

Philosophy Now

Series Editor: John Shand

This is a fresh and vital series of new introductions to today's most read, discussed and important philosophers. Combining rigorous analysis with authoritative exposition, each book gives clear and comprehensive access to the ideas of those philosophers who have made a truly fundamental and original contribution to the subject. Together the volumes comprise a remarkable gallery of the thinkers who have been at the forefront of philosophical ideas.

Published

Thomas Kuhn
Alexander Bird

John Searle
Nick Fotion

Charles Taylor
Ruth Abbey

Robert Nozick
A. R. Lacey

Robert Nozick

A. R. Lacey

PRINCETON UNIVERSITY PRESS
PRINCETON AND OXFORD

qua evil, since in one sense to pursue anything at all is to treat it as good. A sadist might pursue cruelty for its own sake, and cruelty is presumably evil, but if he said to himself, “Cruelty is evil but therefore for me it is good and to be pursued” we might wonder whether he was fully committed to the view that it was evil. Would he not be rather like someone who said, “It is false that grass is blue, but therefore for me it is true and I believe it”? The analogy is admittedly not perfect, because the grass case, but *perhaps* not the cruelty case, remains paradoxical if “therefore” is replaced by “in spite of that”; but it is close enough to be suggestive. Satan may have said “Evil be thou my good”, but apart from being fictional neither he nor his author was concerned with philosophical precision.

Nozick’s own answer is an expansion of the remark I quoted earlier that “value would inspire and motivate us under valuable conditions” (*PE*: 438). If value is to be organic unity it must inspire and motivate us, at least in general. When it does not we may put this down to distorting factors, but to avoid trivializing the case we must limit these to factors that are themselves low on organic unity. He presumably intends us to conclude, although he does not explicitly say so (see *PE*: 438), that envy is so, because it aims at destroying, and so distancing itself from, the value of that which is envied.

Organic unity as value

So value is (at least primarily) organic unity. But Nozick is not content, and goes on to ask whether organic unity is (i.e. amounts to) value, and if so why (*PE*: 441–50). Drawing on a theory he developed in Chapter 1, and which we shall meet later, the closest continuer theory of identity, he insists that organic unity must be not just the only candidate but close enough, if it is to constitute value. “Why isn’t the universe just dark”, as he puts it (*PE*: 441), containing lots of organic unity but no value? This is a question we might well want to ask. Why should organic unity be *good* in particular? If we say, *pace* Nozick, it is *because* it inspires or motivates us, this suggests that what is really valuable is some state of ourselves. In the version that the state must be one of feeling or experience this has been a common assumption of many philosophers, especially of those called hedonists. But there are powerful objections to it, since so many of our values are “intentional”, i.e. directed to something: we find it valuable *that something* (*often outside ourselves*) *be the case*, and the

value we find in our own states (of satisfaction etc.) often depends on this. Here, however, Nozick contents himself with pointing to the possibility we have already noted that there might be some quite different criterion of value, which dwarfed organic unity to insignificance, or even to zero, and was appreciated by “superior beings”. This recalls the reference to such possible beings in *Anarchy, State, and Utopia*, where I suggested his view of them was not radical enough, as they might think in categories that superseded not just organic unity but value itself; perhaps those now adumbrated would have just an intermediate degree of superiority.

Some criticisms

Nozick’s account of organic unity emphasizes both the unity induced and the diversity of the material in which it is induced. Hailwood criticizes this account because of its vagueness (1996: 171–4). One sign of this, he thinks, is that Nozick tends to emphasize the unity at the expense of the diversity, for instance when isomorphism (sameness of structure) is a source of organic unity. Such unity arises here when an abstract structure is realized in material, which therefore becomes isomorphic to the structure, and when several realizations of the same value occur (“value”, because values are treated as abstract structures: *PE*: 424). Hailwood (1996: 172) asks where uniqueness comes in, for the uniqueness of the individual is important for Nozick; his view implies that a collection of exact reproductions of a Rembrandt should have more value than the single original. Nozick explicitly denies that the *intrinsic* value would be affected in this case (*PE*: 425), while adding that a “whole that includes other new types of valuable things will have greater value [than the set of exact copies], because of the increased diversity this involves”. Hailwood could well ask how this fits with the “tightness” (something Nozick approves of) given by isomorphism. The vagueness here comes in the fact that Nozick has given two criteria for organic unity – unity induced and diversity of the material – but has not said how they relate together. (Compare the notorious vagueness of the catch-phrase of utilitarianism: “The *greatest happiness* of the *greatest number*”.)

Hailwood’s other main criticism of the vagueness of how Nozick treats organic unity, and of his emphasis on unification, concerns its effect on his political theory of utopia. A pervading theme of

Hailwood's book is to discuss how far Nozick can achieve the political neutralism that seems to be called for by the "live and let live" philosophy of *Anarchy, State, and Utopia*. He sees it as a "fundamental mistake" of *Anarchy, State, and Utopia* that its attack on patterns "is expressed in terms of a defence of libertarian voluntarism . . . If the aim is supposed to be to avoid the suppression of real plurality involved in the demand for a final solution, or tidy harmony, then the way to do this is not to assert, in the teeth of controversy, a set of universal natural rights" (Hailwood 1996: 158). One might at first wonder why not. One reason would presumably be that one man's rights are another man's duties, and so can be constricting, but Hailwood has more important reasons. (His argument is somewhat difficult and diffuse, and interpreting it is not helped by the book's infuriating lack of any kind of index, but I have been considerably helped by private correspondence, although obviously I am responsible for the use I have made of this.) Organic unity can be seen as "an account of value which supports imperialism via expressivism" (Hailwood 1996: 158), i.e. supports the attempt to impose a single best way of life on everyone via "the view which says that the state should promote a particular conception of the good" (Hailwood 1996: 3; cf. also 172). Nozick's emphasis on unity rather than diversity makes against the proper appreciation, which Nozick wants, of the value of individual uniqueness; there is a tension between organic unity and individual uniqueness. So we need an account of real neutralism, the problem of justifying which "without presupposing a substantive ideal, and thereby sacrificing neutrality with respect to rival ideals" Hailwood calls a "unifying theme" of his book (1996: 179). A view neutral between *all* views would have to be neutral between itself and its own denial, an obvious absurdity which Hailwood does not intend, so how shall we construct a neutralism that avoids this absurd extreme? In Chapter 3 ("Utopia II: objections"), I spoke of the framework for utopia as "still libertarian at the second level, as it were", and it is this distinction between levels that I think we need here. Libertarianism at the first level involves the apparatus of the minimal state, for Nozick, which we have discussed previously. But this libertarianism at the first level is only one factor to be counted along with others in the libertarianism at the second level. It is the tension between these levels that leads to the difficulty Hailwood emphasizes (as we saw in Chapter 3) in deciding what is the real outcome of *Anarchy, State, and Utopia*. As an example we can note, on the question of imperialism, that when Nozick comes to discuss rights later he

