, it is always inconsistent to suppose that different Ks are the same K^*. Since the doctrine of the Trinity implies that different divine persons are the same God, it is inconsistent. The second challenge has this form: even granting that – given the right sort of gloss of 'be' – different Ks can be one and the same K^*, there is still no way of 'parcelling out' the omni-properties to the divine persons and to God in such a way that the doctrine of the Trinity comes out true.

The second challenge seems to me more pressing than the first. To meet it, one would have to show either that (a) – (d) do not exhaust the

possibilities, or explain why a defender of the doctrine of the Trinity can after all (consistently) embrace one of (a) – (d). This is no easy task, and I shall not undertake it here.

Chris Hughes
King's College London

Notes

1   See for example J. Perry, 'The Same *F*', *Philosophical Review* 78 (1970), 181-201.

2   Notably Peter Geach, 'Identity', in *Logic Matters*, (Oxford: Blackwell,1972.)

3   See D. Wiggins, *Sameness and Substance*, (Cambridge, MA: Harvard University Press, 1980), pp. 18 - 23, and J. Perry, 'The Same *F*'.

4   The three-pennies-and-one-copper example was suggested to me by a similar one in Augustine's *De Trinitate* 7, 6; for the graphite-diamond-carbon example, I am indebted to Kevin Suharnic.

Clare Carlisle

[Beyond Good]
Nietzsche's
and Evil: 'Why insist on the truth?'

Friedrich Nietzsche was born in 1844 into a Protestant family in Saxony, in the north east of Germany, and began his academic career as a philologist (a scholar of classical languages). During his early twenties Nietzsche became a Professor of Philology at the University of Basle, but by the time he wrote *Beyond Good and Evil* in 1886 he had left academia and moved to the Swiss mountains in order to concentrate on his philosophy. Nietzsche's enthusiasm for high altitudes, clean air and a decent view was rather fanatical, and in fact his re-location to the Alps illustrates, metaphorically, three of his key philosophical ideas. First, his aristocratic style of thought, his contempt for mediocrity, and his reverence for the creative, exceptional individual are expressed in the image of a solitary philosopher living in the mountains high above the town, elevated (and somewhat isolated) from the common man. Second, Nietzsche rejects the values of traditional philosophy – namely, truth as opposed to falsehood, and moral goodness as opposed to evil – and replaces them with his preference for strength and good health, as opposed to weakness and sickness. Nietzsche liked anything that makes people physically and spiritually stronger; he hated the stuffy, claustrophobic atmosphere he perceived in the academic establishment, and his writing attempts to create an intellectual climate that, like the

invigorating Alpine air, enables individuals to flourish. And thirdly, the panoramic views at the summit of a mountain symbolise a feature of Nietzsche's thought that is often called 'perspectivism'[1]: unlike followers of Plato, Nietzsche insists that there is no such thing as absolute truth, and argues instead that all thinking and perception comes from a particular perspective, and that different perspectives will produce different views of truth. There are *only* these views of truth, or interpretations; there is no objective reality beneath them, no independent standard that they refer to. The task of the philosopher, then, is not to rid himself of perspective – this would not be possible, since to be a person *is* to be a particular perspective, a particular point of view, to be in only one place at once – but to look at things from many different perspectives.

These three aspects of Nietzsche's philosophy will be considered in more detail as I discuss some of the ideas presented in *Beyond Good and Evil*. Before we explore the text further, though, I should add a word of warning about how to approach Nietzsche's rather unusual style of writing. Nietzsche thought that philosophy – or at least, the kind of philosophy he approved of – was more like art than science, and it is important to bear in mind that the way he writes often reflects his philosophical views. For example, he does not attempt to conceal his own perspective by offering a theory that is supposed to be objectively and universally true.

[Truth]
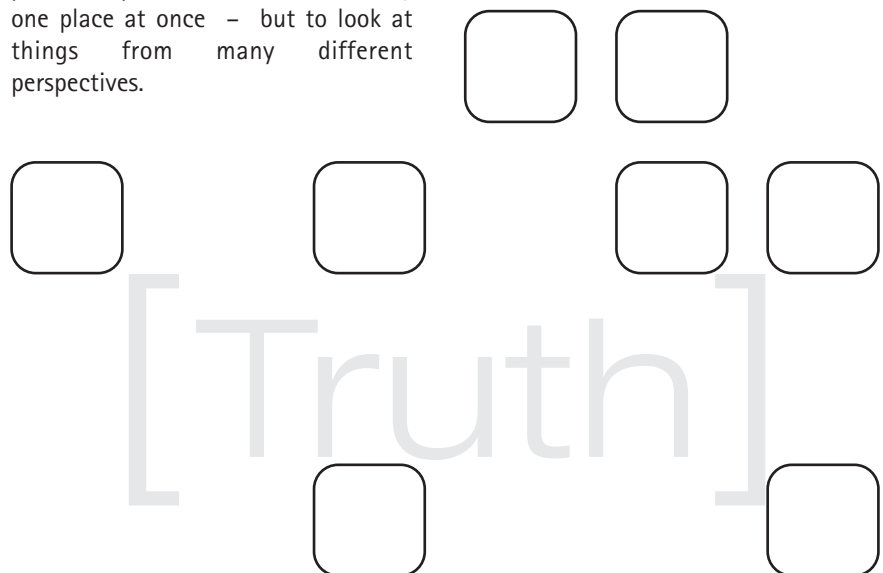
Whereas Socrates argued that knowledge is far superior to opinion, and that our ability to think rationally should control our more emotional reactions, Nietzsche's style of writing is passionate and opinionated – and this is one of the tactics he uses to undermine the philosophical tradition that Socrates influenced so heavily. Instead of constructing logical arguments to support his philosophical views, Nietzsche tends to rely more on colourful language, vivid metaphors, dialogue, myths and humour to win over his readers. This means that when we read Nietzsche we have to pay attention to the images, themes and emotional tones that recur in his writing, in much the same way as we study a novel or a poem in English Literature. It also means that, if we agree with Nietzsche's view of philosophy, we should look out for these more literary aspects of other philosophical texts in order to see how arguments that appear to be purely rational are often expressed using language that is rich in metaphors and strong feeling.

The Value of Truth

If Nietzsche's contribution to philosophy had to be summed up in a single question, this question would be, *what is the value of truth?* This quite simple and apparently innocent question is, in fact, a devastating blow to the dominant philosophical tradition, which until Nietzsche was united in its pursuit of the truth. Philosophers may have disagreed about how to define truth, how to achieve it, and indeed how attainable this goal is, but no-one had questioned the idea that truth is something good, valuable, worth pursuing and perhaps even worth dying for. Socrates, who became the hero of western philosophy, believed that truth was worth more than life itself: on his deathbed, he quite calmly told his friends that he was happy to die because his earthly, embodied existence was like a prison that barred his way to knowledge of the eternal Ideas or Forms. So when Nietzsche, at the beginning of *Beyond Good and Evil*, raises 'the problem of the value of truth', he is questioning a belief and an attitude that lies right at the heart of philosophy. 'Suppose we want truth: *why not rather* untruth? and uncertainty? even ignorance? ...Why insist on the truth?'[2]

These questions have to be understood in the context of Nietzsche's view that European culture was in decline, and heading towards nihilism. 'Nihilism' means believing in nothing, and refusing to recognise value in anything. (As we think about Nietzsche's philosophy, it is worth asking ourselves whether this view applies to the world we live in today.) Although Nietzsche's philosophy is sometimes mistakenly described as 'nihilistic', the opposite is in fact the case, for the purpose of his writing was to halt and to reverse this process of decline – to remove the causes of Europe's spiritual sickness and to create the conditions for recovery and renewal. Nietzsche believes that what he calls 'the will to truth' – that is, the unquestioning faith that truth is the highest value, and the pursuit of truth at all costs – drains the value out of life. This tendency, he argues, is exemplified not only by Socrates and philosophy, but by science and by the morality taught by Jewish and Christian religions. This means that philosophy, science and religion – the three cornerstones of intellectual life – are in fact responsible for the decline of European humanity. When Nietzsche attacks these, he is not just being wilfully destructive, but attempting to restore Europe to spiritual good health and vitality by encouraging people to have a more positive attitude to life, and even to create new values. Instead of using truth as the highest standard of value, Nietzsche argues, individuals need to develop their own powers of judgment and to produce ideas and ethics that will strengthen them and help them to live: 'The falseness of a judgment is for
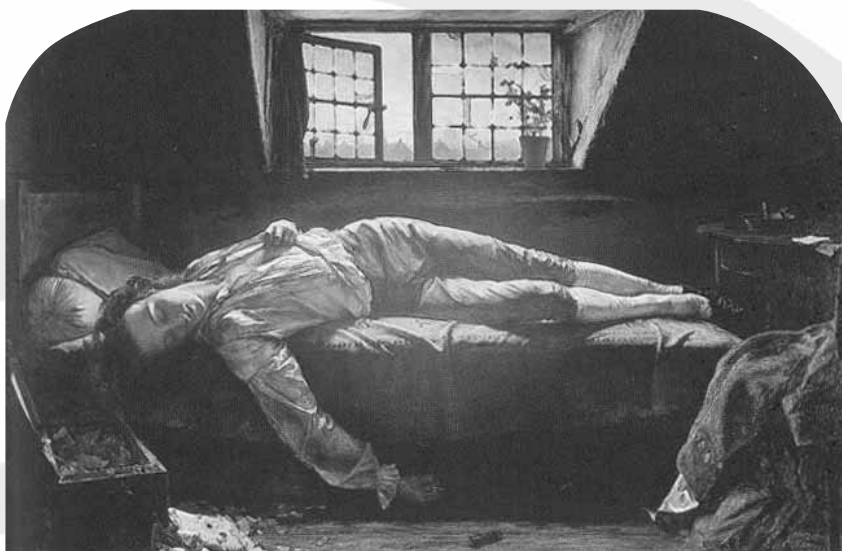
us not necessarily an objection to a judgment. The question is to what extent it is life-promoting, life-preserving, species-preserving, perhaps even species-cultivating.'[3]

