

Analytic Philosophy Without Naturalism

Edited by Antonella Corradini,
Sergio Galvan and E. Jonathan Lowe

Contents

<i>List of illustrations</i>	x
<i>Notes on contributors</i>	xi
<i>Preface</i>	xiv
<i>Acknowledgments</i>	xv
Introduction	1
ANTONELLA CORRADINI, SERGIO GALVAN AND E. JONATHAN LOWE	
PART I	
Epistemology	13
1 Reflection	15
FRANZ VON KUTSCHERA	
Reflection, self-consciousness and intentionality: comment on von Kutschera's paper	24
MICHELE LENOCI	
2 How naturalism implies skepticism	29
ALVIN PLANTINGA	
Methodological and/or ontological naturalism: comment on Plantinga's paper	45
ROBERTA CORVI	
PART II	
Ontology	51
3 Aristotelian substances and the theoretical/practical dichotomy	53
EDMUND RUNGGALDIER	
Aristotelian ontology versus naturalistic ontology: comment on Runggaldier's paper	67
ALESSANDRO GIORDANI	

4	What is naturalism? What is analytical philosophy? PETER VAN INWAGEN	74
	Naturalism, physicalism, and some notes on 'Analytical Philosophy': comment on van Inwagen's paper CHRISTIAN KANZIAN	89
PART III		
	Philosophy of religion	95
5	Contemporary cosmology and the existence of God WILLIAM LANE CRAIG	97
	How to deal with singularities: comment on Craig's paper WINFRIED LÖFFLER	134
6	The design argument: between science and metaphysics ROBIN COLLINS	140
	Metaphysical presuppositions of the argument from design: comment on Collins's paper SERGIO GALVAN	153
PART IV		
	Philosophy of mind	161
7	Rational selves and freedom of action E. JONATHAN LOWE	163
	I see that Martians persecute me: what should I do, if I want to act rationally?: Comment on Lowe's paper LUCIA URBANI ULIVI	178
8	Consciousness and freedom UWE MEIXNER	183
	Which consciousness do we need to have a choice?: Comment on Meixner's paper JOSEF QUITTERER	197
PART V		
	Practical philosophy	203
9	Naturalism, realism and objectivity in ethics ROBERT N. AUDI	205

Non-reductive naturalism versus non-naturalism in ethics: how wide is the gap? Comment on Audi's paper ANTONELLA CORRADINI	218
10 Resisting naturalism: the case of free will HUGH J. McCANN	225
Practical reasoning and action: comment on McCann's paper FRANCESCA CASTELLANI	241
<i>Bibliography</i>	246
<i>Index</i>	257

Contributors

Robert N. Audi is Professor of Philosophy and David. E. Gallo Professor of Business Ethics at the University of Notre Dame, Indiana. He works in ethics, theoretical and applied, and including political philosophy; epistemology, including religious epistemology; and philosophy of mind and action. His most recent books are *The Architecture of Reason* (Oxford, 2001) and *The Good in the Right: A Theory of Intuition and Intrinsic Value* (Princeton, NJ, 2004).

Francesca Castellani teaches Logic and Philosophy of the Social Sciences at the University of Trento, Italy. Her main scientific concerns are regarding action theory, intensional logics, logic and common sense. Her latest book is *Modelli di ragionamento pratico (Models of Practical Reasoning)* (Milan, 2000).

Robin Collins is Associate Professor of Philosophy at Messiah College, with research interests in science and religion, philosophy of physics, and philosophy of religion. One of his most recent book chapters is ‘The Evidence for Fine-Tuning’, in *God and Design: The Teleological Argument and Modern Science* (London, 2003).

Antonella Corradini is Professor of Philosophy of the Social Sciences at the Catholic University of Milan, Italy. Her research interests include the epistemology of psychology, the philosophy of mind and action, the philosophy of biology and the social sciences, and the foundations of ethics. Her latest book is *Modelli epistemologici in psicologia (Epistemological Models in Psychology)* (Rome, 2004, co-authored by M. Castiglioni).

Roberta Corvi is Professor of History of Contemporary Philosophy at the Catholic University of Milan, Italy. Her research interests concern contemporary Anglo-Saxon philosophy with particular attention to issues of knowledge, rationality and language. Among her publications is *An Introduction to the Thought of Karl Popper* (London, 1997).

William Lane Craig is Research Professor of Philosophy at the Talbot School of Theology, Biola University, La Mirada, California. His research

interests include philosophy of time and philosophy of religion. Among his publications are *Theism, Atheism, and Big Bang Cosmology* (Oxford, 1993) and *God, Time, and Eternity* (Dordrecht, 2001).

Sergio Galvan is Professor of Logic at the Catholic University of Milan, Italy. He works mainly on Gödel's theorems and on philosophical logic. Among his publications are *Introduzione ai teoremi di incompletezza (Introduction to the Incompleteness Theorems)* (Milan, 1992) and 'The principle of deontic reflexivity and the Kantian axiom' (*Logique et analyse*, 2001).

Alessandro Giordani is Research Professor of Logic at the Catholic University of Milan, Italy. His research interests concern in particular the application of modal logic to the analysis of issues in the field of ontology and philosophy of science. His publications include *Teoria della fondazione epistemica (A Theory of Epistemic Foundation)* (Milan, 2002).

Christian Kanzian is Professor of Philosophy at the University of Innsbruck, Austria. His main research fields are ontology, analytical philosophy, and the history of philosophy. His latest book is *Ereignisse und andere Partikularien (Events and Other Particulars)* (Paderborn, 2001).

Michele Lenoci is Professor of History of Contemporary Philosophy at the Catholic University of Milan, Italy. He has devoted his research to the theory of knowledge, ontology and anthropology, especially the range of phenomenological movement and analytical philosophy. Among his books are *Pensiero, linguaggio, verità (Thought, Language, Truth)* (Milan, 1986).

Winfried Löffler is Associate Professor of Philosophy at the University of Innsbruck, Austria. His main interests are the structure of arguments in the context of world-views and the philosophy of Bernard Bolzano. A recent publication is 'Epistemically relevant possible worlds' (*Grazer Philosophische Studien*, 2002).

E. Jonathan Lowe is Professor of Philosophy at the University of Durham, UK, specializing in metaphysics, the philosophy of mind and action, philosophical logic, and early modern philosophy. He is the author of *The Possibility of Metaphysics* (1998) and *Locke* (2005).

Hugh J. McCann is Professor of Philosophy at the Texas A&M University. His areas of interest concern action theory and the philosophy of religion. He is the author of *The Works of Agency* (Cornell, 1998).

Uwe Meixner currently teaches at the University of the Saarland, Germany. His research interests are metaphysics, logic, and the philosophy of mind.

He has published many articles and books, among them *Axiomatic Formal Ontology* (Dordrecht, 1997) and *The Two Sides of Being* (Paderborn, 2004).

Alvin Plantinga is John A. O'Brien Professor of Philosophy at the University of Notre Dame, Indiana. He gives seminars in philosophy of religion, metaphysics and epistemology. Among his many books is *Warrant and Proper Function* (Oxford, 1993).

Josef Quitterer is Associate Professor of Philosophy at the University of Innsbruck, Austria. He has published articles on the philosophy of mind and the philosophy of science. He is the co-author of *Unser Selbst: Personale Identität im Wandel neuronaler Prozesse (Our Self: Personal Identity Within the Changes of Neural Processes)* (Paderborn, 2003).

Edmund Runggaldier is Professor of Philosophy at the University of Innsbruck, Austria. His main research field is ontology in the Aristotelian tradition. Among his publications is *Che cosa sono le azioni? (What Are Actions?)* (Milan, 2000).

Lucia Urbani Ulivi is Professor of Theoretical Philosophy at the Catholic University of Milan, Italy. She works on metaphysics and ontology, the philosophy of mind, and anthropology. Her latest book is *Introduzione alla filosofia (Introduction to Philosophy)* (Lugano, 2005).

Peter van Inwagen is John Cardinal O'Hara Professor of Philosophy at the University of Notre Dame, Indiana. His research interests concern metaphysics and the philosophy of religion. His latest book is *Ontology, Identity, and Modality: Essays in Metaphysics* (Cambridge, 2001).

Franz von Kutschera is Professor Emeritus at the University of Regensburg, Germany. His main research interests are the philosophy of mind and ethics. He is the author of *Jenseits des Materialismus (Beyond Materialism)* (Paderborn, 2003).

