Looking Back Over 45 Years of Work

Jonathan Bennett

Philosophers can be known for their doctrines and their arguments for them, or for their methods. I am much less famous than some of the contributors to this book, but insofar as I am known it is for methods rather than for doctrines. I shall describe how I have approached five different regions of philosophy. These regions contain all of my eight books, and more than seventy of the ninety articles that I have published so far. (The others are on a variety of topics which I cannot describe here.)

Early modern philosophy

European-centered philosophy, and especially English-language philosophy, has been strongly influenced by early modern philosophy, the most prominent figures in which were Descartes, Spinoza and Locke in the 17th century, Leibniz later in that century and carrying on into the 18th, and Berkeley, Hume and Kant in the 18th century. A great deal of my work has been concerned with those philosophers: I have published four books and nearly thirty articles on early modern philosophy, and two more books are in preparation. This part of my work is controversial: some philosophers like my approach to early modern philosophy and have been helped by it to find their own way into this area; others

criticize it, sometimes expressing not merely disagreement with my approach but passionate hatred for it.

I confront the great philosophers of the early modern period as I would my own contemporaries: I treat them as real philosophers with whom I can enter into discussion and argument, and I do this with a view to making progress in my own philosophical thinking. For me the study of early modern philosophy is philosophy with a special technique, not history with a special subject matter.

Some philosophers these days hold that this is an inefficient way of doing philosophy: so much has been learned in the past few centuries, they say, that we would do better to read our own contemporaries, who stand on the shoulders of the great philosophers of the past. That might be right if I were looking to the early modern philosophers for a hand-out of valid arguments, true doctrines, and correct solutions to problems; but that is not how I expect them to help me. I can learn as much from their failures as from their successes: when a superb philosopher tackles the hardest problems in a manner that is bold, resolute and highly intelligent, there is much to be learned from how he goes about it, even if he gets most things wrong.

Some of those who devalue this technique for advancing one's understanding in philosophy are themselves early mod-

ernists, but for them this is a branch of history, not of philosophy. One of them, Daniel Garber, recently distinguished his work from mine by calling himself an 'antiquarian': for him the work of a past philosopher is an historical object, to be studied in the same spirit as anything else in the past. The history of philosophy, on this understanding of it, is on a level with the history of the makers of dictionaries or the history of historians. Many of the antiquarians criticize me for doing what I do; I do not criticize their kind of work, but I do not want to join them in doing it. Real philosophy is much more challenging, difficult and interesting than the antiquarian study of its history.

Some philosophically motivated workers on early modern philosophy have criticised my work in terms that sound like those of the antiquarians, though really they are different. To get any kind of help from a philosopher of the past, one must first understand what he wrote; and one may be helped to do that by knowing not only his major works but also the minor ones, his letters, the writings of his contemporaries, facts about the controversies in which he was engaged, facts about his personal life, and so on. There is a vast amount of scholarly work that can be done, and some of it may help with the central task of understanding the thought of the philosopher in question; and those who do such work have sometimes deplored my not doing more of it than I do.

Well, I do not deny that I may sometimes fail to understand a philosophical text completely or accurately because of some scholarly work that I have not done on its antecedents or its surroundings; but one may also fail in understanding of a text because one has not thought about it hard enough and long enough, has not read it often enough, has not brooded over it until one knows it almost by heart. Life is too short for all this; one must choose; and I have chosen to emphasize a strenuous and continuing struggle

with the principal texts, not giving much of my time and energy to scholarly research into antecedents, minor works, and so on. Those who make the other choice—attending more to the periphery, at the expense of the center—produce results that strike me as philosophically thin and superficial, not capturing the tough complex core of the great early modern philosophers. To get at that core one must engage with the philosopher in question as a colleague, as a living struggling thinker with whom one can argue about the issues, rather than as a fragile historical object who is interesting mainly for his place in history and the facts about whom he influenced and who influenced him. The 'scholarly' early modernists find my work irresponsible; I find most of theirs boring.