Let us consider an example of this idea.[4] Imagine a person who believes that a loving God created the world and everything in it – including herself. Such a person would think of her own life as a gift from God, and this belief would give her existence value and meaning. Now suppose that scientists found proof for the theory that all conscious life was created by a 'big bang', an event that has nothing to do with loving purpose or design. This theory may be more *true* than belief in God, but would it be *better*? Might it not be the case that scientific truths make life harder to bear, make people more unhappy, and undermine the value of our existence? As Nietzsche says, the truth may be that nature is 'wasteful beyond measure, indifferent beyond measure, without purposes and consideration, without mercy and justice, fertile and desolate and uncertain at the same time'[5] – but who would want to live their life according to such a truth? It is because this truth is so harsh and difficult to live with that people invent myths, stories, pictures, and a whole variety of different interpretations, to give meaning and value to existence. The question we have to ask when we choose to accept a particular interpretation is not, *is it true?* , but *will it make me stronger?*

Nietzsche's view that the truth may be undesirable suggests that he is not attempting to abandon the notion of truth altogether.   Rather, by questioning the *value* of truth he is undermining its claims to authority over us, and its claims to be absolute. For Nietzsche there is no such thing as *the* truth, objective and independent of ourselves; each person is entitled to their own truth, discoverable only from their particular perspective, but Nietzsche warns against the attempt to impose this truth onto others. He envisages philosophers of the future who, although they continue to be friends of the truth, 'will certainly not be dogmatists'.[6] We may, of course, want to question whether a truth that is valid only for a particular individual can really be called truth at all – surely in order to use the word 'truth' we need to use some criteria to distinguish it from falsehood?   And what, from Nietzsche's point of view, might such criteria be? One possible problem with Nietzsche's 'perspectivism' is that it is reflexive: that is, it refers back to itself and so cannot claim to be 'true' for anyone other than Nietzsche himself. However, Nietzsche would no doubt view this as a strength rather than as a weakness of his philosophy, and would reply to critics of his 'perspectivism' that 'my judgment is *my* judgment: no one else is easily entitled to it… great things remain for the great'.[7]

In his other books (notably *The Gay Science*) Nietzsche argues that, although science takes its rigorous pursuit of objective truth from the moral command to be always truthful, this scientific project has in the end destroyed belief in the God who once provided the authority that enforced moral values.   In other words, Christian teaching, taken to its rational conclusion, eventually undermines itself. This idea is summed up in Nietzsche's famous proclamation that 'God is dead… and we have killed him.'[8] Nietzsche intends this claim as a statement of fact, perhaps a prophecy, rather than as a cry of either triumph or lament. The death of God, indeed, represents both a danger and an opportunity. The danger is that the disappearance of the traditional source of value and meaning will give Europe the final push into nihilism; but this also provides the opportunity to create new values in place of the old religious ones, so that each individual assumes the role of a god by becoming the source of his or her own values. Since Nietzsche believes that truths and values have always been invented by human beings, there is something honest and courageous about this new era that dawns after the death of God, as if we are finally facing up to the way things are.
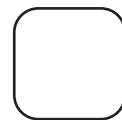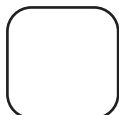
Good and Evil, Good and Bad

Nietzsche's main criticism of the ideal of truth is the same as his criticism of moral ideals: he dislikes their claims to be absolute. The philosophical model for absolute truth is Plato's theory of the Forms, which states that what is true must be unchanging, eternal and independent of the particular perspective, opinion and prejudice of the existing individual. Such a truth stands outside history, outside life itself, and applies to everyone regardless of their culture, language or personal circumstances. In other words, this kind of truth is objective and universal. Nietzsche argues that this ideal of truth is itself a fiction, a falsification of the ever-changing and diverse world we live in: 'it meant standing truth on her head and denying *perspective*, the basic condition of all life, when one spoke of spirit and the good as Plato did.'[9] According to his own arguments against the value of truth, this view would not necessarily be negative, but it does show truth to be inadequate on its own terms and makes the pursuit of truth seem rather ridiculous. More importantly, though, Nietzsche believes that a standard that applies to everyone cannot really be valuable at all, because achieving such a standard gives no distinction. For example, if everybody did A-levels and got 'A' grades, the qualification would have no value; if we told an artist that 'anyone could paint that picture', he would certainly be offended.

Nietzsche uses this argument to undermine morality, and to claim that absolute moral values lead to a culture of mediocrity and nihilism: 'how should there be a 'common good'! The term contradicts itself: whatever can be common always has little value.'[10]

The terms 'good and evil' and 'slave morality' refer to this absolutist morality, and Nietzsche contrasts the values of 'good and evil' with the values of 'good and bad' (or 'master morality'). Things are 'good' or 'bad' according to a particular perspective, in the sense that we say that something is 'good for me' or 'bad for me' (or 'healthy' or 'unhealthy'). For instance, a poisonous plant may be 'bad for me', and from another perspective it may play a crucial role in the ecosystem, but it would not make much sense to describe it as 'evil'. To say such a thing would imply that the plant was bad *in itself*, or objectively, whereas Nietzsche would argue that it is only particular perspectives that make value judgments possible at all.

A year after he published *Beyond Good and Evil*, Nietzsche wrote a book called *The Genealogy of Morals*, in which he clarifies and develops his analysis of the opposition between 'master' and 'slave' moralities. As its title suggests, this text offers a genealogy, or an historical account, of the evolution of moral values - a method which itself undermines the idea that such values are absolutely, eternally, objectively true and valid. In primitive societies, says Nietzsche, stronger people (the masters) will naturally dominate weaker people (the slaves), and so what is good and valuable will be determined by the powerful ruling elites. These value judgments come from the masters' sense of superiority, from their self-confidence and pride in their strength and talents; they would not expect weaker individuals to achieve or even to aim for their high standards, and they would not wish to impose their values onto others. Because the weaker, down-trodden section of society are powerless to take revenge on their oppressors by

means of force, they label the aggressive, arrogant ethic of their masters as 'evil' and 'unjust' ('bad in itself', when really they mean 'bad for me'), and preach values such as humility, meekness and pity. In effect, this 'slave morality' makes a virtue out of necessity, turning weakness into a moral value and expecting everyone to conform to it. According to Nietzsche, the ethical teachings of the Jewish law and, even more so, of Christianity, spring from this thirst for vengeance exercised by the weak upon the strong: 'moral judgments and condemnations constitute the favourite revenge of the spiritually limited against those less limited.'[11]

This interpretation of morality is often misunderstood, so we must clarify a few important points before we can begin to judge and debate it. Nietzsche's description of the evolution of morals may be more or less accurate, but it should be read not as a factual historical account but as a myth that emphasises that values change through time rather than being intrinsic to some sort of eternal 'human nature'. Indeed, as soon as Nietzsche introduces his distinction between master and slave moralities in *Beyond Good and Evil*, he makes it clear that these two types are often combined within one society, and 'even in the same human being, within a *single* soul.'[12] This helps to explain why morality, which is often assumed to be rational and coherent, usually turns out to be quite contradictory 'in the real world', for real human beings. We must also be clear that Nietzsche is not arguing for a return to the values of a barbaric, primitive culture; he recognises that the development of morality has created more civilised, more complex and more profound societies, and that this has empowered humanity as a whole. He admires not

so much the cruelty and violence of the 'masters', but the affirmative, creative origins of their values; and he is critical of the reactive nature of 'slave' morality rather than of particular virtues such as compassion and mercy – although he also emphasises that these virtues are hypocritical in so far as they conceal a desire for power and revenge. Just as Nietzsche argues that truth fails by its own standards, so he suggests that morality rests on a corrupt foundation.

Beyond Good and Evil?

Nietzsche attacks morality because he believes that it produces the values of 'the lowest common denominator' and imposes them on everyone, thus inhibiting creative individuals and preventing them from inventing the new values that could save Europe from nihilism. However, he also recognises that not everyone is strong enough for the task of creating values, and that weaker people need to be given rules and conventions to conform to. This view may seem reasonable enough, but how could it be put into practice? How can we decide which people these rules should apply to – and who decides?
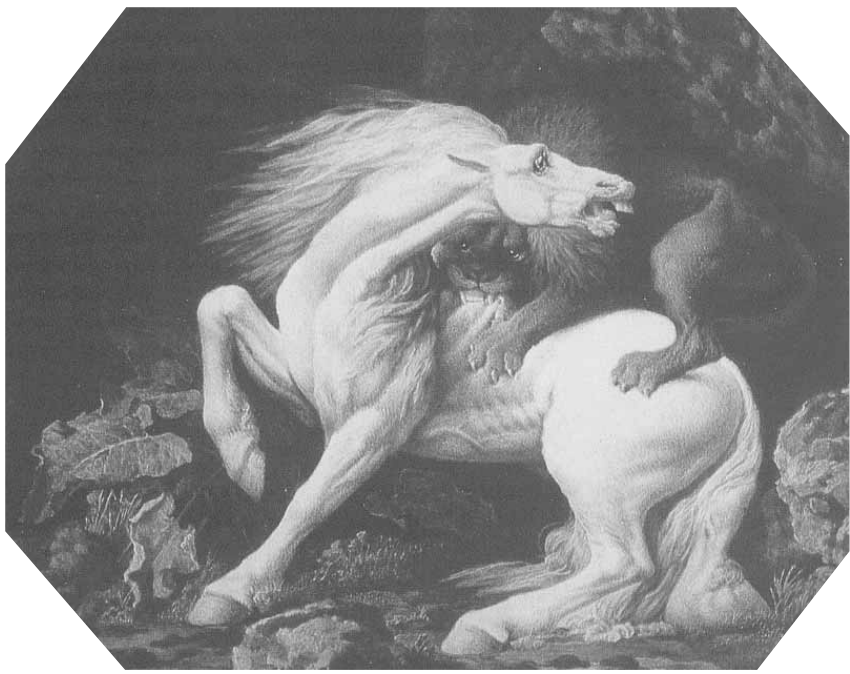
Does the exceptional, creative individual have to leave society? Is isolation the price of greatness – and if so, how is such a person supposed to change the destiny of European culture? Nietzsche leaves such questions open for us to struggle with, although he does offer a poetic presentation of these issues in *Thus Spoke Zarathustra*.