Preface

In recent years, numerous attempts have been made by analytic philosophers to *naturalize* various different domains of philosophical inquiry. All these attempts have had the common goal of rendering these areas of philosophy amenable to empirical methods, with the intention of securing for them the supposedly objective status and broad intellectual approval currently associated with such methods. What exactly does it mean, though, to 'naturalize' a domain of philosophical inquiry?

What similarities and differences obtain between, for example, a naturalization programme in epistemology and one in ethics or action theory? And to what extent can these attempts at philosophical naturalization be judged to be consistent, coherent and supported by solid empirical evidence? Are there any just grounds to challenge the underlying assumptions of such programmes, by appealing to such notions as those of intentionality, rationality, consciousness, free will, and ethical value?

These topics were the subject matter of a major international conference which took place at the Catholic University of Milan in June 2003. The main speakers at the conference were analytic philosophers from both sides of the Atlantic who are both internationally recognized in their respective fields and noted for their record of critical reflection on the credentials of naturalism. The conference was divided into five sections, each devoted to a different area of philosophical inquiry: epistemology, ontology, the philosophy of mind, the philosophy of religion, and ethics. Revised versions of all the papers delivered by the main speakers, accompanied by the replies made to them by their commentators, are collected in this volume, along with an introductory chapter written by its editors.

Introduction

*Antonella Corradini, Sergio Galvan and
E. Jonathan Lowe*

Naturalism probably constitutes mainstream opinion in contemporary analytic philosophy. Its importance, however, lies not merely in its current popularity among analytic philosophers, but also in the fact that the naturalistic approach addresses itself systematically and compellingly to the entire spectrum of philosophical problems, thus presenting a serious challenge to the methods and claims of other philosophical traditions.

The areas of analytic philosophy in which the naturalistic approach has some role to play are many. However, there are four in which it is particularly significant: ontology, the theory of knowledge, the philosophy of mind, and practical philosophy. Moreover, a rich network of themes underlies these areas, forming a common background to them all. These are themes that generate arguments at the heart of each area and, at the same time, emerge from deep interconnections between problems belonging to different areas. Thus, for example, the philosophy of mind is a domain in which certain central theses of ontology, epistemology and practical philosophy are put to the test. As a consequence, if ontological – and, equally, epistemological – claims are guided by naturalistic principles, the philosophy of action will likewise acquire a naturalistic cast, and with it the ensuing conception of the philosophy of mind in general. This point is connected with the fact that the philosophy of mind is an appropriate domain in which to evaluate not only ontological hypotheses of a general kind, but also those with a specific metaphysical relevance for the philosophy of religion. Suppose, for example, that the philosophy of mind delivers conclusions which rule out an analysis of the relationship between the mental and the physical as being one of mere supervenience, thus excluding a physicalist explanation of the mental in these terms. The implication will be that the mental is *emergent* with respect to the physical. But if emergence is taken seriously – that is, if an emergent feature is regarded as one that is genuinely *novel* with respect to the intrinsic potentialities of the subvenient domain – this will mean that such a feature cannot be explained as arising out of the subvenient domain purely in and of itself. This will leave us with just two possibilities: either the novel feature is absolutely inexplicable – that is, it is a brute fact existing without a foundation of any kind, which

conflicts with the principle of sufficient reason – or else it is explicable only within a framework more extensive than is empirically available. Clearly, accepting the latter will mean opting for a transcendent metaphysics, which could even turn out to be a metaphysics of the Transcendent in the strict sense.

Obviously, the emergent status of the mental with respect to the physical is something that might need to be acknowledged even at the level of the philosophy of action and the ethics of moral responsibility. If we are convinced of our responsibility for (at least some of) our actions and choices, then we ought also to be convinced that we are their authors. This, however, means presupposing that we are free, which in turn presupposes a fundamental autonomy of the mental with respect to the physical. It must thus make sense to explain our actions on the basis of the (abstract and axiologically ideal) reasons that are at the root of our intentions and inform our will, not merely on the basis of the causes that generate our bodily movements. But if one recognizes the legitimacy of a rational explanation of our intentional actions, all the more must one accept the irreducibility of the mental to the physical, in accord with the idea, outlined earlier, of the emergence of the mental with respect to the physical.

In conclusion, the four areas of analytic philosophy mentioned at the outset – ontology, the theory of knowledge, the philosophy of mind and practical philosophy, together with the philosophy of religion – are closely interconnected, to such an extent that naturalism in its various forms is a position that deserves only to be accepted or rejected *en bloc*. One cannot, for example, consistently be a naturalist in ontology but a non-naturalist in the philosophy of action. Naturalism is an approach that either permeates every region of philosophical thought or else none.

Now, the thematic and methodological unity of philosophical thought that we have taken for granted so far is accompanied, in many contemporary expressions of analytic philosophy, by the presumption of naturalism. It is therefore quite natural to ask whether the presumption of naturalism, being so widespread, is simply *intrinsic* to the analytical method: that is, whether the structural unity underlying the various areas of philosophical thought is, perhaps, the reflection of a particular model – the naturalistic one – which informs the entire field. Or are we, on the contrary, dealing here with a mere *de facto* alliance, generated by the current intellectual climate, and hence one that is open to reversal once the thematic and methodological unity in question is put to the service of a different presumption? In the context of this issue, the contributions to the present volume may be viewed on two distinct, but related, levels.

On the one hand, the contributions show how it is possible for one to adopt the methodological principles of analytical investigation without accepting the presumption of naturalism. On the other, they demonstrate how the rejection of naturalism in one area implies a consequential change of paradigm in other areas too, entailing the elaboration – though superficially,

perhaps, without a uniform aim – of an alternative model centred on a vision of human beings and the world as constituting a reality that is not purely natural. From the first perspective, the contributions all stem from the question of whether analytic philosophy is intrinsically fated, in virtue of its method, to issue in naturalistic claims, or whether such claims – when and where they emerge – depend on factors that are tied not to the analytical method itself but rather to ulterior assumptions. The authors consider it important to respond to this question – see in particular Peter van Inwagen’s contribution – because it reflects a widespread conviction, especially to be found in philosophical circles rooted in the tradition of continental philosophy. This is the conviction that analytic philosophy is inherently apt to favour naturalistic positions, both for reasons of principle and for historical reasons connecting analytic philosophy tightly to the Vienna Circle and the movements that grew out of it. As Edmund Runggaldier’s contribution demonstrates, it is emphatically maintained in many quarters that the analytical method is, in particular, inherently reductionistic and scientistic in character. According to the continental critics of analytic philosophy, indeed, the analytical method aims to gather together the ultimate constituents of reality – be they atoms or tropes – and to construct, by means of a process of reduction to the basic entities, the entire complex image of reality as a whole. Furthermore, it supposedly assigns to the empirical sciences the task of determining what these basic entities are and how the modalities of the reduction should be specified, allowing philosophical reflection no cognitive role at all, in the strict sense. Seen in this light, the scientistic and naturalistic bias of analytic philosophy seems quite evident, so much so that it appears to warrant the critical verdict that continental philosophy so often pronounces upon it.

On the other hand, the theoretical virtues of the analytical method are sufficient to dissuade even the severest critic from abandoning it in order to embrace some of the perspectives to which the continental tradition appeals. Two of these virtues in particular are unanimously emphasized by the contributors. First, the analytical method is important on account of its systematicity – a characteristic deemed to be essential and inescapable for a philosophy set upon achieving a synthetic vision. Second, the analytical method is welcoming in its attitude towards science, in regard to both its methods and its results. It is not plausible, indeed, to engage in philosophical investigation as though science did not exist, avoiding a robust and systematic engagement with the fields of knowledge that the various scientific disciplines make available to us. The repudiation of naturalism, if it is a practicable option, should not be arrived at *a priori*, on the basis of a gratuitous assumption, but only *a posteriori*, as a result of a critical examination of scientific results and the contexts of their interpretation.

In answer to the anticipated question as to the plausibility of a non-naturalistic paradigm, the individual contributors advance, in their own independent ways, a model of analytical philosophy that is open to a non-naturalistic

vision of the world. And this is the second perspective from which the contributions of the different authors can be read. From this point of view, the contributions – in conjunction with the comments of the discussants – can be placed in three broad thematic groupings: the first concerning connections with the philosophy of religion, the second concerning connections with the theory of knowledge, and the third concerning connections at once with ontology, the philosophy of mind and practical philosophy.