explicitly defends a right to personal liberty and to not being paternalistically forced towards the good (*PE* 500–501).

Another critic of Nozick's treatment of value in general and organic unity in particular is Ellis (1984: 452–4), who asks why we should expect *any* unitary account of value, whether of its structure or of its content. Nozick discusses the structure of value (*PE*: 422ff.), suggesting that values can be regarded as abstract structures, valuable things being realizations of those structures – but we will return to the ontology of value later. Taking values to be organically unified structures then lets him in this case avoid a foundationalist view of values, whereby there must be atoms of value from which all other values are constructed. This might well be an advantage since in epistemology foundationalism, common in the days of logical positivism, is now rather out of fashion (cf. *PE*: 414). The point would be that if an organic unity is one which has a value greater than the sum of the values of its parts taken separately, it could have value even if none of its parts had any at all. This point is suggestive, but of course not conclusive, and Ellis does not discuss it. What he does do is claim that Nozick ignores our need, and ability, to say *why* organic unity confers value in cases where it does. In aesthetics it does so because "our attitude to a work of art is that of a spectator", which would not explain why, say, acts of heroism are valuable. In society, organic unity has been thought valuable "because it was thought that living in an organic society was the truest expression of man's nature"; and in general organic unity "makes different things valuable for different reasons". We cannot therefore assume, he thinks, that organic unity will produce value elsewhere, even when it is present. Ellis's criticism here seems to come to this: there are all sorts of values of quite different kinds, and it may be rather an accident that organic unity appears as often as it does in the account of them. There does seem something in the charge that Nozick puts values into a straitjacket; he passes, for instance, from talking of the value of experiences to talking of that of persons without batting an eyelid.

The basis of value

Why is there any value at all, or rather, how can there be? We have already briefly met the question why the universe is not "dark" (see above, "Organic unity as value"), but Nozick develops his answer

(§ vi), where he offers what he calls “realizationism”, that values exist because we choose that they should, although their nature is independent of us; en passant he compares the dependent existence but independent character of children (*PE*: 556), and draws comfort, if no more, from the idea that on some views mathematical entities and Popper’s “third world” of abstract entities similarly depend on us for their existence but not their nature (*PE*: 556–7).

The prospects for this are not promising. Nozick leaves unclear whether the existence of values is chosen by God or us or both, and whether the choices we individually make commit others or only ourselves (which Hailwood (1996: 171), who makes these and other criticisms, calls “barely distinguishable in principle from choosing to live in an illusion”). Apart from anything else, if we choose that values exist, although not their nature, *what* exactly are we choosing? And does not our very act of choosing presuppose some value? Otherwise why do we bother? A more plausible view, although not obviously Nozick’s, is that only if creatures capable of choice exist can values exist (but one would surely not say there was no value or disvalue at all in a world populated only by animals). Moore (1903: § 50) held that a beautiful desert would have value even if no-one ever saw it or could see it. This is usually considered a rather extreme view. We saw above (“Organic unity as value”) that there are reasons against confining value to experiences. But more plausibly, other things that have value only have it because of some relation to experiences, whether because there are creatures that can appreciate them, or because they consist of (e.g.) happiness being distributed according to merit, or someone making a right choice – since only those who can experience can choose, one would think; if there were no experiences there would be no value. What might depend on choice is moral value: if there were no choices (or the capability of choice – but could this exist without *any* choices?) there would be no moral value perhaps. But we are getting too far from Nozick. I am tempted to conclude with Hailwood (1996: 171): “The less said about realizationism . . . the better for Nozick”; but before going on I will simply refer to the way Nozick brings in the idea of self-subsumingness, an idea that recurs as a motif in many parts of his philosophy. Here it applies to the way we pursue value by choosing to pursue value, because that very choice is itself valuable, something he then illustrates by comparing the way a placebo will cure us if, but only if, we believe it will (*PE*: 560–62).

Our question at the start of this section – why is there, or how can there be, any value at all? – suggests another: what is it for there to

“be” value? More grandly, what is the ontological status of value? Nozick raises the question in this form (*PE*: 424), meaning not “whether organic unities [i.e. values] are imposed on an inchoate and structureless world, or are discovered residing there”, but whether they are adjectival (things are valuable in having organic unity) or entities that we treat as existing parts of reality. Nozick’s answer, mentioned briefly above, is that values are abstract structures, a notion borrowed from model theory. But that still leaves the former question – what is the status of these structures? – and he returns to this question, at least implicitly (*PE*: 562): organic unities are “what value would be if there *were* value”. On the same page he treats the question whether there is value as amounting to, or at least on the same level with, whether there are ethical truths.