There is another set of questions surrounding the positive content of Nietzsche's philosophy. His attack upon absolute truth and moral values is much clearer and easier to understand than his account of the kind of values that are supposed to replace them. This is to some extent inevitable, for it would be contradictory to give someone precise instructions on how to be creative, and perhaps we should accept that Nietzsche's writing aims to inspire individuality rather than to prescribe it. There are, however, a few notoriously enigmatic suggestions about what lies beyond good and evil – most of which are discussed in Nietzsche's notebooks, which were published after his death under the title *The Will to Power*. These suggestions include the *Ubermensch*, or 'higher man'; the attitude of *amor fati*, or 'love of fate'; and the rather obscure doctrine of the 'eternal recurrence'. Since only the eternal recurrence is mentioned in *Beyond Good and Evil*, I shall end my discussion of the text by considering what this may involve.
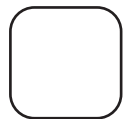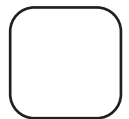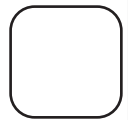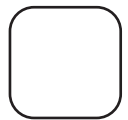
[Truth]

The idea of eternal recurrence was originally an ancient Greek philosophical theory, and it struck Nietzsche with the force of a life-changing revelation one day as he stood contemplating a large rock by the side of an Alpine lake. Instead of viewing time as a linear progression, the doctrine of eternal recurrence teaches that every moment repeats itself continually throughout eternity. This idea may be extremely difficult to grasp, but its significance for Nietzsche is relatively straightforward: if someone was faced with the thought that every moment of her life would recur eternally, could she accept this happily, without fear or regret? To do so, suggests Nietzsche, would be the ultimate affirmation of the value of life – the absolute opposite, in fact, of the longing for death and for another world beyond this one that, he claims, characterises Platonism and Christianity. Nietzsche describes the eternal recurrence as 'the ideal of the most high-spirited, alive, and world-affirming human being who has not only come to terms and learned to get along with whatever was and is, but who wants to have *what was and is* repeated into all eternity.'[13] We must add here that this expresses no naive optimism on Nietzsche's part: he himself endured great mental and physical suffering throughout his life, and he claimed that the joyful acceptance of eternal recurrence was possible only for those who had also faced the abyss of despair. It is this despair, indeed, that makes the total affirmation of life so great and rare an achievement.

Further Reading

Nietzsche, *The Genealogy of Morals* (Oxford: Oxford Paperbacks, 1998)

Nietzsche, *The Gay Science* (Cambridge: Cambridge University Press, 2001)

Nietzsche, *Thus Spoke Zarathustra* (London: Penguin Books, 1961)

Paul Strathern, *The Essential Nietzsche* (London: Virgin Books, 2002)

Michael Tanner, *Nietzsche: A Very Short Introduction* (Oxford: Oxford Paperbacks, 2001)

Arthur Danto, *Nietzsche as Philosopher* (New York: Columbia University Press, 1980)

Don Cupitt, *The Sea of Faith* (London: SCM Press, 1994)

Alexander Nehamas, *Nietzsche: Life as Literature* (Cambridge: Harvard University Press, 1990)

Notes

1 See, for example, Arthur Danto, *Nietzsche as Philosopher*, 68-99

2 Nietzsche, *Beyond Good and Evil*, section 1; section 16

3 *ibid*, section 4.

4 The following example is not Nietzsche's, but it raises issues that Nietzsche discusses frequently throughout his writing.

5 Nietzsche, *Beyond Good and Evil*, section 9; see also section 39

6 *ibid*, section 43

7 *ibid*, section 43

8 See Nietzsche's *The Gay Science*, section 125

9 Nietzsche, *Beyond Good and Evil*, Preface

10 *ibid*, section 43

11 *ibid*, section 219

12 *ibid*, section 260

13 *ibid*, section 56. See also *The Gay Science*, section 341, and *Thus Spoke Zarathustra*, Chapter Three, 'The Convalescent'.

Paul Sheehy

[Individualism]
in the Social Sciences

Introduction

The social sciences study the domain of human interaction – the social world. In much of our social scientific discourse, and indeed our everyday talk, we refer to social groups such as nations, families and gangs. We ascribe actions and properties to these entities. We sometimes explain why an individual is as she is because of her membership of a group; we may praise or blame groups; we may even seek to reward or punish groups. So we may talk of a nation possessing a property such as being 'war-like' or 'cultured'. We might explain why a person is honourable or wicked in terms of the effect a group has had on her. A nation may be praised for its courage or compassion or blamed for events. An instructive example is the debate over whether the German Nation was to blame for the Holocaust and whether it can be appropriate to hold accountable a nation. There is a need to clarify the sense, if any, in which groups are real entities.

That need is a response to two pressures. First, in our description and understanding of the social world the truth conditions of many of our statements depend upon the existential status of groups: put starkly, are there *really* groups which we pick out when we talk of them? A proper understanding of what is said - of what we mean - turns on how we are to treat the references to social groups. Talk about groups has the same grammatical features as our talk about, say, individual persons or tables. Yet should we understand groups, individuals and tables to be real in just the same sense? Second, the justification of moral evaluation, the formation of practical judgements and the articulation of policies and action depend upon the object of such judgements or action being an appropriate one. In particular it must be the kind of thing capable of sustaining the judgements made of it and of being responsive to actions directed towards it.

[Groups]
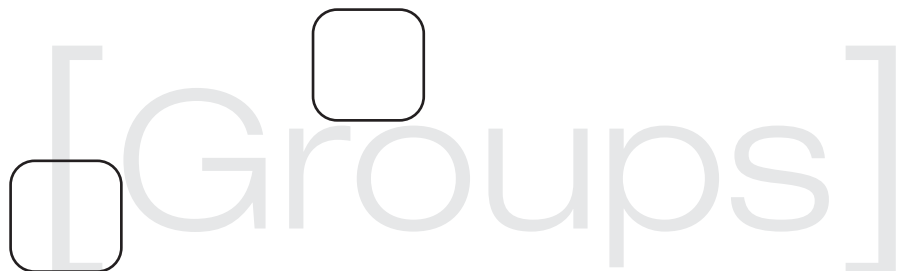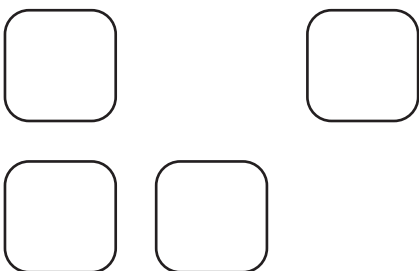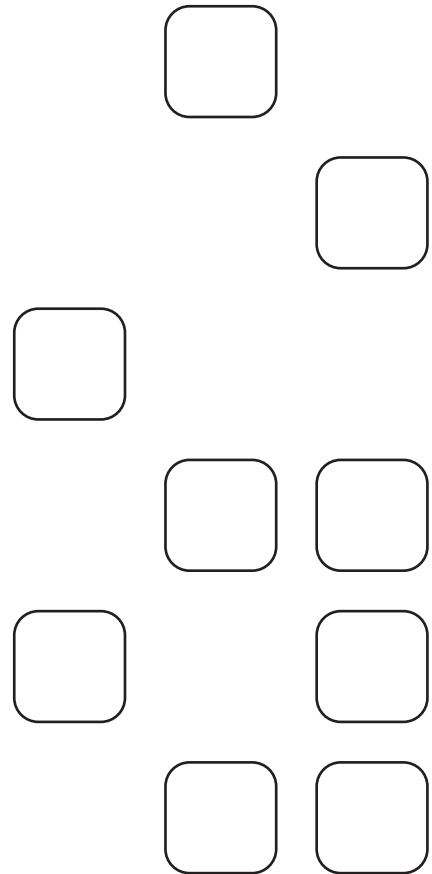
Let us regard the social sciences as seeking to provide true descriptions of the nature of the social world and its constituents, explanations of the phenomena within the world and a theoretical framework of generalisations and laws in which predictions of events or states of affairs may be made. In so doing the social sciences share the approach that quite in general characterises science. One understanding of the social sciences – of such fields of enquiry as sociology, psychology, economics and anthropology – is that they are ultimately the study of individuals. This understanding of the social sciences is individualistic and finds expression in the thesis of *methodological individualism*. The descriptions, explanations and theory-based generalisations are ultimately to be understood in terms of individuals and their properties. This top-down or micro-reductionist strategy promises to elucidate the nature of the social sciences and to account for the continuity of the social and natural sciences. First it would show that notwithstanding the appearance of our ordinary and social scientific discourse we should not think that social groups as such – as entities considered in their own right – are basic to our proper understanding and description of the social world. We could in principle dispense with talk of social groups in favour of talk of individuals. Second, the reductive approach – if applied globally across all fields of science - explains the continuity of the natural and social sciences because ultimately all fields of scientific enquiry can be reductively analysed in terms of the most basic or fundamental of all sciences, physics.

I shall explain that there is a serious challenge to such a reductive understanding of the social sciences.

We cannot reduce our references to social groups such as nations, families or gangs to references to individuals. The view that groups need to feature in social scientific discourse as entities in their own right is *holism*. If holism is true in the social sciences then it suggests that the model of science as unified through reduction to physics is one that ought to be rejected. Or, alternatively, we might preserve the possibility of the unity of science by abandoning the claim that the social sciences ought to be regarded as sciences.

Reduction

Individualism in the social sciences sees groups as composed of individuals and as being individuated through the predication of properties and facts about them. Individuals, though, remain ontologically and explanatorily prior to or more fundamental than groups. The properties of a group, the social generalisations or laws that are predicated of it and in virtue of which it is individuated, are reduced to the properties and generalisations predicable of its individual members. Reduction is an approach whereby one domain of things is shown to be absorbable into or dispensable in favour of another domain.[1]

The truths about groups are held to be expressible, without loss, as truths about individuals. The reductionist about groups accepts that there are groups, but that the science or body of generalisations in which facts about groups are explained can be reduced. Typically this takes the form of the reduction of the social sciences to the psychology of the individual (plus certain aspects of other relevant bodies of knowledge such as biology and ecology). Through this procedure

one domain is said to be reduced to the other. Examples of reductionist programmes include the reduction of numbers to sets, chemical properties (e.g. solubility) to the properties of molecules and atoms, mental properties to physical properties, and the laws of 'special' sciences (in effect all of the sciences apart from physics) to those of physics. Similarly, it has been proposed that social groups, properties and the laws of the social sciences can be analysed in reductive terms. Indeed Pettit has noted that the standard tradition of recent individualism takes the regularities of social science to be reducible to intentional regularities (i.e. mental states such as beliefs, desires, attitudes and so on), with the social-structural properties involved in social regularities being defined in terms of intentional psychology.[2] The appeal of reduction is held to be its ontological economy and conceptual unity in promoting explanations and descriptions couched in unifactorial terms. It may appear, moreover, to touch deep epistemological and ontological truths in revealing to us the gap between our ways of talking and the structure of the world.