In Part III on the philosophy of religion, William Lane Craig and Robin Collins develop two closely connected themes. The former presents a particular version of the Kalam cosmological argument. As is well known, this argument was originally advanced by al-Ghazali in a formulation based on two fundamental principles: (1) it is not possible to have an infinite regress in the temporal series of events. This principle, excluding as it does the possibility that the series of past events has no first member, implies that the universe has a beginning; and (2) that which has a beginning cannot be the source of its own existence, that is, cannot be *causa sui*. Together, the two principles imply that the universe requires an external Cause, whose attributes coincide with those of God the Creator. Craig has also developed such an argument in other contexts, where he has advanced various considerations in favour of the principle of the impossibility of an infinite regress in the temporal series of events, including the *a priori* truth that an infinite series of past events would constitute – impossibly – an example of the actual infinite. Craig's contribution in the present volume is characterized, however, by the fact that the reasons that he takes into consideration in order to exclude an infinite regress are all *a posteriori* reasons of a scientific nature. In fact, Craig surveys a large number of alternatives to the Big Bang cosmological model – according to which the universe has a beginning in time – pointing out the inherent difficulties of each of these. The outcome of this critical review is, thus, an empirical confirmation of the standard model and, hence, of the claim that the universe is not eternal, but had a beginning. At this point, one might have expected Craig to move directly to the conclusion of his argument, with a direct application of the principle that *that which has a beginning to its existence cannot be a cause of itself*. Craig, however, does not regard al-Ghazali's second principle as being unconditionally valid. And this is the second point of interest in Craig's contribution. He, in fact, regards the principle as being valid only in the context of a tensed conception of time. An opportunity to illustrate this point is provided by a critical discussion of the hypothesis advanced by J. Richard Gott and Li-Xin Li, according to which the laws of physics do not exclude the possibility that the universe, even though it had a beginning in time, is self-created. Now, in Craig's view, the self-creation hypothesis makes sense in the context of a tenseless (B-theoretical) conception of time, but not in the context of a tensed (A-theoretical) conception. B-theorists maintain that specifically temporal qualifications such as present, past, and future are subjective illusions of consciousness. Present, past, and future events in time

are all equally existent, differing from each other only in their positions, depending on whether they occur before, after or simultaneously with one another. In contrast, A-theorists contend that only the present exists, so that the process of becoming is a genuine passage from non-existence into being and from being into non-existence. Thus, if one adheres to the tensed conception of time, one cannot consider the universe to be capable of causing itself, since this would require the present world (which is not yet existent at the moment of causation) to be caused by the future world (which does not exist). Only by regarding temporal succession as a B-series is one free to suppose that the universe is self-caused. But, then, indeed, it doesn't even make sense to say that the universe has a beginning in time, because in the B-series nothing comes into being or ceases to exist. To say that the universe has a beginning is then just like saying that a street has a beginning, where the street's having a beginning doesn't signify that it comes into being but only that it starts its course in a certain place. Paradoxical though it may sound, according to B-theorists, the universe is eternal even though it has a beginning, and this is Craig's reason for siding explicitly with the A-theorists.

Collins's contribution on the argument from design is connected with Craig's, not only for the obvious reason that they both deal with important arguments of rational theology, but also for a deeper reason. Even though the arguments – Craig's Kalam argument concerning cosmological models and Collins's teleological argument concerning fine-tuning – draw on empirical considerations, they are both conducted at a metaphysical level or, more accurately, at a level at which metaphysics has a bearing on the empirical. If, for example, we ask why, in formulating Craig's cosmological argument, it is not sufficient to assume the finitude of cosmic evolution beginning with the singularity of the initial moment, we can reply, for reasons given above, that this depends on how one chooses between the A-theory and the B-theory of time, and that such a choice is a metaphysical one. The metaphysical character of such a choice can also be seen from the consequences of accepting one of these theories rather than the other. If one accepts the A-theory, a beginning signifies a coming into existence and, consequently, the existence of the universe is not explained unless one appeals to a cause external to the universe itself. Its existence would not be explained, because in the context of the A-theory the initial moment of the universe is the mark of its ontological contingency – if it exists, it does not do so in virtue of itself – whereas if the universe were ontologically necessary, it would not have need of a transcendent ground. In confirmation of this, if the B-theory were correct, each moment of the temporal series would exist eternally, including the very moment of the initial singularity, so that the series itself would be given eternally – that is, as something ontologically necessary – and there would be no need to seek an external explanation for it. As can easily be seen, the key to Craig's entire analysis of the cosmological argument lies in the point that the universe requires an ontological

foundation. If, then, the universe is ontologically contingent, its foundation is external to it and is transcendent; whereas if the universe is not contingent, then it is its own foundation, but this can be the case only because it is the universe itself that manages to confer upon itself its own ontological necessity.

The ontological point of view – and not only the cosmological – also plays an important role in the context of the argument from design, dealt with in Collins’s contribution. Collins examines the argument from design with the phenomenon of *fine-tuning* as his starting point. This phenomenon consists in the highly improbable fact that the values possessed by cosmological constants at the moment of the Big Bang were, to a very precise and finely determined degree, the only ones permitting the evolution of the cosmos in a way suited to the generation of intelligent life. However, the extreme improbability of fine-tuning is enormously diminished – fine-tuning ceases, indeed, to be improbable – given the hypothesis of design, whereas it remains high given the hypothesis of chance. From this and the *likelihood principle*, Collins draws the conclusion that the fact of fine-tuning supplies strong evidence in favour of the hypothesis of design as opposed to the atheistical hypothesis of chance. However, Collins’s actual conclusion is, in fact, weaker than the one that has literally just been stated. That conclusion is valid, rather, only if one can exclude the many-universes hypothesis – the hypothesis according to which our universe, with its entire evolutionary history, is just one of many (perhaps infinitely many) universes differing in various ways from our own and existing alongside it. If, indeed, the many-universes hypothesis were true, then the initial coincidences of our own universe would not be at all improbable. On the contrary, they would be ontologically necessary, precisely because the many-universes hypothesis would allow us to say that nothing is possible that is not realized in some universe, and this would apply also to the anthropic coincidences. But if fine-tuning were necessary, it would have maximal probability, and thus it would no longer make sense to apply the likelihood principle to it. As one can see also from the reply to Collins’s paper, the argument from design requires, like the Kalam argument, a point of departure that lies beyond the purely empirical framework supplied by Big Bang cosmology. One must argue against the plausibility of metaphysical hypotheses like the many-universes hypothesis and in that way provide warrant for the contingency of the origin of the universe itself. The need for a perspective that is broader than that provided purely by science follows additionally from the bearing that axiological considerations have on the very estimation of the epistemic probability of fine-tuning under the hypothesis of design. This fact comes out particularly clearly in remarks made in the reply to Collins’s paper.

The contributions in Part I dealing with the theory of knowledge present two strongly anti-naturalistic arguments. The first of these arguments is a reworking of the well-known evolutionary argument against naturalism first

expounded by Alvin Plantinga in Plantinga (1993b), which has given rise to a fruitful debate that is still ongoing. According to this argument, it is not possible to explain in terms of evolutionary processes why humans appear to possess a reliable cognitive apparatus, that is to say, an apparatus capable of generating beliefs that are on the whole true. Indeed, the mechanisms on which evolution is based are driven by adaptive constraints, and the revelation of truth does not figure among these. It is not knowledge of truth that is conducive to survival but adaptation of our behaviour to the environment, which only by chance involves knowledge of truth. In other words, adaptiveness is not associated with truth *as such*, but eventually leads to truth only in so far as the latter contributes to bringing about adaptive behaviour. Assuming the naturalistic perspective therefore renders it highly improbable that our cognitive faculties are reliable, from which a sceptical conclusion immediately follows. Plantinga infers, thus, that if one wishes to reject scepticism, one is bound to repudiate naturalism.