This suggests the question of objectivism: are ethical truths true, and are values real, in their own right independently of our attitudes (and of the differing mores of different societies, although this is more properly called absolutism, as against relativism)? To say that they are, denying subjectivism, is compatible with saying that they depend on our existence, in the sense that if we (or beings of some relevant kind) did not exist, neither would values or ethical truths – that is, nothing would be valuable and no ethical propositions would have any application. But Hailwood accuses Nozick of making two confusions, or at least conflation. The first is between objective value (independent of our attitudes) and intrinsic (as opposed to instrumental) value, and here Hailwood allows that Nozick may be simply following a tradition which assimilates claims about these two notions (1996: 130). (Hailwood also refers, privately, to Korsgaard (1983) for a further distinction Nozick, and Hailwood in his book, ignore, between intrinsic and non-instrumental.) More seriously, he thinks, is a confusion between objectivism in this sense about values and platonism (the view that values are substantial though non-material entities, rather like Plato’s Forms) (Hailwood 1996: 167–71). This is not implied by objectivism and Nozick does not need it, Hailwood thinks, and perhaps adopts it only because he thinks, wrongly in Hailwood’s view, that it is implicit normally in ordinary thought; it leads to difficulties for his view that we choose the existence of values. There is indeed a certain tension between Nozick’s treatment of values as “things” whose existence is itself valuable and his view of them as chosen by us, and the discussion of objectivity (*PE*: 728–9 n.45) treats it in terms of knowledge and agreement in a way totally independent of any platonist considerations.

The ethical pull

So far we have discussed the ethical push, which is the dominant element in *Philosophical Explanations*. The ethical pull is the moral claim that others exert on us in so far as they have a certain feature that he calls the “basic moral characteristic” (*PE*: 451). There may be more than one such characteristic – animals may have a different one from ourselves and still exert some claim on us – and Nozick allows a certain flexibility about sentience (*PE*: 730 n.51), but he concentrates on the main one. Three questions arise about it (*PE*: 451–2): what is it, what constraints on our behaviour does it imply, and why? Nozick proceeds to discuss these in order (although the third comes rather later).

The characteristic Nozick settles for is being a value-seeking self or I, which he thinks can embody the uniqueness that we seem to require when considering people; the characteristic is not just multiply instantiated, like having two legs, or even having desires; each person is not just *an* I, nor even just a unique I, but the particular I that he is. This is important when we want (as we presumably do) to avoid treating people as interchangeable. It is not enough to say, without further ado, “This policy will harm X, but it will benefit Y and Z”; we need to know that it will, in some other way perhaps, also benefit X. We cannot compensate Y for X’s loss (a point that was relevant in our discussion of compensation in Chapter 2). To ignore this, which *might* be justifiable when dealing with animals (although this could be disputed, with possible repercussions for Nozick’s choice of characteristic) would be to fall into a crude version of utilitarianism.

To seek value, however, is not for Nozick the same as merely seeking the satisfaction of desire, as an animal might do. To distinguish the two he borrows Moore’s “open question” argument (see Moore 1903: §§ 18–20), although using it slightly differently from Moore (*PE*: 731 n.57). It always makes sense to ask of something I desire for some feature it has, “But is it valuable? Ought I to pursue it?” Only if I ask this question, or rather if it is possible “to make the question seem real and salient” to me, even if I do not in fact ask it, am I a value-seeker.

On the basis of this Nozick claims that the basic ethical principle is to treat value-seeking I’s *as* value-seeking I’s, which he compares to Kant’s injunction to treat people as ends, never merely as means (*PE*: 462). But what does so treating them consist in? We could treat people as intelligent and having desires by hunting them and pitting our wits against theirs as they try to escape. But Nozick thinks that

this sort of behaviour (covered by game theory) would not really treat them as value-seekers in the above sense rather than as mere desirers (see *PE* 463 and 731n. 60). Perhaps not; but could we not refine our pursuit to take full account of them as value-seekers? Could not a certain kind of sadist (or as Nozick might say, of envious man) deliberately try to morally corrupt someone, just for the sake of doing so? Nozick’s own example is of killing value-seeking I’s, whether to get them out of the way or simply because you like doing so; here “the victim’s being a value-seeking I [still] does, qua value-seeking I, move and guide your actions as pursuer” (*PE*: 464). He therefore insists that we are only treating value-seeking I’s as value-seeking I’s if the content of our behaviour is shaped or controlled by their being so. This deals with the killing case but it is not obvious that it deals with the corruption case. Later (*PE*: 466) he distinguishes between treating the characteristic of being a value-seeking I as that characteristic and treating it as valuable. This might seem to deal with the corruption case. But one can anti-V value as well as V-ing it, and here it is the value that one is anti-V-ing, not just the characteristic that makes for value (although *PE*: 467, end of first paragraph, doesn’t bring this out very clearly). However, we saw above (“Value and disvalue”) that anti-valuing value as such has a certain incoherence about it, and Nozick says that anti-responding to a valuable characteristic “rubs against the grain of the characteristic” (*PE*: 467). Perhaps the best way of presenting a Nozickian position then would be to rely on the incoherence point to deal with anti-V-ing value as such, and treat anti-V-ing a valuable characteristic as involving a moral error, justifying this by saying that the characteristic would be valued in valuable conditions, of which the psychological state of the sadist is presumably not one.

Deontology and teleology: rights

So the basic content of the ethical pull is responsiveness to value-seeking I’s, with some allowance made for secondary cases such as the treatment of animals – an allowance, we may note in passing, which displays a welcome flexibility, but like that concerning value and organic unity means that we have not been given a strict analysis of the notion under discussion.

But anyway there are obviously plenty of questions left about how to act in particular cases. Nozick discusses what he calls the

“structure” of the ethical pull mainly by drawing on his early (pre-*ASU*) article on “Moral Complications and Moral Structures” (in *SP*). We need not pursue the details of this, beyond noting that he dismisses two structures, those of maximizing something (which got such a bad press in *ASU*) and deduction from first principles, since they tend to ride roughshod over the particularity and complexity of moral situations, and ends up with a complex structure balancing the right and wrong features of any proposed action and taking account of available alternatives (Principle III at *PE*: 488); this he thinks also takes account of the asymmetry of right and wrong shown by the need to compensate or apologize for wrong done even when that wrong is justified by an overriding right.

One feature of ethical calculations is their inconclusiveness, and here Nozick makes two points. First, “[T]he lack of procedure guaranteed to answer a question does not show the question has no correct answer” (*PE*: 482). He appeals to the results of Gödel and Church in number theory, that there are truths of any given system that cannot be proved within it, which does not make the notion of theoremhood indeterminate. The point has some force, but there is no clear line in ethics as there is in logic between what can and cannot be proved, if indeed the notion of proof applies at all in ethics. The second point is that in respect of requiring intuitive judgement ethics is no worse off than philosophy of science, where different theories must be compared in respect of “explanatory power, goodness of fit with the data, breadth and diversity of evidential support, degree of testability, range and diversity of the phenomena [the theory] covers, simplicity, fit with other accepted theories, and so on” (*PE*: 483). This seems much stronger, and we can no doubt agree with Nozick that ethics need not be denied all objectivity.