Reductionism is associated with a global thesis that every domain can be reduced ultimately to that of physics. The belief that science can constitute a single unified project has come to be known as the 'unity of science'. Such unity has been presented as a 'working hypothesis' supposing there to be an hierarchical organisation of objects in which the objects at each level are formed through the complex arrangement of objects at the next lower level. We can thus envisage an increasing complexity of organisation as we move 'up' from elementary particles through atoms, molecules, living cells, multi-cellular organisms to

social groups. Oppenheim and Putnam were especially influential in developing the case for a reductive approach.[3] Advances in physics and molecular biology appeared to them to confirm the robustness of a reductive research program aiming to explain macro-phenomena in terms of their micro-structure (e.g. the reduction of thermodynamics to statistical mechanics, optics to electromagnetics). Associated with micro-reduction is the idea of theoretical reduction.[4] Roughly, this posits a hierarchy of theories. A theory about one level of objects is derivable from another theory about simpler entities and identities between entities of the reduced (higher level) theory and structures of entities of the reducing theory.

Each level is then subject to a programme(s) of investigation governed by the principles and practices of a particular domain of science aiming to uncover the principles and laws governing the behaviour of the objects at that level. A proper whole within the terms of discourse at a level N is reducible into proper parts in a universe of discourse at a lower level N-1. This reduction also consists in the derivation of the laws governing the behaviour of entities at each higher level from those governing objects at the next lower level. Such reduction calls for the knowledge of bridge principles or laws identifying kinds of objects at the higher or reduced level with arrangements or organisations of objects at the lower or reducing level. Through the transitivity of the reductive relations a unity of science is taken to hold, with the laws of the 'special sciences' being ultimately derivable form those of fundamental physics and the bridge principles.

[Groups]
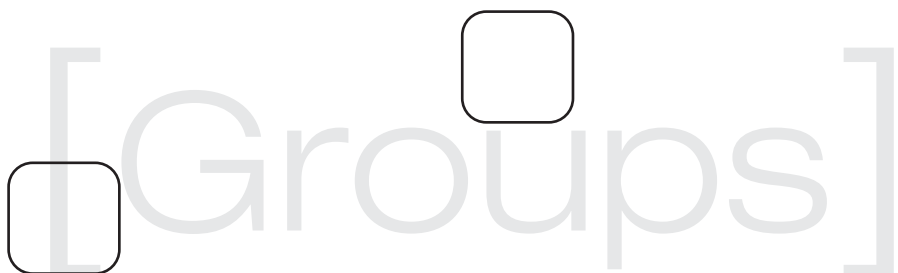
At bottom this is an empirical hypothesis and reduction is an empirical achievement resulting from the identification of suitable bridge principles.[5] We should note that the possibility of reduction is not hostage to the state of our scientific practices or the epistemological limitations of creatures like us. It could be a property of the world that one domain is reducible to another. Whether its reduction becomes part of our body of knowledge depends upon our capacities and our adopting a suitable perspective or theoretical interest to discover the reduction. I assume here that our knowledge and the construction of theories depends at least in part on a world independent of our theory construction, conventions and attitudes. Now, there has been no programme establishing universal and systematic connections between types of entities, properties and laws at different levels, all of which are ultimately connected through the transitivity imparted by bridging principles between adjacent levels.[6] This does not in itself deny the possibility of the reductions of one type of entity to some other type, but to point at the very least to the absence of any practical advance demonstrating that the types of one level are systematically connected to those of another. Sometimes we may find systematic links between kinds at different levels, but the practical endorsement of the global reductionist claim here *is* hostage to counterexample. Furthermore the model of a unified scientific project has been criticised as being thoroughly unsupported by consideration of and reflection upon our practices and standards of taxonomy and of the laws and generalisations employed in different scientific domains. Models of the world distinctly at odds with the

hierarchy, determination and predictability (in principle) of reduction include an ontologically promiscuous realism of countless (cross classifying) ways of ordering nature and a patchwork of laws governing local domains.[7]

I have sketched in the barest outline the thesis of global reduction because if it were to prove our best model of the world, then there would be a sense in which individuals are more fundamental than groups. Equally, though there would be yet more ontologically and explanatorily fundamental levels relevant to our understanding of the social domain. Nonetheless, the dispute between holism and individualism is perhaps best construed as a 'local' one, which seeks to cast light on the descriptive, explanatory and evaluative forms at a particular level of discourse. As such, then, the question is whether there is a persuasive case for a local reduction of social groups to individuals. Naturally, if we feel warranted in recognising the irreducibility of groups, so doubts will grow about the global reductionist programme.

We must remember that it is not the compositional claim at issue here. Groups are just individuals in a compositional sense; just as individuals are composed of their arms, legs, torsos, organs and so on; and these are composed of flesh, muscle, sinews, blood and bones; and these are composed of... and so on. The reduction of groups to individuals is the reductive analysis of the properties, facts, events, generalisations or (social scientific) laws through which groups are individuated as entities in their own right and in which groups figure. It is sufficient for (a weak) individualism to show that the properties of, facts about or generalisations or laws

applying to groups can be reduced to properties, facts or laws applying to individuals. Note that this neither requires nor entails a semantic reduction of group terms or predicates to individual ones. That is, individualism need not demonstrate that the meaning of group terms can be reduced to the meaning of individual terms. The success of reduction could be said to turn on whether it can show that groups lack causal and explanatory potency in their own right. The idea that for something to be real it must possess causal powers is clearly captured by Alexander:

> ...to suppose something to exist in nature which has nothing to do, no purpose to serve, a species of noblesse, which depends on the work of its inferiors, but is kept for show, might as well, as undoubtedly would in time, be abolished.[8]

Reduction must therefore explain the (apparent) causal and explanatory role of groups in terms of the properties of and generalisations pertaining to individuals as such.[9] A reduction of groups to individuals (or more precisely group properties, facts about groups, generalisations or laws within a theory about them) would take the form of an explanation of facts about the group in terms of the dispositions, beliefs, actions, resources, interrelations and situations of individuals. This reductive strategy is at the core of Methodological Individualism. Thus, a characteristic statement of the approach holds that,

> (E)very complex social situation, institution or event is a result of a particular configuration of individuals, their dispositions, situations, beliefs and physical resources and environment. There

may be unfinished or half-way explanations of large-scale social phenomena (say, inflation) in terms of other large-scale phenomena (say, full employment); but we shall not have arrived at rock-bottom explanations of such large scale phenomena until we have deduced an account of them from statements about the dispositions, beliefs, resources and interrelations of individuals.[10]

Methodological Individualism has been criticised as lacking precision in its formulation.[11] For the present, though, it is sufficient to see Methodological Individualism as a suitable framework in which a reductive analysis of groups could be pursued.

The individualistic analysis need not be framed in terms of the actual individual members, but in terms of 'ideal types' or 'anonymous individuals' who characterise the membership of that group. One constructs an ideal type by 'discerning the form of typical, socially significant dispositions and then by demonstrating how, in various typical situations, these lead to certain principles of social behaviour'.[12] The notion of an ideal type, introduced by Weber and employed by Watkins, is used to explain social phenomena in general. An ideal or hypothetical type is an abstraction from the personal preferences, different kinds of individual knowledge persons possess in a particular context, and typical relations between individuals and between individuals and resources.[13] The abstraction is conceived by Watkins to be to an ideal actor, probably without an empirical counterpart, in terms of whose attitudes, beliefs, dispositions, relations, and contextual setting a particular social phenomenon or fact

can be examined. Thus a range of social facts or phenomena could be analysed such as the process of capital accumulation within a market economy, or the tendency of a particular group to display certain properties such as the mercantile spirit of French Huguenots. Within this framework groups can be conceived as entities individuated through their instantiation of patterns of social behaviour and of particular properties. A reductionist programme would aim to explain facts and states of affairs about and associated with the group by correlating a kind of group, say trade unions or peoples, with a kind of individual. The facts about the group would be cashed out in terms of the individual's dispositions, attitudes, actions and relations with others.

This strategy to establish the ontological and explanatory priority of the individual will face problems if we are unable to establish the right kind of connections between the group or social types and the individual types. In the next section I shall explain that reduction in the social sciences is a flawed approach because the same kind of group can be realised through multiply various forms of interrelations, which are not amenable to abstraction to a single ideal kind or family of relations.[14]

Problems for Reduction

If reductionism is to proceed by systematically linking kinds of group with kinds of individuals or kinds of relations, which when individuals stand in those relations constitute such a group, the kinds of the higher level must be co-extensive with those of the lower level. In attacking the global reductionist programme Fodor defines the natural kind predicates of a science as those whose terms are the bound variables in its proper laws, and notes that if (global) reductionism is true, 'then *every* natural kind is, or is co-extensive with, a physical natural kind'.[15] The simple idea is that a natural kind represents a real division or cleavage in the world to which our schemes of categorisation must conform if they are to accurately report the ordering of things. A natural kind records a real distinction in natural or social worlds around which theories are constructed. Using an example of Fodor's, we might remark on the implausibility of reducing Gresham's Law to physics.