Plantinga's argument is the focus of an extremely intense and fertile debate, signs of which are also to be found in the reply to his paper. Some commentators on the argument, for instance, emphasize critically the fideistic presupposition on which it is based. Others regard as implausible the claim that a true belief can, in general, only by accident be useful for evolutionary purposes. Obviously, this could happen in some cases, but not universally and, above all, it is not sensible to suppose that it happens systematically. One could perhaps say that there are certain types of belief – where the type is defined by the content of the belief itself – whose existence is irrelevant to the ends of evolution and hence that such beliefs serve to justify Plantinga's thesis. It is essential, however, to establish that beliefs of this type do genuinely exist, which means that these beliefs will have to be characterized by quite specific and irreducible contents. For example, it is perfectly sensible to say that the truth of certain theoretical beliefs – such as many mathematical beliefs possessing highly abstract contents – is irrelevant as far as survival is concerned and therefore that the genesis of a corresponding cognitive faculty in human beings cannot easily be explained by recourse to the theory of evolution. But the problem at bottom is whether a set of true beliefs characterized in such a way do exist and are irreducible. For only in that case does the problem arise of explaining – in non-evolutionary terms – how a faculty could emerge that is capable of grasping the contents of such a set of beliefs. If the question is posed in these terms, however, Plantinga's anti-naturalistic argument comes in essence to coincide with some of the classical arguments against naturalism – for example, the argument that our cognitive activities exhibit emergent features, such as the capacities for abstraction and *a priori* judgement and others that are not easily accounted for by a naturalistic model.

In contrast to Plantinga's paper, Franz von Kutschera's starts precisely with the analysis of one of the characteristics of human mental activity, our capacity for reflection. This capacity is the source of the distinctively higher

cognitive activities of human beings and von Kutschera devotes the first part of his paper to illustrating some of these, including consciousness in particular as the focus of mental life, both theoretical and practical. The most original parts of von Kutschera's contribution, however, are the two that follow. In the first of these, he argues in favour of a constructivist conception of the notion of a set and, quite generally, a conceptualist conception of abstract entities. The upshot of this line of argument is that being a conceptualist means conceiving of the domain of abstract entities as an open one, capable of unlimited extension through the construction of ever higher levels of reflection. But the very openness of the domain of abstract entities, which takes concrete form in the openness of subjects' beliefs, is a characteristic which, as von Kutschera demonstrates in the final part of his paper, is incompatible with even the weakest notion of the supervenience of the mental on the physical. Without contending that the mental has a strong form of independence from the physical, von Kutschera finally arrives at an argument against the reduction of mental life in all its complexity to mere neurophysiological processes.

The papers in the remaining parts are closely connected with one another in virtue of a common theme which manifests itself in each of the contributions at the different levels on which they operate. This is the idea that anti-naturalistic theses are particularly relevant to human agency and all of the conceptual categories connected with it. It is an idea that is very evident in Edmund Runggaldier's paper, in Part II on ontology. In his contribution, Runggaldier shows how naturalistic ontology, in presupposing a scientific description of reality, gives rise to a dichotomy between theory and practice. Indeed, although it is able to cope with natural properties, by drawing on a scientific account of our objective knowledge of the world, it finds it difficult to adopt the same point of view in dealing with the phenomena of subjective human experience. The problem of the relationship between the theoretical dimension of human knowledge and the practical and subjective aspects of human life impinges directly on the complex of questions concerning agency that forms the heart of the debate about naturalism. What is human action? Is human action free? What does it mean to explain human action? What is the relationship between action and values? The papers in Part II on ontology, in Part IV on the philosophy of mind and in Part V on practical philosophy furnish answers to each of these questions that exhibit a profound unity and a common commitment to anti-naturalism.

The thesis held in common by all the authors in these three parts, and constituting the systematic assumption to which all of their other anti-naturalistic claims can be traced, is the proposition that it is impossible to give an adequate definition of human action without recognizing that such action is conscious and free. There is no human action properly so-called if there is no genuine possibility of choice – in the libertarian sense of the term – between different alternatives. This is, in particular, demanded by the fact that we feel ourselves to be morally responsible, which would not

be justifiable without an acknowledgement that it is in our power to deliberate in favour of this or that choice. But the conditions for free action are many and exacting.

First of all, it is necessary for action to be conscious. Uwe Meixner and Hugh McCann urge in a profound and detailed fashion that consciousness is presupposed by free action. For both authors, it is necessary for one to be conscious of the object of choice, because otherwise one could not have control over the action that occurs as a consequence of a choice. Furthermore, the consciousness involved in the act of deliberation is considered by both to be an irreducible given, not explicable by a reduction of the mental to the neurophysiological level. This is particularly clear in the model proposed by Meixner, in which consciousness plays a role in determining the actions of agents of every kind, including lower animals. A condition of this model is that action occurs in a non-deterministic context. Thus, according to Meixner, to say that an agent x is conscious of p is to say that p is meaningful for x , in the sense that information concerning p , carried by x 's state of consciousness of p , puts x in a position to be able to engender or impede, to his own advantage, the occurrence of p . Hence, for Meixner, to be conscious of a state of affairs is to be able to utilize information concerning that state of affairs in acting to one's greatest advantage, which rules out the possibility that the natural world is deterministic. Otherwise, an agent could not make use of the information of which it was conscious and would be constrained to follow the independent course of natural events, which would determine its actions inexorably. In that case, however, we would have to concede that agents – and in particular human agents – have decidedly limited epistemic capacities, since they would be systematically subject to the illusion of supposing that they could intervene in the course of natural events, without in reality having a power to do so.

As we can see, Meixner's paper opens up all the problems of action, various of which are approached by the other authors from several different angles. The diverse aspects that they deal with can all be related to two broad issues connected with the last two questions posed above. On the one hand, there are problems to do with the explanation of human action and, on the other, problems concerning practical knowledge. These two themes are, obviously, intimately interrelated.

The issue of the explanation of human action is principally examined by Jonathan Lowe and Hugh McCann. If action is understood to be conscious, the product of freely executed deliberation, then it is evident that action cannot be explained in accordance with models deployed in the natural sciences. More particularly, one must distance oneself from a causal model of explanation, in which causation is understood as a relation of production or succession among natural events. One must assume a teleological model, in which the explanation of action hinges on the intentions of the agent and the agent's beliefs about the world. This is insisted upon with particular force and clarity in Lowe's paper, which devotes several pages to deepening

the differences between causal explanation – which is valid in the domain of the natural sciences, including cognitive science itself – and explanation by means of reasons, which is valid in the field of intentional action. The conclusion that Lowe comes to is that human action is not suited to causal, but rather to rational, explanation. This allows us to formulate the problem of free will correctly, as one concerning a choice between alternatives that is not determined by causal mechanisms beyond the agent's control, but is left up to the agent, who ponders over various reasons – without being determined by any of them – before coming to a final decision. In this way one avoids the danger of having to say that our decisions occur merely by chance as a consequence of denying that they are caused. Denying that human action has a causal explanation is not equivalent to regarding our actions as the results of chance, understood simply as the absence of causes. Such a conclusion is not necessary, because the will makes its choices on the basis of reasons and all that is required for the purposes of freedom is that the subject should have control over the final decision. Of course, maintaining that the subject has control over the decision is not trivial, but it is the very least that is required – particularly in McCann's view – in order to secure a non-naturalistic vision of human beings. Clearly, the theory of rational action-explanation defended by Lowe and McCann also embodies a characteristically anti-naturalistic assumption of the sort that Meixner highlights. Just as it would make no sense to attribute to human beings a consciousness incapable of intervening in the course of natural events – for the simple reason that, as remarked above, this would be a consciousness perpetually doomed to illusion – so likewise it makes no sense to suppose that one can explain action rationally without postulating that the acting subject is capable of control. Similarly, this is for the simple reason that explaining in terms of reasons means supposing that reasons can be pondered over, evaluated and, in a word, controlled, right down to the final act of choice, which in its turn is an act of control *par excellence*.

With these last observations we come to the second problem area connected with the general theme of action: the place of values and preferences in the whole fabric of reasons underlying choice of action and the role played, from this point of view, by practical knowledge. The analysis of action advanced in the papers by Lowe, Meixner and McCann leads to an explanation of action in terms of reasons. From a logical point of view, this means that the explanation of human action proceeds according to a scheme conforming with principles of practical inference, rather than the Hempelian scheme or anything like it. But this in turn implies that the premises of any argument capable of explaining an action are – reducing the schema to its bare bones – volitional and epistemic in nature. There are no empirical laws in the *explanans*, and the nexus between *explanans* and *explanandum* is constituted by a logic specifically capable of connecting volitional and epistemic premises to a description of the action. Such a logic is the *logic of volition*, characterized precisely by the principle that unimpeded

volition analytically entails the realization of the willed action. On the other hand, the beliefs, desires and ideals of the agents are important factors bearing on their action: for they choose what action to perform in the light of their own beliefs, desires and ideals. How, then, is it possible that such factors should have a bearing on the choice of action and, as a consequence, be efficacious for the purpose of its performance, without its being the case that the action is actually caused by them? As was remarked above, this is possible because a genuine action is something that the agents choose to do in the light of their beliefs and desires and not because of them. If, therefore, the logic of volition is the logic of explanation in terms of reasons, we must show exactly how such reasons interact with the logic of volition.