There is a personal reason for my coming to relate to early modern philosophy in the way I do. My philosophical studies through M.A. were done in a one-man department in New Zealand. The one man was the late Arthur N. Prior: he was a wonderful teacher and an even better role model; I loved him, and revere his memory. However, he was so overworked that he did not have time to learn very much; and when I left New Zealand to undertake graduate studies I knew little of philosophy or its history. I did my graduate studies in 1953–5 at the University of Oxford, which I enjoyed and to which I am grateful. I learned much from Quine (who was there as a visitor), Waismann, Price, Dummett, Hampshire, Grice, and a few others; but I had no teachers whose job it was to educate me in a systematic way, and at the end of those two years I was still very ignorant. Then after one year of teaching in the USA I began twelve years in Cambridge, where my main work was to lecture on early modern philosophy. To a large extent I taught myself philosophy through wrestling with those texts; and that is presumably the reason why I find it natural to view their

2

authors as contemporaries, almost as colleagues, and to view their main writings as sources of philosophical discoveries and insights. I do not mean that they wrote the truth and I read and believed it. On most issues, indeed, they were at least partly wrong; but their struggle with the problems has always been one of my chief sources of illumination.

Philosophy of mind and language

The philosophical area which has had the second-largest place in my work (two books and about twenty articles) is the philosophy of mind and language. My relationship to that part of philosophy is a little peculiar. I shall try to explain it.

Over the past thirty years, English-language philosophers have worked hard to get a better understanding of attributions of 'contentful' mental states. When we say that a person or animal believes that there is a cat in that tree, or wants it to be the case that there is water in the barrel, a good philosophical question arises: What is the role of 'There is a cat in that tree' or 'There is water in the barrel' in describing an animal's mind? What are we doing when we characterize minds with help from propositions about the outside world? One approach to this which has seemed right to many of us, and which has been called 'functionalism', says the following. Beliefs and desires are best understood in terms of their roles in explaining behavior. When we have the facts about how an animal moves in relation to its different environments, and how those environments are perceived by the animal, we have a chance of forming a theory that will help us to predict further facts about how it will behave in various circumstances; this theory will make use of the concepts of belief and desire; and according to functionalism that is the best way to explain what beliefs and desires are—namely, they are two kinds of mental state which relate in such-and-such ways to perceptions and to behavior. An essential feature of this program is it makes the concepts and belief and desire work in harness: you cannot explain a dog's digging a hole in the ground by saying 'It wants a bone' unless you also suppose that it thinks there is a bone in that place; and you cannot explain its digging by saying 'It thinks there is a bone there' unless you also suppose that it wants a bone.

Although it has turned out to be complicated, difficult, and sometimes controversial, functionalism has made a good start towards the truth about contentful mental states; or so I believe, in common with many other philosophers these days. My own work has not contributed to it any widely accepted doctrines, but I think it has contributed in a different way, which I shall explain.

In about 1963, at a time when functinonalism as now understood did not exist, I was writing my first book on Kant, and I found myself in a muddled frame of mind about what it is for a person or an animal to have beliefs. I had been working with the assumption that for someone to be intellectually capable of believing that there is a fire in the next valley (for example) is for him to be capable of saying that there is a fire in the next valley. Then I suddenly realized that I did not believe this! If it were right, then non-human animals would not be capable of beliefs, because they do not have languages in which beliefs could be expressed; but in non-philosophical contexts I was quite sure that some animals do have beliefs—e.g. that a dog can think that a cat has run up a tree. I tried and failed to develop a 'quick fix' for this trouble; so then I set the Kant work aside for many months, and worked on the question on what the basis is for attributing beliefs to non-human animals. The upshot was my book *Rationality*.

Although I came later to think that the central doctrines

of that little book were probably wrong, and I don't think anyone was convinced that they were right, the book had some success as a topic of discussion in graduate student seminars, and I think it had some influence in turning people's thoughts towards that kind of question and thus helping to launch the movement that led to functionalism.

About twelve years later I produced another, longer book which was partly a contribution to the philosophy of mind. It began as an exploration of an analysis of the concept of meaning, by the late Paul Grice. What is it for someone to utter a sound (or write a mark, or make a gesture) thereby meaning that such and such is the case? Grice rightly distinguished 'natural' meaning, which is involved when we say 'Those clouds mean that there will be rain' from the 'non-natural' kind of meaning that words have. (I do not know whether any Chinese language contains anything like this ambiguity in the verb 'to mean'.) Grice's concern was with non-natural meaning, which he sought to explain in terms of intention. Here is an idea that has appealed to many theorists of meaning:

> To say 'When he pointed towards the door he meant that his wife was to leave the room' is to say that in making his gesture he intended to get his wife to leave the room; and to say 'When he placed his hands on his cheeks he meant that he was sorry' is to say that in making that gesture he intended his audience to think that he was sorry.

Perhaps meaning is conected with intending to affect the behavior and beliefs of others, but the connection cannot be as simple as that. There are many ways in which someone might act, intending to get his wife to leave the room, that nobody would describe as his uttering something meaning that she was to leave the room. For example, if he starts making a loud noise which he knows she will want to get

away from, he may do this intending to get her to leave the room; but this behaviour of his will not be anything like uttering a command; it will not be an instance of non-natural meaning. So the above suggested account of meaning cannot be right as it stands.