The account referred to above of the structure of the ethical pull goes back to pre-*Anarchy, State, and Utopia* days. But there is nevertheless a shift of emphasis between *Anarchy, State, and Utopia* and *Philosophical Explanations*, which is carried further in *The Examined Life*, as we shall see a little later. *Anarchy, State, and Utopia* was decidedly deontological in outlook, despite the reservations of Chapter 2 above (“Blurring the distinction”), but Nozick’s present view has softened enough to be presented “*despite* [his] obvious leaning toward deontology” (*PE*: 498; my emphasis). Referring to a notorious passage (*ASU*: 30n.) about violating side constraints to avoid “catastrophic moral horror” – notorious because of the casual way in which it is introduced in a footnote – he now says (*PE*: 495) he

had imagined teleology would take over there, but without any smooth transition to it or rationale for just when teleology became relevant.

After discussing various ways in which deontology and teleology might split the work between them of contributing to an ethical view, Nozick now constructs a view that combines them both, distinguishing two versions of teleology according as what is aimed at is the best consequence or the best action, but where the best action is no longer automatically defined as the one with the best consequences (*PE*: 497). Two actions with the same consequences may differ in value according to whether those consequences are all aimed at. Nozick gives the example of the “double effect” doctrine where an action that involves some evil differs in value according to whether the evil is intended as a means to some ultimate good or is merely a foreseen but unintended step towards achieving that good. Consider the difference between bombing civilians as a means toward terrorizing an enemy into surrender, where their deaths are essential to the aim, and bombing them as a foreseen but unwanted consequence of destroying an arms factory – one would be happy enough if they escaped first. Ironically what makes the former case morally worse is the greater organic unity it involves between the means and the end (cf. *PE*: 497–8). Could Nozick say that the overall organic unity was greater in the latter case (compare his treatment of the concentration camp example at *PE*: 419)? Perhaps, but it does not seem obvious and would require some arguing to show it. He does talk (*PE*: 498) of the “more intimate connections to value in action” as against that in pursuing valuable results. But will this necessarily exceed the organic unity in the terror-bombing case?

This allows Nozick to subsume certain deontological views under a maximization strategy, and thus under teleology, by extending that strategy to cover maximizing the value of one’s actions as well as, and sometimes instead of, that of the world. (The earlier discussion (*PE*: 475) where the maximization structure was dismissed already spoke of maximizing the score of an act as well as of the world in respect of some natural property (organic unity perhaps?), but presumably there the score of the act was still being calculated in terms of the result it produced.)

How does this compare with Nozick’s dismissal in *Anarchy, State, and Utopia* (28–9) of what he called a “utilitarianism of rights”? There he pointed to cases where one should not necessarily maximize the avoidance of violating rights, presumably even by oneself, which

would seem to mean, in his later terminology, that one should not necessarily maximize even the value of one's own actions. This need not, I think, be inconsistent with his *Philosophical Explanations* view because the point is that one should not necessarily act now so as to maximize one's own future avoidance of violating rights (e.g. by putting the development of one's future character before fulfilling some important claim on one, where they clash). In fact, the point is brought out again with explicit reference to one's own avoidance of violating rights (*PE*: 547–8), where he treats it as a side constraint rather than a maximization view, adding that “ethics binds us in the first person”. But it can still be a maximization view in the sense of the last paragraph.

Closely linked with deontology is the notion of rights, which formed the central notion for *Anarchy, State, and Utopia*. They are still important, and we saw above (“Some criticisms”) that Nozick defends a right to personal liberty against paternalism, as he did in *Anarchy, State, and Utopia*. But he emphasizes perhaps rather more the distinction between how others ought to treat us in responding to our status as value-seekers and what we can demand from them (and enforce) as of right without ceasing to respond to *their* similar status. It is important that we have a sphere of autonomy where we can exercise our personal liberty, but the limits to the extent of that sphere are now given greater emphasis, but with no discussion of what happens when rights clash in the way that in *Anarchy, State, and Utopia* led to the awkward difficulties about compensation. The general attitude to rights is now more relaxed. It is still true that they form a framework within the interstices of which other moral oughts must flourish, but it is now possible that they can be transcended, although Nozick seems to have primarily in mind the sort of spiritual development that makes some of a person's previous rights no longer relevant as a response to what the person now is (*PE*: 503–4). This may seem a rather abstruse point, but it does illustrate that rights no longer have the prominence they had in *Anarchy, State, and Utopia*.

The is/ought question

Ever since Hume, philosophers have puzzled over how statements about what we ought to do, or what ought to be the case, are related to statements about what is the case, or how we can get from an is to an

ought without committing what Moore (1903: §§ 5–14) called the “naturalistic fallacy” of assuming that an is could logically entail an ought. A similar puzzle, in fact Moore's form of the puzzle, concerns how to get from statements about what is the case to statements about what is valuable or good, and these two forms of the puzzle are commonly lumped together, a practice I will follow when convenient and not misleading, as Nozick himself does. Some writers have simply denied that the “fallacy” is a fallacy at all, but most seem to agree at least that the situation is not quite as simple as that.

Nozick does not enter the arena of the debate about the fallacy as such. After discussing briefly but inconclusively three possible relations between is and ought, he turns to a fourth, which accepts the gap, or chasm, as he calls it, but tries to explain it by seeking some property that belongs to one of the terms but not the other. The discussion is difficult because Nozick does not make very clear whether he is looking for something that holds of and so explains ought-statements as such (as *PE*: 538 suggests), or something that explains one ought-statement rather than a rival to it (as *PE*: 545 suggests). The difficulty is perhaps that what looks as if it will be a complete explanation in the first passage becomes only a partial explanation in the second. A further important point is that the feature we are looking for will only explain why something is true if it *is* true; it will not serve to prove *that* it is true. This emphasis on explaining rather than proving is in line with Nozick's new approach to philosophy that we discussed in Chapter 1, although he does eventually go on to what does underpin his favoured ethical principle, responsiveness to a value-seeking I (the realizationism we discussed above).