[Groups]
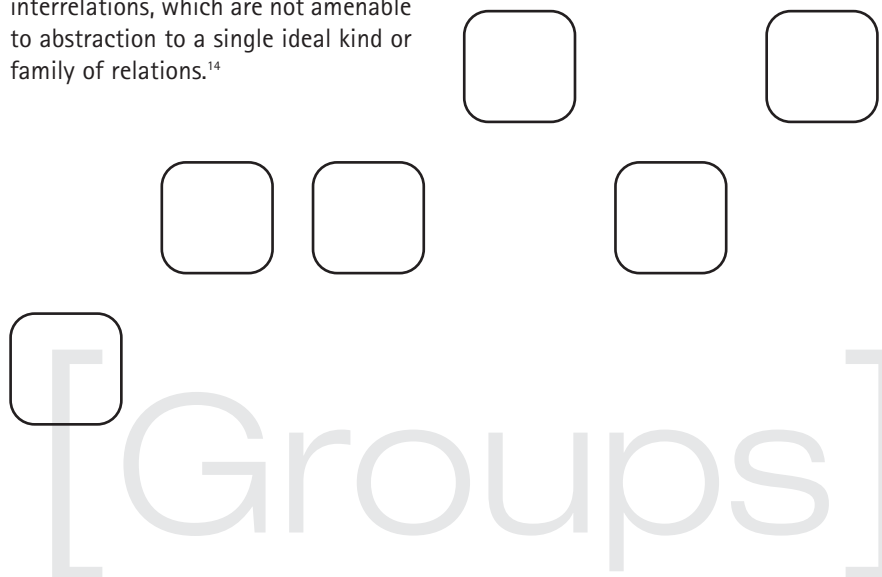
Gresham's Law makes a generalisation about monetary exchanges in certain conditions – bad money drives out good money. Those exchanges can take an indefinitely large number of distinct forms - the exchange of beads, pieces of paper designated as dollar bills, pieces of paper issued as cheques and so on. It looks unlikely that a disjunction of physical predicates covering all such events expresses a physical natural kind with which monetary exchange could be identified. The reasons Fodor suggests for thinking it 'intolerable' that every natural kind term of a special science must correspond to a physical natural kind term apply directly to the case for rejecting the identification of social or group terms with individualistic ones. Adjusting Fodor's original for the 'local' social scientific case we could say that the reasons it is unlikely that every social or group kind corresponds to an individual kind or kind of interrelations between individuals are that:

(a) interesting generalisations (e.g. counterfactual supporting ones[16]) can often be made about groups, their properties, relations and associated events whose individualistic descriptions have nothing in common;

(b) it is often the case that whether the individual events, properties or relations subsumed by such generalisations have anything in common is irrelevant to the truth of the generalisations, their interest or significance from the point of view of the science or perspective in question;

The social sciences attempt to establish (to some degree of approximation to practical and predictive usefulness) counterfactual supporting generalisations about groups such as the ruling class, the poor, tribes, families, gangs, work units, cultures and religious communities. Understanding a group may involve an analysis of the individual beliefs and values. However, generalisations about a (kind of) group can not always be reduced to generalisations about individuals, their relations, beliefs and practices. Consider religious groups or communities. The methodological individualist is committed to the view that religious groups can be reductively analysed in terms of the kind of individual or set of relations between individuals.

There should then be a correspondence between group type and types of interrelations individuals when we consider the diversity of religious groups and institutions. However, we may doubt that a social kind 'religious group' can be identified with any kind or ideal type of individual or pattern of relations. This is because the beliefs, interrelations and dispositions of the individuals who together constitute a religious group could vary considerably. For example, the kind of individual ideal type constituting a single kind of group could vary considerably. One set of individuals may be characterised by monotheistic beliefs, focusing on a principle of resurrection and by a disposition to defer to the authority of those individuals playing certain roles within the doctrine of the religion. This would contrast with individuals committed to polytheism, blood sacrifice and disposed to follow the dictates of the priest of their favourite deity. Of course they share the property of being members of a religious group, and as such their behaviour under certain conditions may be predictable. For example, members of religious groups may tend to respond to external threats in a way that gives the fullest expression to the central tenets of the religious doctrine.

This kind of response can vary in form and, moreover, it seems to be an explanation couched in terms presupposing the explanatory salience or autonomy of the group. Individualism faces a difficulty in establishing the priority of individual dispositions and beliefs so that the explanatory currency of the social sciences is coined in individuals alone. Moreover, if reductionism were committed to an identity of kinds it must regard a potentially indefinite
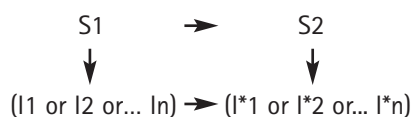
disjunction of kinds of individual beliefs, practices and dispositions as fitting the right hand side of statements such as:

x is a religious group if and only if I1 or I2 or... or In

where each of I1 etc. represent a kind of combination of individual dispositions, practices, beliefs and relations that constitute religious groups.

We can formulate generalisations such as 'the increasing formality of religious groups correlates with an increased bureaucratic sophistication of secular governance'. For example, the conversion of pagan kingdoms to Christianity in Saxon England may have been directly linked to the rapid growth of kingly power, centralisation and the revival (in part) of a sophisticated Romano-British taxation system.[17] The association of the secular authority with spiritual authority and the organisational infrastructure of the church enhanced the capacity of the secular ruling groups, because the nature of those groups underwent certain important changes through the conversion. On the basis of these kinds of generalisations social science can engage in counterfactual analysis. Historians and sociologists may find, for example, that generally it is true that whenever religious groups are characterised by a formalised set of practices secular government grows in its bureaucratic sophistication.

However, short of a complete enumeration of the ideal or constructed kinds that form religious groups we will not analyse the claims reductively, where reduction entails generalising from one disjunction of individual facts type-correlated with group ones to another disjunction of individual facts type-related to another group fact. This model of type reduction in social science would look like this where S1 and S2 are facts about groups, and I1 etc. facts about the beliefs, dispositions and actions of individuals:

$$S1 \longrightarrow S2$$
$$\downarrow \qquad\qquad \downarrow$$
$$(I1 \text{ or } I2 \text{ or}... In) \longrightarrow (I^*1 \text{ or } I^*2 \text{ or}... I^*n)$$

The problem here is not with whether there need be a strict nomological (law-like) link between the social and the individual kinds. Rather, and leaving that issue to one side, the criticism is that the reduction of a social or group level generalisation does not *explain* what is happening at the social level. The reduction just tells us that any one of a whole set of individual facts could give rise to or 'realise' the fact about the group via the generalisation between one set of individual kinds and another. To gather such individual facts together as a kind does not furnish any explanatory gain. In particular it does not secure an explanation of why the generalisation is a valid one at the group level. Now, to the extent this casts a shadow over the identification of kinds of groups, or kinds of facts about groups, with kinds of individual dispositions, beliefs and actions and facts about them, it is not yet an argument that can secure holism. Individualism can hold that groups are eliminable from our best explanations just because they do not function in any causally salient fashion, and so there are no generalisations about the social world in which reference to groups is anything other than metaphorical. Or, and perhaps more promisingly, individualism may accept the failure of type-type identity and embrace what we may call 'non-reductive individualism'.
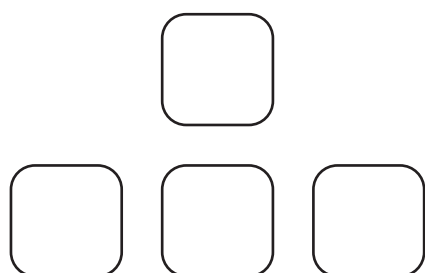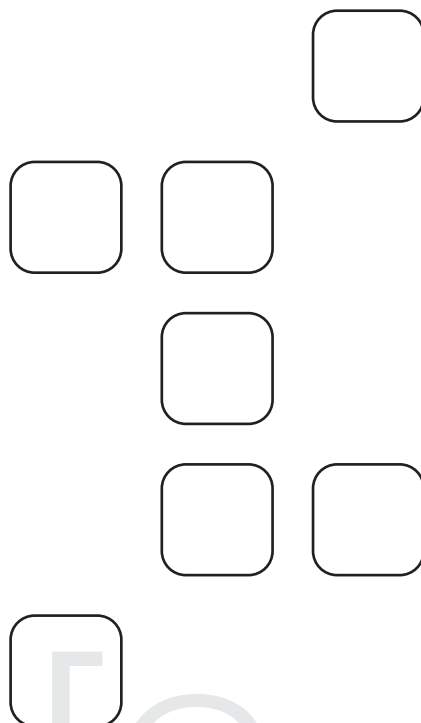
[Groups]

However, a discussion of how the individualist response goes must await another day. I have attempted to motivate a case against a reductive understanding of social scientific discourse. The importance of this issue goes beyond the philosophy of science, and as is often the case in philosophy we find that a range of concerns connect in perhaps surprising ways. For what we mean when we refer to social groups impacts upon how we should understand not just the sociologist, but what we mean in much of our descriptions, explanations and recommendations of, about and for the sphere of human interaction. In order to have knowledge of, and perhaps even some degree of influence over, our world we had better be clear about what kinds of thing there are in that world.

Notes

1. C.f. J. Kim in his 'The Myth of Non-reductive Materialism', *Proceedings and Addresses of The American Philosophical Association* 63 (1989) pp. 31–47 and 'Downward Causation and Emergence' in Beckermanm, A., Flohr, H., Kim, J., (edd.) *Emergence or Reduction: Essays on the Prospects for Nonreductive Physicalism* (Berlin: de Gruyter & Co., 1992).

2. C.f. P. Pettit, *The Common Mind (With Postscript)* (New York: Oxford University Press, 1996) p.145

3. The *locus classicus* of this notion of 'micro' reduction is P. Oppenheim and H. Putnam, 'The Unity of Science As A Working Hypothesis', *Minnesota Studies in The Philosophy of Science* II (1958) edd. H. Feigl et al.

4. c.f. J. Kemeny and P. Oppenheim, 'On Reduction', *Philosophical Studies* 7 (1956) pp. 6–19

5. C.f. D. Mellor, 'The Reduction of Society', *Philosophy* 57 (1980) pp. 51–75 (esp. pp.51–2).

6. For arguments against unity as a working hypothesis see e.g. J. Fodor, 'Special Sciences (Or The Disunity of Science As A Working Hypothesis)', *Synthese* 28 (1974) pp.97–115; J. Dupré *The Disorder of Things* (Cambridge, Mass.: Harvard University Press, 1993).

7. By Dupré *op cit* and N. Cartwright (in e.g. 'Fundamentalism vs. The Patchwork of Laws', *Proceedings of The Aristotelian Society* (1994) pp. 279–92) respectively.