That much was common knowledge when, in 1957, Grice offered a refinement of the above suggested analysis of the concept of non-natural meaning. He argued that non-natural meaning is intending to affect *in a certain way* the behaviour or beliefs of others; the 'way' in question fits the gesure of pointing to the door and does not fit the making of a disagreeable noise. I cannot go into the details of it here; I can only report that I am one of many philosophers who think that this ingenious work of Grice's constituted a fundamental advance in our understanding of the concept of meaning.

In the middle of the 1970s I wrote a book called *Linguistic Behaviour*, basing it on Grice's work. In it I tried to defend Grice's analysis against various criticisms which I thought wrong, and to improve it in certain ways so as to meet other criticisms which I thought to be right. The second half of that book was a success, I still believe, though I doubt if it has had much influence, partly because the whole topic became unfashionable at about the time the book was published. The first half of the book was a sort of contribution to the literature of functionalism. It came about in this way.

When I expounded Grice's ideas to students, I would say that his analysis of meaning gives us a better hold on the concept of meaning, which we can then use to get a better hold on the concept of language. Quite often I was met with the following objection:

> 'That whole procedure is circular. You are trying to clarify language with help from meaning, which you are trying to clarify with help from intention. But we could never be entitled to attribute intentions to any

creature unless we were already satisfied that it had the use of language. So your analysis is not the kind of building-up from the simpler to the more complex, or from the more to the less basic, which you believe it to be. It is just a circle—from language to intention to meaning to language.'

This is wrong. I believed then and still do believe that we can be entitled to attribute intentions (or desires) to a creature without attributing to it a capacity for anything like language. I decided to try to show this in the first half of *Linguistic Behaviour*, in which I revisited the scene of *Rationality* and made a much better job of it. The first half of *Linguistic Behaviour* is a contribution to the functionalist literature, though it has not been as influential as the work of some other writers with similar views, notably Daniel Dennett.

Ethics and action theory

I have done work in one part of ethics: a book and about a dozen articles. I started thinking about this at a small conference in Oxford in about 1963, where the philosopher Elizabeth Anscombe gave a speech in defence of the validity of absolutist moral principles, that is, ones of the form 'It would absolutely always be wrong to do such-and-such, no matter what the consequences would be of not doing it'. Some people, she said, have criticised such principles on the grounds that there is a difficulty about the drawing a line between an action and its consequences. She dismissed this as 'absurd', without explaining why anyone might have accepted it. (I admire philosophers who can be depended upon to present fairly the reasons for philosophical views that they reject; and I try to be like them, though no doubt I often fail. Miss Anscombe has never tried.) That started me on a train of thought that continued, intermittently, for more than thirty years.

There is a problem about dividing actions from consequences. When we ask 'What did he do?' or 'What action did he perform?', the answer might be 'He moved his foot like this' or 'He dislodged a rock on the bank of the river' or 'He started a slide of rocks into the river' or 'He diverted the river into a different channel' or 'He saved the village from being flooded'; and each of these might correctly describe the same behaviour. His 'basic' action, one might say, was the movement of his foot; all the rest were consequences of that; but we can report them without speaking overtly of 'consequences', instead giving suitably enriched accounts of 'what he did'. I know this to be true of the English language and of many other European languages; I confidently expect that it holds for all natural languages, because it reflects such a deep human need. We are seldom interested in the basic fact about how someone behaved—that is, in the fact about how he moved his limbs or his tongue—but we often care about what consequential state of affairs he brought about through such a movement.

Now, that does not rule out such absolutist principles as 'It would absolutely always be wrong to kill an innocent person'. But it does imply something about how such a principle needs to be defended. I shall try to explain this. Consider these two statements:

(a) Agent moved in a way that had Patient's death as a consequence.

(b) Agent killed Patient.