Before introducing the feature that is to do the explaining, Nozick gives an interesting list of nine other cases where one kind of statement cannot be derived from another; one example is metaphor and another the synthetic necessary identity that Kripke sees between light and electromagnetic radiation. He leaves open whether this latter holds between value and organic unity, but claims that none of his nine applies to is and ought.

The feature Nozick chooses is self-subsumption, of which we shall see more later. A principle subsumes itself if it falls under or is an instance of itself, but this gives no guarantee of its truth. Nozick gives “Every principle containing seven words is true” as an example of a false self-subsuming principle (*PE*: 541). Having Rawls in mind, he then offers as at least plausible that any principle that would be

agreed on by a majority in certain conditions ought to be followed. This principle itself might be agreed by a majority in the same conditions, although this does not follow automatically, since people might fear the tyranny of the majority, for instance (Nozick originally writes “unanimously” for “by a majority”, but then softens it to “by a majority” (*PE*: 541–2), presumably to make this last point easier).

How does this serve to do any explaining? Nozick points out that a moral principle that subsumes itself in this way follows deductively from that fact together with itself. But it also follows, of course, from itself alone (it is a triviality of logic that any proposition can be deduced from itself), so why should adding a superfluous premise help? Nozick elaborates an example, which we can omit, before saying more (at *PE*: 543–4). It now turns out that the explaining principle must also be “sufficiently deep”. This is evidently in part because it will then explain “many other more particular principles”, in the straightforward sense that they can be deduced from it, but also presumably because only then is it likely to subsume itself. “Pay your debts” does not subsume itself (it is a principle, not a debt). But what we are concerned with is explaining the self-subsuming principle itself, not the subordinate principles it straightforwardly explains. Nozick insists that the explaining principle must be not only “sufficiently deep” but “correct”, which as we have seen introduces a quite independent point. We cannot know whether a principle explains itself, by subsuming itself, until we already know that it is correct. If it is not, it will not explain itself or anything else although it will still subsume itself. So subsuming itself is not sufficient for explaining itself, any more than for its being correct. So why does it help? Nozick talks of it as stopping the principle in question from being a merely brute moral fact, where “brute” seems to mean just “unexplained”. But although subsuming itself does mark a principle out as being of a certain sort and not just any old principle, it does not seem clear that any explaining has been done.

However, Nozick is anyway not satisfied, because rival moral principles too, such as egoistic ones, might be self-subsuming, and for all we have said so far the correct moral principle, even though not merely a brute fact, has not been shown to be necessary, which it surely must be. Before going on to his realizationism Nozick toys with a possible Kantian structuring of ethics whereby “we structure the world so that the [ethical] statements come out true” (*PE*: 546; cf. 547). If this could work it would be a highly important development in ethics, but we need not discuss it further because Nozick rejects it

for three reasons. First, it is not easy to see how it works exactly. Secondly, it angles ethics too much towards the moral push rather than the moral pull. It makes our duties depend too much on the nature of ourselves rather than of the value-seeking I we are supposed to be responding to. This is hard to assess. Is the point that the pull should pull us however we structure the world? But it does not pull just anything (animals, for instance). But to say more would involve a long excursus into the details of the theory. Thirdly, it militates against our desire to track independently existing values (see Chapters 5 and 9 below), even though it lends some support to our desire for autonomy. Again it would take us too far afield to discuss this in detail, or ask how it relates to Nozick’s own view, whereon the nature of values but not their existence is independent of us, beyond saying that Nozick is well enough aware of the narrowness of the gap between being unwilling to pursue ends we have created ourselves and preserving our autonomy, yet he thinks we must still try to slip through it (*PE*: foot of 551).

Political implications: symbolic utility

What are the effects of all this on the political philosophy of *Anarchy, State, and Utopia*? We have already seen that Nozick has moved a great distance from there. The major role played by value is a large element in this change, and so is the partial reconciliation of deontology and teleology in the policy of maximizing the value of actions. Side constraints are still there in *Philosophical Explanations* (although not in the index, which should refer to pages 495, 497, 548, 734 n.74 at least); but they are less prominent and there is less emphasis on their exceptionless nature. Similarly rights play a lesser role and can sometimes be transcended. The emphasis now is more on responsiveness than on the mere avoidance of anti-responsiveness, so that the hard shell surrounding negative rights (rights in particular to non-interference) tends to dissolve. We saw above (“Deontology and teleology: rights”) how one’s spiritual development may affect the way others ought to treat one, so that one’s previous rights are no longer so relevant and can to some extent be transcended. One might wonder whether this could be expressed by saying that one simply acquires new rights based on responsiveness, but this might distort the development of Nozick’s thought if the rights are thought of as hard-shelled and absolute, as they were

earlier. The effect of the push and pull account of morality on Nozick's views about rights is discussed at some length by Hailwood (1996: 120–24, 178).

The political implications of this are far-reaching. The inheritance laws are modified (*EL*: Ch. 3), and the dreaded word “taxation” has now become respectable (*EL*: 288; in *Philosophical Explanations* he was still refusing to accept research grants funded by public taxation (*PE*: 523), and called such acceptances “wrongs”, even if “committed for a good purpose” (*PE*: 507)). In fact his defence of compulsory taxation (albeit allowing conscientious objection) as constituting society's “solemn marking and symbolic validation of the importance and centrality of . . . ties of concern and solidarity” (*EL*: 289) might have been intended, whether or not it was, to answer R. P. Wolff, who had said of the Nozick of *Anarchy, State, and Utopia*: “He portrays social interactions as marginal to the existence, integrity, and coherent identity of the individuals who participate in them” (Paul 1981: 95). “The bonds of concern for others”, Nozick goes on, “may involve not simply symbolically expressive and (it is hoped) effective policies through the general tax system, but also particular limitations of liberty concerning kinds of action” (*EL*: 291) (a most un-ASU-ish remark!) (The kinds in question are illustrated by discrimination.)