8. Alexander *Space, Time and Deity*, vol II (London: Macmillan, 1920) quoted in Kim (1992) p.134.

9. For generalisation one can read 'ceteris paribus, special science or higher level law'. Unlike the laws of physics those of say biology, engineering and sociology are not exceptionless. Whether the laws of physics should be regarded as exceptionless or truly basic is itself subject to much debate (c.f. Cartwright *op cit*, Dupré *op cit*.).

10. J. Watkins 'Historical Explanation in the Social Sciences', *The British Journal for the Philosophy of Science*, 8 (1957) pp. 104–117. repr. J. O'Neill, *Modes of Individualism and Collectivism* (London: Heineman 1973) p.168. Watkins provides classic statements of the principle of methodological individualism. Also see his 'Ideal Types and Historical Explanation', *The British Journal for the Philosophy of Science*, 3 (1952) pp. 22–43. and 'Methodological Individualism: A Reply', *Philosophy of Science*, Vol 22 (1955) pp. 58–62.

11. For example D. Ruben *The Metaphysics of The Social World*, (London Routledge & Kegan Paul, 1985); A. Carter, 'On Individualism, Collectivism and Interrelationism', *Heythrop Journal* XXXI (1990) pp. 23–38.

12. Watkins ((1955) repr. O'Neill (1973) p.165)

13. Watkins ((1952) repr. O'Neill (1973) p.144)

14. The reductive approach will also be undermined if it turns out that the best explanation of the nature of the individuals and their relations is cast in terms of the influence on them of the groups to which they belong. For a discussion of the explanatory indispensability of social groups see e.g. P. Sheehy, 'Social Groups, Explanation and Ontological Holism' *Philosophical Papers*, forthcoming July 2003.

15. Fodor *op cit* p.102.

16. A counterfactual conditional is one in the form: if it were the case that p, then it would be the case that q. For example: if Oswald had not shot Kennedy, someone else would have. Or, if the German nation had not been ruled by Hitler, it would have been ruled by another anti-Semitic nationalist.

17. For a discussion of this period see e.g. J. Campbell, *Essays in Anglo-Saxon History* (London: The Hambledon Press 1986).

Joaquim Siles i Borràs

Evil, Freedom and [Responsibility]

An essay on Kant's Moral Philosophy

The notion of evil has become common currency in political speeches, newspaper articles and everyday discussions. Nonetheless, this is not a new debate. The question of evil has haunted philosophers for centuries. Questions such as 'why do human beings commit evil acts?' or 'how does evil happen?' are some of the questions that philosophers have been asking for a long time. Immanuel Kant, a German philosopher from the XVIII Century, dedicated much time to thinking some of those questions through. Kant's moral philosophy is usually introduced as made up of duty and law, but, as this essay will show, it is also about evil and responsibility. The present essay attempts to give a brief account of Kant's thoughts on evil, and the necessary relationship that such a notion maintains with the questions of freedom and responsibility. The aim of this discussion will be to show that Kant's moral philosophy sees morality as the most original dimension of the human being, and how, consequently, a philosophical understanding of what it is to be a human being requires an exploration of this moral dimension. For Kant, evil is not a question of religion and of the divine, but of reason and of the human.

Evil and the Question of Freedom

In 1792 Kant wrote an essay called 'On The Radical Evil in Human Nature'.[1] The first remarks on evil in this essay are fundamental. He says: 'we call a human being evil [...] not because he performs actions that are evil (contrary to law), but because these are so constituted that they allow the inference of evil maxims in him'.[2] In other words, there are what Kant calls evil maxims underlying evil acts, and what makes a human being evil is our choice of evil maxims or principles; i.e., one is not evil because he or she performs evil acts, but because one chooses evil maxims.

What Kant seems to be saying with this argument is that evil is not innate in human beings as a species, and that, instead, it is we who bring it on ourselves by choosing evil maxims, which, in their turn and by default, lead us to commit evil acts. But whilst Kant maintains that evil is not innate in human nature, at the same time he suggests that we do have a natural or innate disposition towards the adoption of maxims (good or evil) that guide our actions. It is in human nature to choose maxims or principles before acting, rather than simply acting without awareness of what and why we are doing what we are doing.

[Evil]
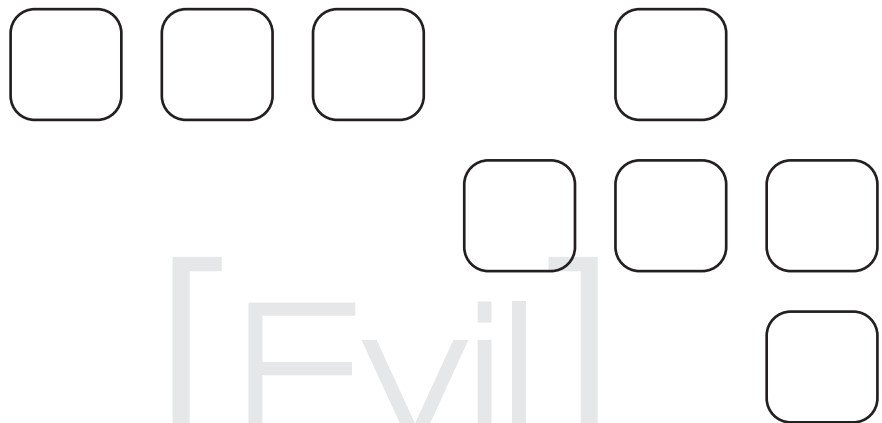
To choose maxims prior to acting means that we decide first and then act according to our decisions. Nonetheless, and although it is in our nature to choose and adopt maxims, whatever maxim we choose depends entirely on our freedom. Kant, therefore, posits evil with our decisions because they are free, instead of with our acts which are the result of such decisions. Freedom is, then, the exercise whereby we constantly adopt moral maxims through which we act and live our lives. It is in the free exercise of adopting evil maxims (rather than good moral ones) that evil comes about. Evil, in other words, is the result of our own freedom. But how and why do we choose evil maxims rather than always choosing 'good moral maxims'?

Even though Kant argues that it would be impossible to try to explain why our freedom makes us choose good or evil maxims, he gives an account of how we choose evil over good and how evil comes about.[3] Kant begins this discussion by saying that there is 'an original predisposition to good in human nature'.[4] Nonetheless, and at the same time, he also says that human beings have a 'propensity to evil'.[5] How can we have both, and what is the difference between predisposition and propensity? We shall deal first with the question of the predisposition to good.

Kant thinks that our natural predisposition to the good has three levels. The first level belongs with the dimension of what Kant calls the 'animality' of the human being; i.e., the human being as a physiological living being.[6] In this physiological sphere, we have a propensity to self-preservation, to propagation of the human species (through the sexual drive), and to live in community with other human beings through our social

drive. Kant calls the predisposition to animality a 'mechanical inclination' to our physiological needs, or a 'self-love' (for myself, my species or my family), which, insofar as it is mechanical or merely instinctive, does not require reason.[7] Whilst Kant stresses that the non-rational nature of these mechanical inclinations does not necessarily mean that they are negative, he says that such inclinations can be perverted and lead to 'the bestial vices of gluttony, lust and wild lawlessness (in relation to other human beings)'.[8]

The second level of predisposition is a predisposition towards 'humanity', this being 'an inclination to gain equal worth in the opinion of others' by means of comparing ourselves to others.[9] We could understand this as a socio-political predisposition to gain respect from others and even a respectable position in society in relation to others. Even though reason is a requirement for a socio-political life (for we need to reason in order to compare ourselves to others), here reason can be at the service of our socio-political purposes. As a result, Kant emphasises that the predisposition to humanity may also lead to 'unjust desires' such as jealousy, rivalry and even hostility to other human beings, and, in extreme cases, can even lead to 'diabolical vices' such as envy or joy in others' misfortunes.[10] But if these two predispositions in our physiological and socio-political dimensions of the human being are good in themselves, as Kant seems to be arguing, what makes them become vices?

The answer to this question seems to reside in the third dimension of the human being; i.e., the predisposition to what Kant calls 'personality'.[11] This is the human predisposition to respect what is morally good or what Kant

calls 'the moral law' (I shall return to this later). Personality is the level at which the human being is not a mere physiological and socio-political being, but a moral being. The human being is not only a being that has an instinct to physiological and social self-preservation and that can use its rational faculty strategically to gain respect and equality in society, but is also a moral human being that has the predisposition towards the acceptance of the moral law. The acceptance of the moral law also implies speaking of reason, with reason now being an end in itself rather than as a means to acquire our social aims (as it could be the case in the predisposition to humanity). What this means is that human beings do not only act according to their own interests and in order to reach their individual and social aims. Furthermore, they can also choose to act rationally, regardless of whether the outcome suits their interests or not. When the human being chooses to act rationally, then he or she is acting according to what they ought to do, according to their moral duty.

To posit the predisposition to respect the moral law at the highest of the three levels is not to say that we must obliterate our predispositions to animality and humanity (because they could lead us to horrible vices) in favour of our moral predisposition. The human aim is not to live solely on a moral pedestal. Kant is clear about this point when he emphasises that: 'all these predispositions in the human being [...] are predispositions to the good'.[12] In fact, and even though we can use our predispositions to animality and humanity 'inappropriately', these predispositions cannot be 'eradicated' for a human being is the combination of the three predispositions and would not be a

Kant's argument so far shows that we can only speak of evil within the framework of freedom. The freedom of my will resides in the fact that I may choose to obey the moral law in the same way that I may choose evil maxims and let myself be driven by my individual, social or political interests. Nonetheless, such an original freedom also means that one is equally aware of the morally good and the morally bad and that we are aware that we are acting in accordance with the moral law or not. This implacable freedom is what makes the human being be a moral being. The question of freedom, thus, occupies a key role in Kant's moral philosophy. If we are to speak of evil at all, and if to act morally means anything at all, this is because the human being is free. Evil is evil because I am free to choose it. If I did not have the capacity to choose to act according to my interests and if I were only able to act according to the moral law, then I would not be free, I would not be a moral being and the moral law would mean absolutely nothing.

human being without the three.[13] Our physiological instincts to self-preservation and our political interests are not negative per se, but they 'demand compliance with the moral law'.[14] In other words, the source of evil does not reside in the fact that we, as humans beings, have animal and socio-political needs, but in the priority we grant to such needs and inclinations in detriment of the moral law. What Kant is saying here is that our morality, our predisposition to respect the morally good, must guide our predispositions to animality and humanity. When this is not the case, when we push the moral law into second place, and choose instead to act according to our physiological inclinations and socio-political interests, then we are choosing evil maxims and our inclinations turn into diabolical vices.