Important pointers towards such an account are supplied, not only in the papers by Lowe and McCann, but also in Robert Audi's, as well as in the replies to McCann and to Audi himself. McCann urges in particular that choosing for a reason is a teleological relation, not a causal one. It follows that if an agent's desire to do something, together with relevant beliefs of the agent, are understood by the agent as reasons for doing that thing, then it is not the agent's mental states of desire or belief as such that have explanatory force, but rather their contents. Such a content is not an event or a state, but a 'thought', which is not an entity of an appropriate type to explain anything in causal or statistical terms. The mental contents of reasons for an action make reference to the end to which the action is directed, to the extent that this is anticipated by the agent as being positive or desirable, and to the most opportune means sufficient in the circumstances to realize that end. Finally, the explanatory value of reasons resides in the particular logical relationship that obtains between the relevant contents.

Here it is particularly important to distinguish between the content of the volitional premise and the contents of other relevant premises. This distinction gives a great practical significance to many of the points made in Audi's paper and the reply to it. In fact, the content of the volitional premise is traceable, in the final analysis, to the acting subject's set of preferences. Simplifying somewhat, these can be represented in terms of (somehow) ordered structures, whose elements are states of affairs that are objects of evaluation on the part of the agent. Obviously, the evaluation of such states of affairs – and the consequent ordering of them which can for this reason be termed a preferential one – is determined by the subjective preferences of the agent. But often in the process of forming these, an important role is played by the ordering of values in which the subject believes as values of an objective or ideal kind. Now, it is not easy to provide a purely naturalistic account of such values: for, being deemed objective, they cannot be treated as values that are subjectively traceable to individual desires and, being ideal and therefore abstract, neither can they be seen as natural states of affairs rooted in the biological structure of the human species. Such values need to be regarded as possessing an epistemological status that secures both their autonomy with respect to natural

qualities and their supervenience upon them, noting that a supervenience relation in this domain is not to be understood in a nomic sense. The complex of problems connected with a non-naturalistic conception of values, applicable not only within the theory of action but also in the broader context of practical philosophy and metaethics, is the target of overall concern in Audi's paper and the reply to it.

Returning to the question with which we began this introduction, we can, following our overall evaluation of the different contributions, confirm our initial thoughts concerning the relationship between naturalism and the analytical method. The question at issue involved two aspects. On the one hand, it had regard to the problem of whether the presumption of naturalism is inherent in the analytical method or whether, on the contrary, it is tied to the latter in a contingent way ungrounded by principle. On the other hand, it inquired whether there is a non-naturalistic approach that is fundamentally unified, global in scope and methodologically comparable with that of naturalism, but constituting an alternative to the latter. The answer is positive in both cases. The analytical method is not united to the naturalistic approach, so we can coherently claim to be both analytic philosophers and non-naturalists. And, indeed, the most significant thing to emerge from the papers in this volume is that all of them contribute, in a unified and convergent fashion, towards elaborating the fundamental theoretical outlines of an alternative picture to that of contemporary naturalism. The common feature of all the contributions, which gives precise sense to the anti-naturalistic picture that they disclose, seems to be their affirmation of the *openness of knowledge and reality*: the openness of reality to the transcendent, the openness of knowledge to forms of cognition free from the constraint of adaptation, the openness of the mind to the construction of mental structures that are ever more complex and abstract, the violation of the causal closure of the physical world, the emergence of the mental with respect to the physical, the openness of the will to free choice – in sum, the openness to ontological, epistemic and practical *novelty*. In contrast to anti-naturalistic conceptions of a traditional kind – which are often representative of philosophical paradigms incommensurable with that of contemporary naturalism – an anti-naturalistic picture of this sort is set within the framework of analytic philosophy and its characteristic method. The great advantage of this is that the differing models of contemporary naturalism and anti-naturalism are able to confront one another. Moreover, the unity of the non-naturalistic model, together with the global nature of the themes that serve to contrast it with the natural sciences, enables us to keep alive the dialogue with science while at the same time avoiding the insidious danger, rooted in the scientific method, of abjuring all knowledge of a synthetic kind.

4 What is naturalism? What is analytical philosophy?

Peter van Inwagen

The title of this conference is ‘Analytical Philosophy Without Naturalism?’, or, if I read the Italian title right, ‘Can Analytical Philosophy Be Other Than Naturalistic?’ Now I am an analytical philosopher, and I am not a naturalist – or, at any rate, I have not been a naturalist for at least twenty years. So I have been a practitioner of analytical philosophy without naturalism for twenty years. If I had ever been a naturalist, I certainly ceased to be one when I became a Christian. Before I was a Christian, I was neither an adherent of some other religion nor some sort of ‘philosophical’ theist. Nor did I hold any philosophical position – dualism, for example – that was inconsistent with naturalism. For all that, I’m not sure I ever was a fully paid-up card-carrying naturalist, but I can say that there was a time when I at least thought that naturalism was *probably* true.

My becoming a Christian had little or nothing to do with my philosophical work or my philosophical training – beyond the fact that I am the sort of person who is influenced by arguments, and all the arguments against Christianity that I knew of were, or so my philosophical training told me, either bad arguments (most of them) or inconclusive arguments (a few of them). But, of course, that all the arguments against some position are bad or inconclusive is hardly a reason for accepting it. (For one thing, at least in my view, to adopt the rule ‘Accept any position such that all the known arguments against that position are bad or inconclusive’ would lead to contradiction, since most if not all substantive philosophical positions have the following property: the arguments against them are at best inconclusive, and many of these positions are the denials of or are otherwise inconsistent with one another.)

When I became a Christian, I carried on with my philosophical work, just as physicians and cabinet-makers who have been converted to Christianity presumably carry on with theirs. In one sense, at least, my philosophical work did not seem deeply relevant to my religious convictions; no doubt most physicians and cabinet-makers would make a similar judgment. A physician who had been in the practice of providing the means of suicide to selected patients and who became a Christian would have to change his ways, of course, but he might well not have been doing that or anything

else inconsistent with Christian teaching. If I had been a defender of naturalism or of the moral permissibility of euthanasia, or of some other thesis inconsistent with Christian teaching, I too should have had to change my ways when I became a Christian. But, although I believed things inconsistent with Christian teaching before I became a Christian, I had never been interested in defending any of those things in my philosophical work. When, therefore, I ceased to be a naturalist I carried on being an analytical philosopher without noticing any break or discontinuity in my philosophical practice. And so I have done for twenty years. I have never seen any inconsistency in being both an analytical philosopher and a non-naturalist. (I should say that I don't see non-naturalism as a philosophical position as naturalism is a philosophical position. And I don't see opposition to naturalism as the constituting factor of any possible community of intellectual interest. If I learn that a notorious atheist is a Cartesian dualist, I don't say to myself, 'Well, at least he's a non-naturalist.' I say only: 'Well, that makes *two* things we disagree about.') I have never seen any connection between being an analytical philosopher and being a naturalist. Analytical philosophers, after all, disagree about a wide range of things; in my view, the truth or falsity of naturalism is one of them. There have been analytical philosophers who have believed that there were true contradictions, that unrealized possibilities were physical objects, that human persons were tiny material particles inside human brains, and that there were no such things as tables and chairs. If analytical philosophy is so tolerant a hostess as to admit such as these into her house, she will, surely, not gibe at granting entrance to a few non-naturalists.

It is of course true that most analytical philosophers are naturalists. Perhaps it is even true that most analytical philosophers are propagandists for naturalism, team players, proselytizing enthusiasts. (As I certainly was not when I was a naturalist – *if* I was a naturalist.) But the explanation of these facts, if the latter suggestion attains to facthood, must be psychological or sociological or something of that order. It is certainly not logical or philosophical, not even in part. Nelson Goodman once said that the opposition between empiricism and realism about universals (an opposition that in some sense certainly exists) was like the opposition between being a truck driver and being a lover of the ballet. That is, the opposition is real and undeniable, but to be explained otherwise than by appeal to the content of empiricism and realism. I would say the same about the connection between analytical philosophy and naturalism.

If I see matters this way, then, what in my view is there for a conference called 'Analytical Philosophy Without Naturalism?' to be *about*? – and what place is there in it for a philosopher with my views? I shall consider a few possibilities.