These are not strictly equivalent, because **(a)** could be true and **(b)** false. This could happen in several different ways, of which I shall mention the two that I think to be the most important. **(1)** If Agent instructs Servant to kill Patient, and the order is obeyed, **(a)** is true but **(b)** is false; the person who orders the killing may be as bad as the killer, but he

is not the killer. **(2)** If Agent walks out of the room without giving an injection which Patient needs to stimulate his heart, and if Patient dies as a result, **(a)** is true but **(b)** is false; in this case Agent *allows* Patient to die, but does not *make* him die, and so does not kill him (killing involves making-die). Now, in several articles and recently in a book called *The Act Itself*, I have looked carefully and in depth into the difference between making something happen and allowing it to happen; and into the difference between bringing something about through the actions of another person and bring it about by means that do not involve another person. I have argued that each of these differences, when properly understood, can be clearly seen to have no moral significance at all. That does not immediately refute the absolutist principle:

(i) It would absolutely always be wrong to kill an innocent person;

but it implies that if that principle is right then so are the principles:

(ii) It would absolutely always be wrong to allow an innocent person to die, and

(iii) It would absolutely always be wrong to act in such a way that in consequence someone else kills an innocent person.

No moral absolutist would accept either **(ii)** or **(iii)**, I believe; so this work of mine does serve as a sort of refutation of **(i)**, and thus of absolutist moral principles generally.

My main interest in it, however, concerns method rather than doctrine. I want to encourage moral philosophers who rest weight on various concepts that we use in characterising human behaviour to think harder, more deeply and more accurately about those concepts—what they mean and how they work. If that were more generally done, the literature in analytic ethics would progress faster and be more interesting.

Events

That work in ethics has connected with work I have done—in one book and five articles—on the concept of an event. We use this concept a lot: we speak of storms, earthquakes, weddings, dinners, executions, departures, arrivals, rises, falls, and so on; we have in fact many thousands of words that stand for kinds of event. In the past few decades—largely under the impetus of the work of Donald Davidson—there has grown up a considerable interest in understanding what sort of thing an event is, and how the concept of an event fits in with the rest of our system of thought about the world. Much of the work in this area has been organized around a big division of opinion which I shall now describe.

Jaegwon Kim has offered a plausible metaphysical view about what sort of item an event is, from which he has inferred an implausible semantic view about what is needed for two different expressions to refer to the very same event. Davidson rejected Kim's views about the semantics of event names, but offered no alternative to his metaphysic of events, that is, his account of what events are. In my book *Events and their Names* I argue that Kim is right in his metaphysics and Davidson is right in his semantics: the source of the trouble was Kim's mistaken inference of his semantics from his metaphysics. In my book I show that inference to be invalid, and explain the underlying error which prevented Kim and others from seeing it to be so.

That underlying error involves some details in the English language. I don't think they carry over to other European languages, and I have no reason to think that they carry over into any Chinese language. But the metaphysic of events which Kim has asserted and I have explained more deeply can presumably be presented in any language. I shall sketch it here very briefly.

6

First I must introduce a technical concept which was born in medieval European philosophy, and is still with us in various guises. The world contains **particular concrete** things (in the philosophical sense of 'concrete'), such as the house I live in; it also contains **abstract universal** properties, such as whiteness; and—some philosophers have thought—it also contains **abstract particulars** such as the whiteness of my house. This is like my house in being particular: it exists at a particular place, and through a particular stretch of time, unlike the universal whiteness which exists all over the place for a much longer time. But it is like the universal whiteness in that there is nothing to it but whiteness: it does not involve shape, texture, weight, or any of the other properties of my house. The following diagram may be a help:

	Universal	**Particular**
Abstract	whiteness (possessed by my house, the Taj Mahal, etc.)	the whiteness of my house
Concrete		my house (which is white, large, wooden, etc.)

The medievals gave various different names to items like the whiteness of my house, the circularity of this coin, and so on; one natural term for them is 'property instances'; but in recent years many of us have adopted one philosopher's suggestion that we call them 'tropes'.

I contend that an event is a trope. That fall of that sparrow was a particular instance of the property *falling*; it was distinct from the sparrow, and from the fact that the sparrow fell, and from the universal property falling, but it is metaphysically very close to each of these.

This account of what events are explains all the features that events clearly have: they are particulars, they are located in time (and usually in space), they are closely associated with *things* but are not identical with them, they exist because things have certain ·properties·, and so on. All of this falls into place as soon as we accept that an event is a trope. This metaphysic of events seems to have been Kim's, but I have set it in a clearer light by separating it from Kim's wrong semantics of event names.