This reference to the symbolic marks a new entrant on the scene in the later ethics and politics, symbolic utility, which is only hinted at in *Philosophical Explanations* but is appealed to in the political revisions in Nozick's next book, *The Examined Life*, and discussed at some length in his last main book so far, *The Nature of Rationality* (*NR*) (especially 26–35). As the name suggests, something has symbolic utility not in virtue of its intrinsic nature, nor of what it causes, but of what it symbolizes. The most obvious examples are perhaps Freudian. If handwashing symbolizes freedom from guilt then an act of handwashing takes on the utility not of making physically clean but of making free from guilt (*NR*: 28). A more everyday example might be when on receiving some frightful white elephant from great-aunt Bertha at Christmas we say, “Oh well, it's the thought that counts”; and similarly there is the sentimental value we often attach to otherwise useless objects. Nozick also points to the importance of symbolic meanings in anthropological contexts (*NR*: 32).

The obvious question that arises about symbolic utility concerns its rationality: surely in some sense it is not “real” utility? Certainly symbolic and causal utility may clash to the disadvantage of the latter. It is irrational to insult one's employer to symbolically get at

one's father. The family affection displayed by great-aunt Bertha's white elephant may outweigh the small cost she paid at the jumble sale where she got it. Nozick in fact goes further: “A large part of the richness of our lives consists in symbolic meanings and their expression, the symbolic meanings our culture attributes to things or the ones we ourselves bestow” (*NR*: 30), and he goes on to add that because certainty often has a different symbolic meaning for us from mere high probability, symbolic utilities should be treated as a separate component in a decision theory (*NR*: 34).

We may perhaps conclude that symbolic utility has no direct causal effects in the physical world but it may well have causal effects in the psychological world, and we are psychological creatures as well as physical creatures, and inevitably so. However, there is one further question: if symbolic utility has purely psychological value, but still is an important part of our total value-scheme, how does this fit with Nozick's rejection, never retracted, of the experience machine and its virtual reality sophistications? (Nozick does indeed allow a limited recourse to the machine (*EL*: 108), as Hailwood has pointed out to me; but this seems a far more trivial matter than the large role allotted to symbolic utility in his later philosophy, perhaps, as I will say again shortly, for reasons connected with organic unity.) Nozick never discusses this, I think, but perhaps one answer would be that although symbolic utility directly has only psychological effects these are not confined to a single mind. One's symbolic actions can, qua symbolic, have real if only psychological effects on other people; and so far as the sophistications of the experience machine go (which allow one to have real effects on other people but only by cheating, as it were, in acquiring one's achievements) the objections to this, that it makes one's actions somehow inauthentic, have no bearing on the present questions about symbolic action.

Symbolic utility is one of the things Nozick says he had ignored in *Anarchy, State, and Utopia* when announcing the death of his libertarianism (*EL*: Ch. 25): “It neglected the symbolic importance of an official political concern with issues or problems, as a way of marking their importance or urgency, and hence of expressing, intensifying, channelling, encouraging, and validating our private actions and concerns toward them” (*EL*: 287). Although symbolic utility as such does appear to be a new entrant, as I have said, it may well be, as Hailwood has suggested to me, that Nozick intends it to fall under the rubric of organic unity, in that the relation of symbol to symbolized is a case of the sort of linkage Nozick requires. If this is true,

although Nozick never says it explicitly, I think, it would provide another instance of the role of organic unity as a unifying theme in Nozick's philosophy itself.

In the ten pages of *The Examined Life*, Chapter 25, Nozick does not of course develop the details of an alternative political system. But this lurch from libertarianism in the direction of communitarianism, despite its dismissal as "vague and ill-defined" (*PE*: 631n.), or perhaps even the dreaded socialism, shows how far Nozick has travelled from the *Anarchy, State, and Utopia* for which he is so famous.

Summary

Nozick's next book, *Philosophical Explanations*, exemplifies his philosophical switch from proof to persuasion, and asks, in the relevant chapter, how ethics is possible, shifting away from *Anarchy, State, and Utopia* in content as well as method, and introducing the notions of moral push and pull. Starting with the moral push Nozick asks what motivates morality, answering that the immoral person lives a less valuable life, and even if he does not in fact value living morally he would in valuable conditions. But it is not clear that this distinguishes moral from other values, and Nozick has been criticized for appealing to value to solve problems about morality.

Value itself Nozick thinks is primarily organic unity, i.e. unity in diversity – an important theme running through much of his work. Value has the function that we should "V" it (a generic term for various pro-attitudes), and V-ing value is itself valuable, a point he relies on in claiming to account for the positive nature of disvalue, which he admits has its own allure, but this he thinks is parasitic on that of value. He elaborates to some extent on the nature of organic unity, but how adequate it is for the role he gives it has proved controversial at best.

Critics have seen a tension between unity and diversity in Nozick's appeal to organic unity, e.g. concerning the value of uniqueness in a painting. Hailwood also sees Nozick's emphasis on unification as giving rise to the tension he sees (as we saw in Chapter 3) between libertarianism and the framework for utopia in *Anarchy, State, and Utopia*. Ellis asks why value should have any unitary account, and why organic unity confers value. He thinks Nozick unduly straitjackets value in his appeal to organic unity.

To the question how values can exist Nozick offers his "realizationism", that we choose their existence but not their nature, a doctrine left in some obscurity. What values are is abstract structures, although Hailwood sees a confusion of objectivism and platonism.

Turning to the moral pull, Nozick introduces the "basic moral characteristic" of "being a value-seeking I", which must be treated as such. This leads to a certain softening of his deontology and the admittance of some teleology, and even a maximization strategy, when we distinguish the value of actions from that of their consequences. His emphasis on rights is also somewhat relaxed.

After some preliminaries Nozick accepts the famous gap between "is" and "ought", and uses one of his favourite tools, self-subsumption, to help explain (not prove) moral principles (whether some as against others or morality as such is unclear), but admits this is not satisfactory, because it will not show that the principle is correct, let alone necessary (which it should be), and rival principles too might subsume themselves. Nor will a "Kantian structuring" of ethics satisfy him.