Someone might take the title of the conference to imply that a plausible *prima facie* case could be made for the thesis that there was an intimate connection (a logical or philosophical connection) between analytical philosophy and

naturalism. And that person might go on to infer that (because of the presence of the word 'without' in the conference title, and because the conference took place under the auspices of a Catholic university) the purpose of the conference was to provide a forum for philosophers who had something to say against this *prima facie* case. But if this is the purpose of the conference, I cannot contribute to it for the simple reason that I see no *prima facie* case for a philosophical connection between analytical philosophy and naturalism.

Here is a second possibility. Someone might suppose that the title of the conference was chosen in recognition of two facts: that most analytical philosophers are naturalists, and that this fact had no ground in the nature of analytical philosophy or in the content of naturalism. This person might suppose that the purpose of the conference was to underscore the second of these facts; that its purpose was to provide a forum for analytical philosophers who were not naturalists, to provide an occasion on which a series of philosophical papers were presented that were free from naturalistic presuppositions. Might I have something to say at a conference that fitted this description? I might indeed, but it is so easy to have something to say at such a conference that one can be reasonably certain that our second 'someone' has not got the intentions of the organizers of the conference right. Imagine that I had proposed to present a paper to you on one of the following topics:

- An examination of the account of variables given in Quine's 'Variables Explained Away'.
- An argument for the conclusion that agent-causation does not solve what writers on free will call 'the control problem'.
- A discussion of David Lewis's argument for the conclusion that any theory of composition besides universalism must have the consequence that the linguistic theory of vagueness is false.

(I have indeed written on each of these topics.) I can assure you that if I had presented a paper on any of these topics, it would have had no naturalistic presuppositions. And nor would it have had any non-naturalistic presuppositions. In fact, every proposition I should have defended in a paper on any of these topics would have been consistent both with naturalism and with the denial of naturalism. And that, I think, is exactly what would have rendered papers on these topics unsuitable for presentation at this conference. Whatever the purpose of the conference may be, it cannot be simply to afford philosophers an opportunity to present papers on topics that have nothing whatever to do with naturalism.

Here is a third possibility. It represents *my* guess as to the intention of the organizers of the conference: they propose to provide a forum for analytical philosophers who are anti-naturalists, to provide an occasion for the presentation of philosophical papers that presuppose or defend the denial of

naturalism. (If this is right, a better title for the conference would have been 'Analytical Philosophy with Anti-naturalism'.) Have I anything to contribute to a conference of that nature? Well, no, not really. As I have said, most of my philosophical work is simply irrelevant to naturalism. Most of my philosophical convictions, moreover, are regrettably close to those of the typical naturalist. True, I think that there are non-human immaterial intelligences, such as God and St Michael. I believe that Christ was raised from the dead, and that one day we, like him, shall be raised imperishable. I believe that every Sunday I eat and drink – in a mystery, but an entirely non-metaphorical mystery – the body and blood of Christ. These beliefs of mine are obviously inconsistent with naturalism; if I belonged to the Federated Society of People Who Are Naturalists Because They Are So Smart, and if I publicly confessed these beliefs, I should certainly be expelled. But these beliefs don't play any role in my philosophical work. It's true that I have written essays on topics like the Holy Trinity and biblical inspiration and miracles and the problem of evil. But I don't think that these essays contain much that couldn't be accepted by a naturalist. (At most, the occasional sentence.) The purpose of my essays on religious topics has always been apologetical. My purpose in each of them was to examine an attempt to demonstrate the falsity of some Christian doctrine: arguments for the incoherency of the doctrine of the Holy Trinity, for example, or various forms of the argument from evil. My conclusion in each case was that the argument under consideration was a failure. And that a certain argument for the falsity of some Christian doctrine is a failure is a conclusion that a naturalist can consistently, if not happily, accept.

If one turns from the title of the conference to the 'Presentation' that accompanies the conference materials, one finds that the subject-matter of the conference is there said to be those programs that propose to 'naturalize' some field of enquiry or some part of philosophy – to produce a naturalized epistemology, for example – or to provide naturalistic accounts of certain problematical concepts or entities or phenomena (of truth, of concepts, of intentionality). (The present paper is officially part of the 'Ontology' section of this conference, although I confess it contains little in the way of ontology, a fact I am now in the process of excusing.) There is, it seems to me, a certain tension between the title of the conference and the 'Presentation', for naturalization projects do not make up a very large part of analytical philosophy. If one opens a recent volume of any uncontroversially analytical journal to a randomly chosen page, one is not *very* likely to find on that page a part of some attempt to naturalize something. So, it would seem, a great deal of even the most recent analytical philosophy has no connection with the various naturalization projects that some analytical philosophers are engaged in. In any case, I have nothing to say about these projects. I know little of them, and what I do know of them suggests that they are vague, pretentious, amorphous, programmatic, and have not achieved anything of value. I should say, however, that, except on that ground, I am not

particularly opposed to them. (My attitude toward them is like Francis Bacon's attitude toward magic. 'Magicians', Bacon said, 'attain little to greatness and certainty of works.') In principle, I could accept the results of many of them, if they had results, although I'd probably balk at any attempt to naturalize religious belief. Suppose, for example, that someone proposed to naturalize morality. I am not thinking of someone like Michael Ruse, who has attempted to show that morality is an illusion that can be explained in naturalistic terms. I am rather thinking of someone whose project is to defend just that meta-ethical thesis that was called naturalism long before the current fashion for naturalizing things got under way: the thesis that moral properties are to be identified with certain natural properties. Grave difficulties attend this project, but, supposing them to have been overcome, I can see no objection to accepting its results.

So, this suggestion, too, the suggestion that the papers presented at this conference concern themselves with projects that propose to 'naturalize' various things offers me no prospect for contributing a paper to the conference.

What, then, am I to do? I can think of nothing to do but what my title suggests I am going to do: to try to say what naturalism is and what analytical philosophy is. If I cannot contribute directly, as it were, to what I take the purpose of this conference to be, at least I can try to say something that is relevant to that purpose.

I begin with naturalism. The word 'naturalism' obviously has something to do with 'nature'. There can, I think, be no objection to saying that a naturalist is someone who says that there is nothing besides nature, nothing in addition to nature, nothing outside or beyond nature – or no objection besides this one: it's not a very informative thing to say. For what is this 'nature' besides which, if the naturalists are right, there is nothing? The etymologies of the words '*natura*' and '*physis*' are connected with the ideas of birth and growth. And this root meaning is preserved in the usual understanding of the English word 'nature' by most people: as most people use the word 'nature', nature is something that is mostly green. (Indeed, in its primary sense, the English word 'naturalist' applies not to a person who holds a certain philosophical position but to a person who knows all about birds and flowers.) But if naturalism is the thesis that there is nothing besides nature, the word 'nature' must be understood as denoting something that is not mostly green but mostly black. (I owe the 'nature as green'/'nature as black' distinction to C.S. Lewis.) 'Nature', in the sense that interests us, is simply another name for the physical universe or the cosmos. Naturalism is, therefore, the thesis that there is nothing besides the physical universe. The physical universe is the sum total of all physical things. Naturalism is, therefore, the thesis that all things are physical things. Naturalism is, therefore, physicalism. (We should certainly be very puzzled if someone said, 'I'm not a naturalist, but I'm a physicalist' – or 'I'm not a physicalist, but I'm a naturalist.') But this is of very little help, for it

merely replaces the question, What does 'natural' mean? with the question, What does 'physical' mean? – and the latter word has come a very long way indeed from its archaic Greek genesis in the growth of plants. And what *does* 'physical' mean? Well, the word (in the sense that is of present interest) obviously has something to do with the science of physics. If we were satisfied with the current state of physics, of standard working physics, as opposed to speculative physics, we could define 'physical' this way. There are two basic kinds of physical entities: space-time and elementary particles. A physical thing is either space-time or some part of space-time or an elementary particle or something composed of elementary particles. But no one is in fact satisfied with the current state of physics, a composite of the general theory of relativity and the so-called standard theory of elementary particles. One very good reason for this is that the general theory of relativity and the standard theory of elementary particles are inconsistent with each other, although in most situations that physicists are concerned to study, the inconsistency can be walled off and ignored. A second reason is that, while the general theory of relativity is beautifully plausible, the standard theory, for all its effectiveness, contains much unlovely arbitrariness. It is hoped that one day physics will be a unified and beautiful whole, that the science of physics will be coextensive with a single theory that explains both gravity (the present business of general relativity) and everything else (the present business of the standard theory). But it is generally conceded that the ontology of this dreamt-of final theory will almost certainly be very unlike the ontologies of general relativity and the standard theory. Perhaps there will in the end be no space-time, but only particles (or something analogous to particles; maybe they won't be point-like); or perhaps there will be no particles but only space-time (particles being reduced to singularities in space-time); or perhaps there will be neither, but some third kind of thing, a kind of which we do not at present have a clear picture.