[Evil]

The Moral Law

But what does Kant understand by the moral law and what does it mean to obey it? Simply put we can say that Kant understands a moral law to be an a priori rational maxim, rather than what we could call an a posteriori maxim deduced from sensible experience.[15] When I formulate a maxim that is going to guide my acts, this can be either an a priori rational maxim that is to be universally valid for all human beings at any time or a subjective maxim derived from experience. If it is subjective and derived only from experience, then it will be a relative maxim, and, therefore, a maxim that is to be obeyed only on certain occasions. We can think of the maxim that says that it is permitted to kill on certain occasions. When this is the case, such a maxim cannot be considered a moral law, because it is not universally valid for all human beings in every place and time. When a maxim is only applicable to the one or ones who have formulated it or to everyone but only on certain occasions, then we can say that the maxim is not properly rational, for even though we have to employ reason in order to formulate it, reason is here a means to an end. In this case, we would be formulating a maxim that guarantees us our interests; we would be using reason in order to justify killing. When this is the case, we cannot speak of laws but only of maxims, or, more precisely, evil maxims. So, what is the requirement for a maxim to become a law, and, thus, a moral law?

Even though all moral laws emerge from a maxim formulated by a subject, the moral law is not solely deduced from my own experience, but is reasoned on behalf of all human beings. What is formulated, is formulated bearing in mind all human beings ever to exist: the ones I know, the ones I do not know and the ones I will never know. This is the difference between formulating maxims and formulating laws. The formulation of a law is the formulation of a maxim that I wish to be applicable to everyone at any time. Kant says that whenever we make a moral decision and we choose a maxim that is to guide our actions, 'I ought never to act except in such a way that I can also will that my maxim should become a universal law'.[16] In other words, every time I choose a maxim I should measure the maxim I am choosing not according to whether it fits my purposes and avoids certain undesirable consequences, but according to its universality; i.e., as if it would have to be followed by all human beings at any time. This is what makes a maxim a law, and therefore moral. A maxim is a law when it is universally valid, and it is moral because it is not formulated with regards to ends but because it is rationally good in itself. Thus, under no circumstance can I morally formulate a law that says that I can kill in such and such a case. If I decide to kill in certain circumstances, then I am formulating a maxim that serves my purposes here and now, rather than formulating a moral law. When this is the case, and according to Kant's argument, I am an evil human being.

Kant understands the moral law to be a command or what he calls a 'categorical imperative' that we must follow because it is our moral duty to act according to what is good for the whole of humanity.[17] As Kant says, 'act only on that maxim through which you can at the same time will that it should become a universal law'.[18] But to speak of the moral law in terms of a command that one must obey does not imply the loss of our freedom. To obey the moral law is not a mere act of obedience in which I follow a command emptily. Far from it, every time I choose to obey the moral law for duty's sake I am formulating the law myself or, in other words, I am legislating it. The moral law is not just inscribed in me before I can think but, on the contrary, is formulated as a result of my autonomous reasoning and my freedom. The human subject legislates every time she or he chooses. To obey means here to legislate a law for the whole of humanity.

Moral and Radical Evil

Now that we have dealt with the questions of the predisposition to the good, freedom and the moral law, we can return to the question of evil. According to Kant's argument, I either act in accordance with the moral law (what is good for the whole of humanity), and I do so because it is my moral duty towards all other human beings, or I subordinate the moral law and what is universally good to my own interests (whether these are aims, ambitions or fears). It is for this reason that the question of evil is important for Kant. For every time I choose not to obey the moral law and I obey instead an evil maxim, I am perpetrating evil against the whole of humanity. By adopting an evil maxim that says that I can kill in a particular situation, I am prescribing that it is good to kill on such occasions and, therefore, that to kill can be good and that, as such, we ought to do it. What I am therefore doing is condemning the whole of humanity to an evil principle. Evil, as we said at the beginning, does not reside in our acts but in the immoral maxims that we choose as principles on which our actions are grounded. This is why Kant says that evil is always 'moral evil'.[19]

It is at this very moment of the argument that Kant speaks of moral evil as 'the propensity of the power of choice [freedom] to maxims that subordinate the incentives of the moral law to others (not moral ones)'.[20] Whilst Kant believes that we have an original predisposition to the good (as we have seen above), at the same time he thinks that we have a propensity to evil. This propensity is the propensity to moral evil, and it consists in the free subversion of the moral law for other inclinations, aims or interests. Even in cases in which the subversion might appear to be necessary and even in line with what is considered to be legal, the subversion of the moral law leads to evil in a moral sense. Thus, we can imagine a case in which my country asks me to join the army. It would be my legal obligation to comply with and obey the state law. Simultaneously, I am aware that to obey such a legal requirement would contravene the moral law, for it is universally good that I should not kill under any circumstance, and therefore I should not go to the army to learn to kill. It is clear that, according to Kant's moral philosophy, we ought to say 'no' and stick with the moral law, regardless of any other reason or any punishment the state may threaten us with. Failure to stick with the moral law would make us evil human beings. To comply with the state law in this specific case would mean to renounce our moral obligation to the whole of humanity, no matter what reason we give. To put the whole of humanity in second place with regard to our own interests means, according to Kant's position, to commit evil against the whole of humanity.

It is because of this notion of the subordination of the moral law for our own interests that Kant says that moral evil is always 'radical evil'.[21] According to the argument followed here, the human being is constantly located at a junction whereby one must take moral decisions on behalf of the whole of humanity. This decision is always either-or and there seems to be no room for any middle ground. One could think now of the following situation. Confronted with the legal obligation of having to join the army, one could say 'no' to the legal demands (and thus comply with the moral law), but then choose the moral law not for its own sake (because it is good for the whole of humanity) but because going to the army would interrupt one's career or personal interests. In such a case, and even though one's personal interests coincide with the moral law, this decision is not morally better than simply joining the army against the command of the moral law. On the contrary, it would be equally evil.

For in both cases we are subordinating the moral law to our own interests. The human being making such a decision would be 'evil because he [or she] reverses the moral order of his [or her] incentives in incorporating them into his [or her] maxims'.[22] In other words, moral evil is always radical because what we do when subordinating the moral law to our own particular interests is to eradicate the moral law from the maxims that guide our actions.[23] The eradication of the moral law signifies for Kant that one freely 'corrupts the [moral] ground of all maxims' by uprooting the moral ground from the maxims that make us act one way or another.[24] The result of this is that our actions are left to be guided by naked interests. Moral evil is always radical evil, not because of the high degree of suffering that it brings, but because we make it the root of our maxims and, consequently, of our actions.
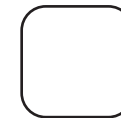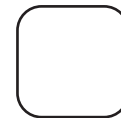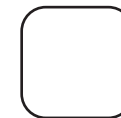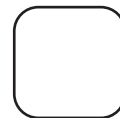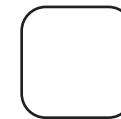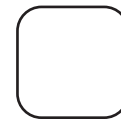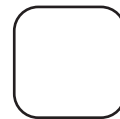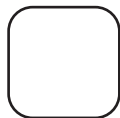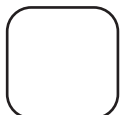
[Evil]

Responsibility and the Moral dimension of the Human Being

Kant's moral philosophy does not seem to leave us with any advice that would help us to live an easy life. Either we choose the moral law and, therefore, the whole of humanity, or we put our well-being above the whole of humanity and choose evil. Against this, one could argue that Kant's position is naïve, given that it is difficult to imagine a human being that is not or that has never been evil. At some point or other we are bound to disobey our reason and commit ourselves to our own interests, either due to the fact that we put our survival before everything and everybody else, or due to the interests of our whole nation or culture. Kant would not deny that. In fact, this is why he emphasises that even though we have a predisposition to the good, we also have a propensity to evil. On the one hand, evil is not innate in the human being, on the other hand we seem destined to bring evil on ourselves at some point in our lives. Nonetheless, this does not imply that evil is simply part of the human condition and, therefore, we must accept it as it is. This is an important point. To speak of the human propensity to evil does not mean to postulate that since we cannot avoid bringing evil upon ourselves, then our responsibility is only a responsibility 'in theory' and that it is justifiable to take some time off from our moral duties. Propensity to evil is not a pretext that authorises us to choose evil maxims when we wish, and thereby subvert the moral law for our inclinations, interests or fears. Far from it, propensity to evil (however unavoidable it might be for us human

beings) is precisely what situates us at a moral junction every time we must choose. Or, in other words, propensity to evil shows the depth of the human responsibility that Kant wants to emphasise. Without such a propensity, human freedom and responsibility would not have any depth. The avoidance of evil is ultimately in everyone's hands every time we have to make a moral decision that concerns the whole of humanity, and that is the responsibility that Kant wants every human being to face. Even though for Kant we have a propensity to evil, the commitment to evil is not justifiable in any case.

Whilst we must admit that this argument has a resonance with Christian belief, Kant's moral philosophy cannot be restricted within the limits of religion and god. Kant's notion of the propensity to evil does not simply assert the impotence of the human being and the ultimate need of god. By acknowledging that human beings cannot avoid bringing evil upon themselves, Kant is not suggesting that it is therefore necessary to put our trust in god in order to avoid evil. On the contrary, Kant aims to constrain religion within the boundaries of reason, of human freedom and, above all, of human responsibility towards the moral law

and the whole of humanity. As we have seen in this essay, the human being is a moral human being, and it is moral insofar as it is a wholly free being that has the capacity to choose rationally and autonomously. In other words, god cannot tell us what to choose, what is good or what is bad, for if that were the case then the human being could not be considered an autonomous rational being and, therefore, a wholly free moral being.