Our last section summarized Nozick's new approach, emphasizing responsiveness rather than rights, and discussed some of its political implications. It also discussed a new and important idea that will reappear in epistemology (see Chapter 6), symbolic utility. All in all, let us add, Nozick's ethics and politics have been more influential for his early challenging and detailed advocacy of libertarianism than for either his novel application of it to utopias or the notion of organic unity that so governed his later and both methodologically and ethically more relaxed approach.

Chapter 5

Epistemology

Introduction: internalism and externalism

It is generally agreed among reviewers that Chapter 3, on epistemology, is the best chapter in *Philosophical Explanations*. It is not very original. Its leading ideas were already put forward by Dretske in particular a decade or more earlier, as Nozick acknowledges in some detail (*PE*: 689 n.53). But it is with Nozick that the ideas really come to the fore, perhaps because he has developed them more extensively. The chapter has three sections, on knowledge, scepticism and evidence, of which the first two are very closely bound together.

We can set the stage by introducing a contrast – prominent in recent discussions of epistemology – between internalism and externalism. How does knowledge go beyond mere belief? (I say “mere” belief because nearly everyone agrees that knowledge, at least of facts as opposed to people, places, etc., does involve belief, although there are exceptions (see McGinn 1984, especially 547ff.). Clearly to be knowledge a belief must be true. We may think we know, or feel absolutely certain about, something false, but once we agree it is false we must withdraw any claim to know or have known it. (This is sometimes expressed by saying that “know” is a “factive” verb.) But more interesting is the question: what must be true about the knower himself? This raises issues about the role of knowledge in our lives. Why do we have the concept of knowledge at all? Do we want knowledge primarily for our own sakes or for the sakes of our relations with others? When we ask whether someone has knowledge, are we assessing him and asking whether he is a knowledgeable person, or are we mainly concerned with whether he is a reliable source of information for ourselves?

Internalists can be thought of as concentrating on the first question and externalists on the second. For the internalist what matters is the state of mind of the knower. If I am to know something I must be able to produce a justification for my belief. This justification must either logically entail what I am believing or else support it in some weaker although still pretty strong way, and it must itself be true. But the important point for the internalist is that it must be accessible to me – I must be able to produce it. Otherwise what would be the point of attributing the knowledge to *me*?

Externalists on the other hand, to put the point rhetorically, might rather phrase the question like this: what would be the point of attributing the *knowledge* to me? What matters, if my belief is to be knowledge, is that it be linked in some suitable way to what it is about. Whether I am aware that it is so linked is a secondary matter. The link may be of various kinds. An obvious case is causation. If my belief is caused by its object, or what it asserts to be the case, one might expect it to represent the object accurately – indeed how could it fail to do so, since a belief's object must exist or be the case if it is to cause the belief? A causal link is not the only link the externalists may choose, however, for the object may cause the belief in a roundabout way and rely on various extraneous circumstances that might not be on hand in other similar cases, so that the belief, although true, was only so by good fortune and could not be taken as reliable. The idea that what makes a belief knowledge is that the process by which it is arrived at is one that reliably results in true beliefs is called reliabilism. The process need not be causal, and the mechanism underlying it, if any, may be quite unknown. Consulting tea-leaves would do, if beliefs formed in this way turned out regularly to be true. (A case often referred to in the literature derives from a story by D. H. Lawrence about a boy who always managed to predict the results of races when he was sitting on his rocking horse. If you knew such a boy with an established track record, would you not consult him before placing your bets?) We shall see later that reliabilism itself may not entirely escape the chance of relying on good fortune. But whatever knowledge turns out to be, Nozick brings in organic unity again (*PE*: 417, 524), suggesting that knowledge has value because it unifies a person with a fact.

Many objections have been raised to each of these two approaches, some of which we shall meet as we go along, and various compromises between them can be constructed; but let us now turn to Nozick.

Outline of Nozick's theory

Nozick is an externalist. There is one passage (*PE*: 267; we shall see more later) that might be thought to show some hesitation on this, where he says, "it seems plausible that justified true belief is a necessary condition for knowledge; but I prefer to leave this question unsettled". He has just said that it is not a sufficient condition, because of the notorious counterexamples produced by Gettier to the effect that what makes it justifiable for someone to believe something may be quite different from what makes the belief true, so that, as we saw with the externalist causal theory, the belief might happen to be true but could not be relied on. But the reason why he thinks justification may be a necessary, even though not a sufficient, condition for knowledge seems to be (*PE*: 267 is not very clear here) that whenever one knows one will be using a reliable method of one sort or another, and this will carry justification with it as a sort of byproduct. This is different from the sort of motivation I attributed to internalism above, and he has just said (*PE*: 265) that he will follow "the externalized treatment of justification as reliability" (but cf. also *PE*: 281).

The theory Nozick offers was originally developed, he tells us (*PE*: 169), in connection with action rather than knowledge, a topic we shall meet in a later chapter. A causal account of knowledge sounds plausible, but in the sphere of action causation has been thought to threaten responsibility, and he seeks "a way for action to parallel belief, to be so connected to the world, even causally, in a way that is desirable", adding that if we succeed in this "determinism would be defanged" (*PE*: 170–71).

As we have seen, it has generally been agreed that for a subject *S* to know a proposition *p*, *p* must be true and *S* must believe *p*. Nozick agrees too, but he realizes that we cannot simply add a role for causation, because of the problem of deviant causal chains we saw earlier. His solution is to appeal to counterfactual conditionals, or counterfactuals, so called because the antecedent is presented as though it were false, whether or not it is so, and the consequent tells us what *would* then be, or have been, the case. The difference between counterfactuals and ordinary indicative conditions can be brought out by appealing to a famous example: "If Oswald did not kill Kennedy someone else did"; "If Oswald had not killed Kennedy someone else would have". Plainly these are not equivalent. The two conditions that *S* must satisfy if he is to know *p* are that were *p* not true *S* would not believe *p*, and were *p* still (in somewhat changed

circumstances) true *S* would still believe *p* (and would not also believe that not-*p*, i.e. would not contradict himself). These conditions Nozick calls the "variation" and "adherence" conditions respectively (let us label them "Variation" and "Adherence".) When a belief satisfies these conditions Nozick says it "tracks" the truth, a metaphor that has become a hallmark of this theory.