However this may be, most people would agree that we cannot plausibly define 'physical' by reference to the ontologies of the two components of present-day physics. What, then, are we to do if we wish to define 'physical'? Most thinkers who have tried to answer this question would say something like this: we must abstract from the physics we know its essential core, something that, because it is essential to physics, will survive any future transformation of physics. We can do this; and when we have done it, we shall not have something so abstract and general as to be philosophically uninteresting. We shall have a really interesting concept, and it will reveal to us the essence of or allow us plausibly to define the physical (and hence the natural and hence naturalism).

Attempts at 'abstracting the essential core' of present-day physics have been of two types: epistemological and ontological. The former finds the essence of physics in an unchanging method. Although (the proponents of the epistemological attempt say) the physics of the future will no doubt

employ concepts very different from those of the present day, physicists will always use the same methods in their investigations, and these methods will one day lead them to a final theory and this final theory will be a *physical* theory (and not something that grew out of physics but is not physics, as experimental psychology grew out of philosophy but is not philosophy) because it is historically continuous with the discipline now called 'physics', its continuity with present-day physics consisting precisely in the continued application of those unchanging methods of investigation. This final theory will postulate entities of certain sorts; it will assert their existence. We may therefore plausibly define physical things as things of the sorts that will be postulated by the final theory. Or, at any rate, we may define physical things as things that are composed of things of those sorts, for no doubt there are special contingent structures – human beings and stars, for example – whose existence will not be a consequence of the theory itself, although their existence will be a consequence of the theory plus certain boundary conditions. Naturalism, or physicalism, may be defined as the thesis that there exist only things of the sorts that will be postulated by the final theory or else things composed of those things.

I don't know what to say about this suggestion, other than to voice some suspicions. I'm not a philosopher of science, after all, and I know little of these matters. My unschooled suspicions are these: I suspect that there may be no such thing as the unchanging methods of investigation in physics that the definition depends on; and I suspect that if there are such methods, human beings, even given world enough and time, may not be capable using them to reach a final theory. (Who knows?: maybe there will be competing candidates for the office 'final theory' such that the unchanging methods of investigation will tell physicists that one could choose between these candidates only by performing an experiment that would require a linear accelerator a thousand light-years long, a device human beings will never be in a position to construct.) I do not know what others would say about this, but, for my part, I should not want the validity of the account of naturalism I endorsed to depend on the assumption that the suspicions I have voiced are ungrounded.

The ontological approach seems to me to be the more promising. It is based on the following idea: to see what is essential to physics, look at the kinds of theoretical entities postulated by physics and ask what is essential to entities of those kinds. (I use 'theoretical entity' as an epistemological, not an ontological term: theoretical entities are entities such that our reasons for believing in them are theoretical.) When we do this, we find that the following things can be said about the theoretical entities of physics with such assurance that it seems reasonable to suppose that any possible future physics will postulate only entities that have these properties. First (whatever the 'new pan-psychists', people like David Chalmers, may try to tell us), nothing in the nature of these entities has anything to do with the mental: nothing in their nature pertains to personality, consciousness, or

intentionality. Human persons and all other entities that are conscious or have intentional states are complex, special, physically contingent structures composed of the fundamental entities postulated by physics. (It seems plausible to suppose that these structures will turn out to be very large relative to the fundamental entities. But this is uncertain: perhaps the fundamental entities are not the particles or tiny loops we are inclined to expect but all-pervasive fields or some such.) The mental properties of these complex structures must supervene on the properties of the entities physics postulates, of course, but those entities themselves do not have the least tincture of the mental about them. Just as the theoretical entities of physics are in no way alive (although all living things are ultimately composed of them), they in no way think or feel. The theoretical entities of physics, moreover, have no teleological properties. Even if certain complex structures entirely composed of these entities have ends or purposes or functions (even if the function of hearts is to pump blood, for example), teleology is in no way present in the theoretical entities that are the ultimate parts of these structures. Finally, these properties are numerically quantifiable. If, for example, the theoretical entities have mass or charge (properties that may not belong to the theoretical entities of future physics), one entity may have twice the charge or one-half the mass of another. And what applies to the properties of these entities applies, *mutatis mutandis*, to the relations that they bear to one another. If, for example, one entity is separated from another by a certain space-time interval (a relation that may not figure in the theories of a future physics) – well, there is nothing mental or teleological about space-time intervals, and they have numerical measures.

Physicalism or naturalism may now be defined: nothing exists but (a) entities having such properties (non-mental, non-teleological, numerically quantifiable properties) and bearing such relations (non-mental, non-teleological, numerically quantifiable relations) to one another; and (b) composite objects that have these entities as their ultimate parts. Some among the mereological sums *may* have mental or teleological (and non-quantifiable) properties; whether they have such properties or not, their properties, *all* their properties, supervene on the properties of the entities that are their ultimate parts and the relations these ultimate parts bear to one another. The distribution of properties in any world depends on the distribution of properties among the ‘ultimate’ entities. As a matter of metaphysical necessity, two worlds that are alike in respect of the properties of and relations among their ultimate entities are alike in *every* respect. According to naturalism, the truth of what is expressed by any of the following sentences, supposing them to express truths, supervenes on the (non-mental, non-teleological, numerically quantifiable) properties of and the (non-mental, non-teleological, numerically quantifiable) relations that hold among the ultimate entities:

- Tokyo is the capital of Japan.
- Alice’s favorite color is blue.

- It is wrong to bear false witness against one's neighbor.
- There are at least half a million species of beetles.
- It's easier for a German to learn Dutch than it is for an Italian.
- Most mathematicians have a low opinion of mathematical logic.

There is one issue that divides naturalists that I will simply note. Some naturalists are unhappy about affirming the real existence of platonic or abstract entities (such as properties or propositions or mathematical objects like numbers or functions or tensors) and some are not. Rather than take sides in this dispute, I will simply insert one additional word into my statement of naturalism: 'Nothing *concrete* exists but . . . ' This is naturalism *simpliciter*. We may then say that there are two schools of naturalism: anti-platonic naturalism, which insists that there are no platonic entities, and the more liberal school, which simply does not worry about platonic entities (which, after all, have no causal powers) and happily allows quantification over them. One might, of course, worry about what the opposed pair of terms 'concrete' and 'abstract' mean, but this is a problem that confronts everyone with ontological interests, and there is no reason to suppose that a naturalist and a non-naturalist (or an anti-platonic and a 'liberal' naturalist) would disagree about the merits of any proposed way of making the abstract/concrete distinction clear.

This, then, is naturalism. I turn now to the topic of analytical philosophy. It used to be, until quite recently I think, that the general public, insofar as they were even vaguely aware of them, referred to analytical philosophers as logical positivists, whether they were logical positivists or not, rather as some people refer to citizens of the United Kingdom of Great Britain and Northern Ireland as 'Englishmen'. And the label 'logical positivist' fits most analytical philosophers about as well as 'Englishman' fits a schoolgirl in the Scottish Highlands. Calling all analytical philosophers logical positivists is, in fact, even more misleading than calling all Britons Englishmen, for while lots of Britons are Englishmen, few if any of the analytical philosophers of the present day are logical positivists. The application of 'logical positivist' to analytical philosophers generally is important for our present concerns, because a commitment to naturalism (in the form of the following thesis, which the logical positivists shared with the nineteenth-century positivists: scientific knowledge is the only knowledge) really is an essential component of logical positivism. The logical positivists were, properly speaking, a school of philosophers that originated in Vienna in the 1920s. The members of this school were murdered or driven from the European continent by the Nazis (many of them were Jews; they were all liberals or socialists). Logical positivism in exile, so to call it, flourished briefly in the English-speaking countries till the late 1940s, and was then abandoned by its former adherents for purely philosophical reasons. (We should honor the logical positivists for this. I believe that logical positivism is the only important philosophical position to have been abandoned by its own inventors because rational argument convinced them that it was false.)