The notion of deep responsibility for the whole of humanity that Kant wants us to face cannot be reduced, however, to a mere question of 'good conscience'. One does not say 'no' to the legal demand of joining the army just because we aim to be satisfied with ourselves or because we simply want to avoid a bad conscience. If one chooses depending on self-satisfaction then this decision is not taken according to the moral law and what is good for the whole of humanity but only according to the consequence of the decision, and, therefore, according to what is good for ourselves. In such a case, one is still subverting the moral law for one's inclinations towards one's own individual preservation, and therefore, and following Kant's argument, one is still an evil human being.

This argument emphasises the rigour that Kant assigns to the moral dimension of the human being. This moral rigour resides not just in human freedom, but in the unconditional responsibility that every human being holds with every single moral decision. Every moral decision belongs to every one of us, and must be conditioned by absolutely nothing or nobody else. Emil Fackenheim captures this idea beautifully when he writes with regards to this moment in Kant's moral philosophy: 'Nothing in heaven or earth is more important than the

moment in which a man –any man– makes himself good or bad. And whenever a man makes such a decision, the universe, so to speak, holds its breath'.[25] All our moral decisions are of the highest importance. It is as if the future of the whole of humanity depended on every one of us and on every one of our moral decisions.

What is at stake in Kant's discussion on evil is the notion of the human condition and the ethical value of human life in all its moral radicalness. Kant's moral argument could be read as an attempt to make us think morality as the ultimate and deepest realm of the human being; as that which makes us fully human. This deepest realm is life at a constant moral junction in which every one of us must freely choose what we ought to do. To be a human being is, however, not an easy matter for Kant, for every time we make a moral choice, the future of the whole of humanity is held in suspense, awaiting our decision. Such is the depth of the moral dimension of the human being according to Kant.
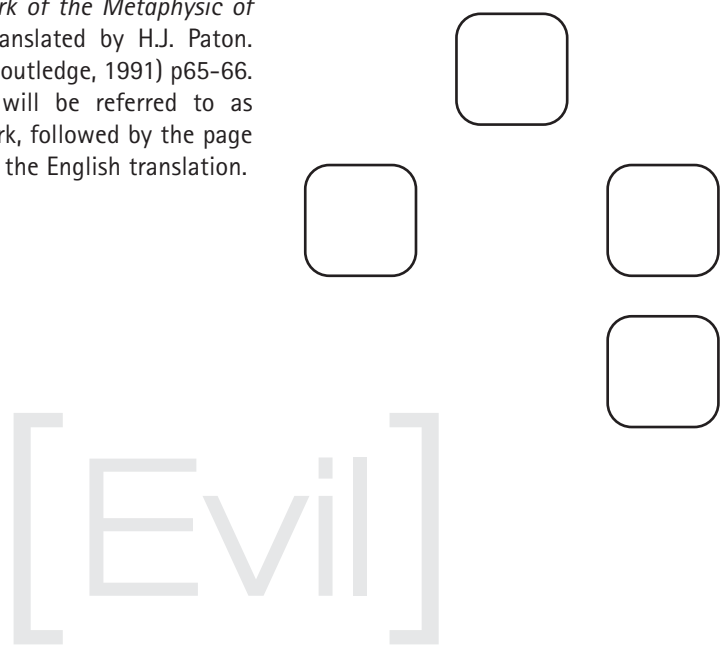
Notes

1   Immanuel Kant (1792) 'Of the Radical Evil in human Nature', published in: *Religion within the Boundaries Of Mere Reason and other writings* (Cambridge, translated by Allen Wood and George Di Giovanni, Cambridge: Cambridge University Press, 1998). This text will be referred to as Religion, followed by the page of the English translation.

2   Religion, p46.

3   See Religion, p47 footnote. Kant's argument here seems to be that if we could give an account of why

our freedom makes us choose good or evil, then freedom would not be the free power of choice, and, instead, we would be saying that freedom is determined by a previous ground. Were such the case, freedom would not be a free power but a ground conditioned by a deeper ground. See Religion, p47 footnote.

4   Religion, p50

5   Religion, p52

6   Religion, p50–51

7   Religion, p51

8   Religion, p51

9   Religion, p51

10 Religion, p51

11 Religion, p52

12 Religion, p52

13 Religion, p52

14 Religion, p52

15 Immanuel Kant, *The Moral Law. Groundwork of the Metaphysic of Morals.* Translated by H.J. Paton. (London: Routledge, 1991) p65–66. This text will be referred to as Groundwork, followed by the page number of the English translation.

16 Groundwork, p67

17 Groundwork, p68

18 Groundwork, p84

19 Religion, p54

20 Religion, p54

21 Religion, p56, p59

22 Religion, p59

23 We must bear in mind that Kant is using the term 'radical' in a literal sense. The term radical comes from the Latin 'radix radicis', which translates into English as 'root' or 'root of roots'. Evil, for Kant, would always have to be radical insofar as it eradicates the moral root of our actions and subverts it for evil maxims.

24 Religion, p59

25 Emil L. Fackenheim, 'Kant and Radical Evil': *University of Toronto Quarterly*, volume XXIII (1953-4), p353

[Evil]

Notes On [Contributors]

Joaquim Siles i Borràs

is a doctoral student at Manchester Metropolitan University, where he is working on the ethical dimensions of Husserl's phenomenology. He is also a part-time lecturer at MMU, where he teaches courses in phenomenology, ethics, and existentialism. He is currently working on a chapter of a book on Husserl's logic.

Clare Carlisle

studied philosophy and theology at Trinity College, University of Cambridge. Last year she completed her doctoral research on Kierkegaard, and since receiving her PhD from Cambridge she has taught philosophy at Manchester Metropolitan University. She is currently writing a book on Kierkegaard, and she will soon be beginning new research into the philosophy of religion at the University of Leeds.

Paul Sheehy

is joint editor of the Richmond Journal of Philosophy.

John Burgess

is professor of philosophy at the University of Princeton, where he specialises in logic, philosophy of mathematics, metaethics and 'pataphysics'. He acquired his PhD in logic and methodology from the University of California at Berkeley, before moving to Princeton in 1975. He has published widely, including (with Gideon Rosen), *A Subject With No Object* (Oxford: Clarendon Press, 1997), and articles on subjects such as Frege and modal logic.

Chris Hughes

is a lecturer in philosophy at King's College London. He first studied at Wesleyan University, before completing his PhD at the University of Pittsburgh. He has also taught at Cornell, and his interests include metaphysics, the philosophy of religion and mediaeval philosophy. His long list of articles includes pieces on the Cosmological Argument, bundle theory, Aquinas and scepticism.

Galen Strawson

is professor of philosophy at the University of Reading. His range of interests includes metaphysics, philosophy of mind, Kant, ethics and the British empiricists. He has published widely, including *Freedom and Belief* (Oxford: Oxford University Press, 1986), and *The Secret Connexion (Realism, Causation, and David Hume)* (Oxford: Oxford University Press, 1989). He is currently working on a book on the self.

Notes For [Contributors]

Content

We welcome articles on any area in philosophy. Papers may be broad or narrow in their focus (for instance a discussion of the mind/body problem, or an analysis of Hume's treatment of causation in the *Enquiry*). We would particularly encourage contributions which reflect original research on the following philosophical themes: epistemology, metaphysics, philosophy of religion, ethics, philosophy of mind, philosophy of science, political philosophy, religious ethics; and texts, such as: *The Republic, The Nicomachean Ethics, The Meditations, An Enquiry Concerning Human Understanding, Beyond Good and Evil, On Liberty, Existentialism and Humanism, The Problems of Philosophy, Language Truth and Logic.*

The articles should be around 3000–4000 words.

Style

The language used in the articles should be as non-technical as possible whilst preserving the richness of the arguments. Where technical terms are unavoidable they should be explained and examples offered.

Notes should be presented as endnotes. Textual references should be presented in the following format: Barry Stroud, *Hume* (London: Routledge, 1977), 77-91.

Presentation

Articles should be written in *Word* (any version).

Contributions

Articles for this journal are currently written by a panel of philosophers from a variety of universities in Britain, Australia and the United States, whose work is edited by the journal's editorial board. We invite unsolicited contributions from philosophers working in any field. The contributions should be submitted via email attachment to rjp@rutc.ac.uk

Copyright

The RJP retains the option of reprinting published articles in later RJP publications. Authors may republish articles with the journal's permission provided that they acknowledge that those articles were first printed in the RJP. Papers should only be submitted if the author is willing and able to be bound by the conditions set out in this paragraph.

[Contributors]

Richmond upon Thames College

is a large further education college located in
Twickenham, West London
offering 16-19 students one of the widest
curriculum choices in the country.

Last year, we came top of all
London further education colleges
in the Times league tables
and we are proud of our reputation
for achieving excellent results year after year.

We are well known nationally for
our high quality staff, excellent student support systems
and the inclusive education we offer to all our students.

If you would like to find out more about us -
please contact our Course Information Unit on
020 8607 8305 / 8314
or visit our website on
www.rutc.ac.uk

Richmond upon Thames College

[Subscribing]

to the *RJP*

The RJP comes out in Autumn, Spring and Summer. To subscribe you need to select the appropriate price from the table and complete the mailing information at the bottom of the form. All prices include post and packaging. The bottom half of the form should then be detached and sent with a cheque made payable to Richmond upon Thames College to the address below:

RJP Subscriptions
Philosophy Department
Richmond upon Thames College
Egerton Road
Twickenham
London TW2 7SJ
United Kingdom

Please allow one week for delivery in the UK, and two weeks for the rest of the world.

Annual Subscription : Current Rates

Institutional Subscriber in the UK	£33.00
Individual Subscriber in the UK	£18.00
Institutional Subscriber from the rest of the EU	63.00 Euros
Individual Subscriber from the rest of the EU	40.50 Euros
Institutional Subscriber outside the EU	$67.50 US
Individual Subscriber outside the EU	$45.00 US

Your name: .... Your organisation's name (if appropriate) ....

Address ....

 .... Post code/Zip code...

Telephone number (including full international dialling code) ....

Email address ....

I enclose a cheque for the amount of ....to purchase an annual subscription

Signed .... Date ....

[Philosophy]

RJP

The Richmond Journal of **Philosophy**

Richmond upon Thames College