My work on events has led me to three somewhat negative conclusions. The first concerns a range of philosophical issues that supposedly concern the nature of events as such, i.e. the metaphysics of events: Are there any instantaneous events, as distinct from ones that take up a period of time? Can two events occur in the same place at the same time? Could an event have occurred in a different place, or at a different time, from the one at which it actually occurred? What does it take for two events to be parts of a single larger event? These and other comparable questions have been debated in the philosophical literature, and I now think that most of this debate has been a waste of time. Although the questions are posed as though they concerned the nature of events (metaphysics), really they only concern the language in which we talk about events (semantics). Consider for example the questions about whether a given aggregate of tropes constitutes one event or several. Was the battle which raged on Tuesday the same battle that had begun on Monday? Was the fire that burned down my house on Tuesday the one that had burned down yours on Monday? We answer such questions by consulting our semantic conventions governing phrases like 'same fire' and 'same battle'. It cannot be the same fire, we hold, unless there was a continuous fire linking the two; but we allow that it can be the same battle even if the two episodes are

not linked by continuous fighting (the armies can sleep and then resume their battle, their one battle). This difference is purely conventional; we can imagine handling 'same battle' differently, insisting that if the armies slept then the next morning they started a new battle. This is a trivial verbasl matter, with no philosophical interest.

As well as writing about the concept of an event, philosophers have often used that concept in their theories. Certain problems in the philosophy of mind have been approached by asking how events in the brain relate to events in the mind; theories of causation have been expressed in terms of causal relations between events; and so on. These explorations have mostly been inconclusive and unsatisfying, and I contend that this is because of their use of the event concept. In my book I argue that our event concept is not sharp-edged or versatile enough to be useful in disciplined theories. That is my second negative conclusion.

The third is that the event concept—though quite useful for giving small, vague pieces of information—is one that we do not seriously need. We can say that *some people fought* rather than that *there was a fight*, that *rain fell and winds blew* rather than that *there was a storm*, that *he collapsed because he fell* rather than that *his fall caused his collapse*; and so on. Virtually all our uses of the event concept can easily be replaced by forms of speech in which that concept is not used at all.

Conditionals

My interest in events has connected, in rather indirect ways which I shall not explain here, with an interest in conditional statements. I have published five articles about the correct analysis of these, and I have a draft of a book which I shall probably call *Conditionals: An Opinionated Survey of the*

Literature. In the European languages, and no doubt in the Chinese languages too, we have two different kinds of conditional, of which I shall give examples. Suppose that you hire me to mend the roof of your house while you are away on a visit to another city; when you return, the roof is mended; so you have good reason to accept this:

> **(1)** If Bennett didn't mend the roof my house, someone else did.

But you may have no reason to accept this:

> **(2)** If Bennett had not mended the roof of my house, someone else would have.

You would accept **(2)** only if you believed that I had made some arrangement for someone else to do the job if I was unable to, and you may have no such belief. But the mere fact that the roof is mended gives you adequate grounds for accepting **(1)**. So although each of these conditional goes from the thought of Bennett not mending the roof to the thought of someone else mending it, they are profoundly different from one another—so different that one could be acceptable at a time when the other was entirely unacceptable.

This apparently trivial little matter is actually of great importance. Each of these two kinds of conditional does important work for us; I don't see how we could dispense with either of them (as we could easily dispense with the concept of an event). We need conditionals of type **(1)** for handling states of partial information, assessing evidence; and we need conditionals of type **(2)** in statements about how the world is structured—for example **(2)**-type conditionals are involved when we say that sugar is soluble, that wood is flammable, and so on. Furthermore, each of the two kinds of conditional raises many problems all on its own, and then there are further problems in seeing clearly how they relate to one another. Many fine philosophers have contributed

to our understanding of these matters; I have not been a pioneer in this area, but some of my contributions have made a difference, and I feel that I have contributed something to the communal progress towards the truth.

Getting a clear understanding of conditional statements is an exciting and challenging project. The last time I taught a graduate seminar on conditionals, each member of the class told me that this was one of the most instructive courses he had ever taken but also by far the most difficult.

Books by Jonathan Bennett:

Rationality (Routledge and Kegan Paul: London, 1964; re-issued Hackett: Indianapolis, 1989)
Kant's Analytic (Cambridge University Press, 1966)
Locke, Berkeley, Hume: Central Themes (Oxford University Press, 1971)
Kant's Dialectic (Cambridge University Press, 1974)
Linguistic Behaviour (Cambridge University Press, 1976; re-issued by Hackett: Indianapolis, 1990))
A Study of Spinoza's Ethics (Hackett: Indianapolis, 1984)
Events and their Names (Hackett: Indianapolis, 1988)
The Act Itself (Oxford University Press, 1995)
Learning from Six Philosophers: Descartes, Spinoza, Leibniz; Locke, Berkeley, Hume (Oxford University Press, 2001)
A Philosophical Guide to Conditionals (Oxford University Press, 2003)
(List updated in 2003.)