Such is the theory in bare outline, but before discussing it we had better introduce one or two complications. Variation is often regarded as the more prominent and important of the last two conditions, but there are at least two cases where Nozick thinks it is insufficient on its own. One he illustrates with an example from Harman (*PE*: 177). A dictator is killed and the media at first report this but then are made by the Government to deny it, so that people think the dictator is still alive. Smith, however, does not hear the later reports and so continues to believe, truly, that the dictator is dead. But does he know this? He satisfies Variation, for had the dictator not been killed he, like everyone else, would believe he was alive. But can he really be credited with knowledge merely because he failed to hear the later reports? In the changed situation where he did hear them he would have believed them (we assume), and so would have believed, falsely, that the dictator was still alive. He does not satisfy Adherence, and that is why he does not know.

The other case concerns necessary propositions, e.g. that twice two is four (*PE*: 186–7; I am simplifying slightly to avoid the next complication to be mentioned). Variation does not apply here, Nozick thinks, for it hardly makes sense to ask what would happen were twice two not four. But Adherence comes to the rescue. A child who believes that twice two is four merely because his teacher says so, but would not believe it were his teacher to say otherwise, does not have knowledge.

We have said nothing so far about how we come by our knowledge. Sometimes this does not matter, and our two main conditions as stated can represent the theory well enough. But sometimes they can lead to odd results if left unadorned. Nozick gives this example (*PE*: 179): "A grandmother sees her grandson is well when he comes to visit; but if he were sick or dead, others would tell her he was well to spare her upset". So she violates Variation, yet surely she knows he is well. So Nozick adds that we must take account of how she knows, namely by seeing, for in the case where she forms the belief that he is well even though he isn't, she would be using the method of relying on others, not that of seeing. If *p* is true then and *S* believes it via

method M, Variation should be rephrased to say that were p not true S wouldn't believe, via M, that p , while Adherence will say that were p still true (in somewhat changed circumstances) S would believe, via M, that p (I have adopted a small modification of Variation suggested by Luper-Foy in a difficult discussion (Luper-Foy 1987a: 225) where line 14 seems to misprint "not- p " for " p "; see also 1984b: 28–9). Luper-Foy sees problems with Adherence whichever formulation is adopted.

Nozick adds various other complications and refinements to deal with cases such as where S uses several methods at once to form a belief, and then runs through a set of problem cases from the literature which he claims his approach will solve. But we have now got a sufficient view of the theory to let us start discussing some of the issues it raises.

Counterfactuals

The notion of tracking the truth depends essentially on that of counterfactuals, which claim that something would be (or have been) the case if something else were (or had been). But what are we to make of such claims? The model Nozick appeals to is that of possible worlds, due mainly to R. Stalnaker and D. K. Lewis (although going back ultimately to Leibniz). Imagine all the logically possible scenarios laid out in a great array round the actual world, with similarities between them represented by closeness in the array and similarity to the actual world by closeness to it. Then the general idea is that a counterfactual such as "If p were the case, then q would be" (or for short, "Were p , then q ") will be true if the nearest worlds or scenarios to the actual one in which p is true (the nearest " p -worlds") are also q -worlds. The situation is more complicated than this mainly because there are infinitely many possible worlds and they probably form a "dense" set, i.e. between any two of them, however close, there are infinitely many others.

Nozick considers some of these complications (PE: 680–81 n.8), but mostly they do not affect the basic idea. One point is worth noticing however, namely the use of the plural ("worlds or scenarios") in the above statement of the general idea. Here Nozick differs from Lewis, for whom the nearest world to the actual world is the actual world itself, so that "Were p , then q " will always be true when p and q themselves are both true. But this trivializes the issue; we want there to be some connection between p and q , be it causal or of some

other kind. We don't want "Were grass green, snow would be white" to be true merely because grass *is* green and snow white. Again suppose that I am standing at the roadside when a bus sweeps past – I might say, "Had I stepped out then, I would be dead now". This will be true, on the theory, if in the scenario which is as close, i.e. similar, as possible to the actual world except that I step out, I promptly die. But it surely ought to be true in other scenarios too, such as where everything is the same and I step out but am wearing socks of a different colour. I would hardly comfort myself after my narrow escape by the thought that had I been wearing different socks I'd have been OK.

In the footnote just mentioned (PE: 680–81) Nozick deals with this sort of problem by appealing to a set of scenarios close to the actual world and having in common that p , or whatever is at stake, is true in these. This set is then called a "neighbourhood" – in this example the p -neighbourhood of the actual world – and "Were p , then q " is true just when q is true throughout this p -neighbourhood. The notion of a neighbourhood is given a somewhat intricate, but clear enough, definition and is then developed still further to deal with some of the complications mentioned above. These details need not concern us as Nozick himself feels "little inclination to pursue" them, and does not insist on a rigorous and polished account.

Counterfactuals are slippery customers. We all assert them cheerfully enough in our daily lives, but we would probably be embarrassed if asked to say just what made them true when they were true. They don't seem in any direct sense to describe the world as it is; indeed hence the talk of merely possible worlds. McGinn (1984: 535–6) thinks Nozick treats them as simply true (or false), without being grounded in anything, although Nozick (PE: 266) seems to presuppose an underlying, if unknown, "mechanism". Any appeal to counterfactuals to support a philosophical theory faces objections of two kinds: general ones like these about their nature and ones about the actual use made of them in supporting the theory. Let us start with the former (we will come to the latter in "Some criticisms of Nozick" below).

A rather scathing attack on the appeal to possible worlds is made by Fumerton (Luper-Foy 1987a: 167–9), who says, "the attempt to 'explicate' subjunctive conditionals [i.e. counterfactuals] using the possible worlds metaphor is at best disingenuous". He doesn't say what it is at worst. Nozick himself, as Fumerton goes on to point out, is well aware that we cannot, without circularity, analyze counterfactuals in terms of the closeness or similarity of possible worlds and