Although the widespread identification of analytical philosophy with logical positivism was an error, it was not an inexplicable error. Logical positivists *were* what came to be called analytical philosophers (just as Englishmen *are* Britons). No doubt a page of philosophy written by a Viennese logical positivist in the early 1930s and a page of philosophy written by any analytical philosopher in the early twenty-first century would seem very much the same sort of thing to those philosophers whose primary twentieth-century texts are Husserl and Heidegger. Analytical philosophy is not in any useful sense of the word a school of philosophy (like Thomism); it is rather a philosophical community, a community of philosophical discourse (like scholasticism). It is because the logical positivists (on the one hand) and present-day analytical philosophers (on the other) belong to the same community of philosophical discourse that their respective writings can seem very much the same sort of thing to members of other communities of philosophical discourse.

What, then, it is that defines this philosophical community to which both present-day analytical philosophers and the logical positivists belong? One obvious answer to this question is that, as the phrase 'analytical philosophy' suggests, this community is defined by the fact that it assigns a central place in philosophy to something called 'analysis'. Let us examine this suggestion.

The post-war popularity of the term 'analytical philosophy' was largely a product of the popularity of a certain answer to a question that troubled anglophone philosophers in the 1950s: What is it that we philosophers *do*, anyway? This was not, of course, the first time in the history of philosophy that this question had been posed. But, like all perennial philosophical questions, it had been answered in different ways in different eras in the history of philosophy. The following answer has had its advocates in every era: philosophy provides a special kind of knowledge. But this answer immediately raises a second question: How is this knowledge related to the knowledge provided by the other sciences (or disciplines, as we should say today)? In the Middle Ages it was said that philosophy was the handmaiden of theology, the queen of the sciences. When physical science had become the queen of the sciences, many said that philosophy was the handmaiden of science (in the modern sense of science): if biology and chemistry and astronomy are the branches of the tree of human knowledge (Descartes said) and physics is its trunk, philosophy comprises its roots. Others said that philosophy was not an indoor but a rather less prestigious outdoor servant of the sciences, the under-laborer of the sciences, whose job was to clear away intellectual litter that might cause science to stumble or otherwise impede its progress. A further view, perhaps not held to be inconsistent with the under-laborer view, was that it was the business of philosophy to understand the sources of human knowledge. And there was always a school that assigned a much grander role to philosophy: Spinoza and Leibniz believed, as Plato had, that philosophy could produce, that it was the

business of philosophy to produce, a kind of knowledge that would exhibit the whole of reality as a rational system. Or was the business of philosophy (as still others said) to show what we, we thinkers and knowers, must be like if we have the kinds of *a priori* and *a posteriori* knowledge that we know we have? Well, you know the story of these conceptions of the purpose or function of philosophy as well as I (in many cases, no doubt, much better). The point is: they eventually came to be seen as wholly unsatisfactory by many who reflected on the nature of philosophy. Just as philosophy came to be seen by many as unable to provide indisputable knowledge of God, freedom, and immortality (not to mention morality, knowledge, and all the other matters of concern philosophy had once claimed to provide indisputable knowledge of), philosophy came to be seen by many as unable to provide indisputable knowledge of (or indeed any halfway satisfactory account of) itself, its own nature and function.

In the 1950s, as I have said, some anglophone philosophers became interested in finding some little thing for philosophy to 'do' that would not run afoul of the claims and accomplishments of the physical and biological and behavioral sciences. (And their interest was not in finding something for philosophy to do such that it would be clear that this was what philosophy had always done. Their interest was in finding something that philosophers could do *from then on*; something for philosophers – that is, members of university faculties of philosophy – to retreat to and occupy themselves with, as one might say.) And this was their solution: The business of philosophers, the *proper* business of philosophers, the only thing philosophers should ever have been concerned with (admittedly, philosophers have had other concerns, much grander and more ambitious concerns, concerns whose disastrous consequences are amply illustrated in the history of philosophy), what philosophers do whenever they produce anything of lasting value, is the analysis of concepts. This was not the first time this thesis had been put forward; it had been a central thesis of logical positivism. The post-positivist philosophers (who were for the most part the same people who had been the positivist philosophers) retained the thesis but did not retain the theory of meaning that had led the logical positivists to maintain that the proper vocation of philosophy was the analysis of concepts. (It had led them by the scruff of the neck: the verification theory of meaning, the central dogma of logical positivism, left no possible vocation for philosophy other than the analysis of concepts, unless it were as a kind of non-cognitive poetry.) The post-positivist account of philosophy had thus no theoretical basis (according to itself, it could have had no basis but an analysis of the concept 'proper vocation of philosophy', since the question, What is the proper vocation of philosophy? is pretty evidently a philosophical question.) And it soon became evident that philosophers, even English-speaking philosophers who were paradigmatically analytical philosophers, were not going to abide by it. There were theoretical attacks on the idea; certainly the attack on the idea of analyticity in Quine's enormously influential

essay 'Two Dogmas of Empiricism' was, among other things, an attack on this conception of philosophy, for, if there is an unproblematical notion of the analysis of concepts, there is an unproblematical notion of analyticity. Much more important to my mind, however, is the fact that analytical philosophers, paradigmatically analytical philosophers, simply refused to restrict their activities to the analysis of concepts. One remembers A.J. Ayer's sad reaction to Alvin Plantinga's Oxford lectures on modality: 'I've lived in vain.' And Plantinga was hardly an isolated case: it could hardly be said that the writings of Chisholm and Kripke and David Lewis contained nothing but the analysis of concepts. In a way, this was a pity: that the proper business of philosophy is the analysis of concepts is a nice theory and a fairly clear theory as philosophical theories go. And if that is what philosophy is, it seems plausible to say that *analytical* philosophy is philosophy that, as a matter of self-conscious methodology, restricts itself to philosophy's proper business, the analysis of concepts. The only trouble with this is that it isn't true. The proper business of philosophy is not coextensive with the analysis of concepts, and many central figures of the analytical movement have demonstrated by their choices of topics and methods that they do not understand philosophy in that way.

If analytical philosophy cannot be defined by reference to its supposed method, the analysis of concepts, might it be defined by reference to its historical roots? Well, perhaps, but those roots are various and tangled. We certainly can't say that analytical philosophy is philosophy that is wholly rooted in the native philosophy of the British Isles. (I say 'the native philosophy' to take account of the fact that for a significant part of the nineteenth and the early twentieth century, British philosophy looked to Germany for its fundamental ideas.) It has for some time now been uncontroversial that analytical philosophy cannot be defined as that philosophy that has its roots in British empiricism and the work of Moore and Russell. Michael Dummett has said that analytical philosophy, which is sometimes called Anglo-American philosophy, would be better called Anglo-Austrian philosophy. There is something to this. The work of Bolzano and Brentano and Meinong is of immense importance to an understanding of the history of analytical philosophy. (The Viennese logical positivists, those of them who *were* Austrians, were not so much Austrian philosophers as philosophers who were Austrians. And, although it is unfashionable to say this, I will say it: I think that this is also true of Wittgenstein.) And of course, we must not forget the contributions of the great Polish philosophers and logicians between the wars to analytical philosophy and the important post-war contributions of Danish and Swedish and Finnish philosophers. And then there is Frege. If Germany, speaking generally, has had little influence on analytical philosophy until very recently, no one is more important to the history of analytical philosophy than Frege – not Hume, not Moore, not Russell, not even Wittgenstein. The history of analytical philosophy is so complex and involves so much historical contingency – from the point of

view of the history of *philosophy*, the Nazi control of the universities in Germany and other countries of central Europe, the Second World War, and the Soviet control of the nations of central Europe, are matters of historical contingency – that it is very hard to base any account of analytical philosophy on its history.

In the end, I think, we can say only that, although we can make some remarks about analytical philosophy, we can give no useful account of its essence – if it has an essence. We can say nothing that is as helpful to someone who wants to know what analytical philosophy is as statements like ‘Analytical philosophy is philosophy that, as a matter of self-conscious methodology, restricts itself to the analysis of concepts’ or ‘Analytical philosophy is philosophy that has its roots in British empiricism and the work of Moore and Russell’ would be if only they were true. But we can, as I say, make some remarks:

- Analytical philosophy, as a general rule, aspires to clarity of expression. This does not of course mean that the writings of analytical philosophers are always clear. Wilfrid Sellars, for example, was an analytical philosopher if anyone ever was, and he was notoriously obscure. But he was *trying* to be clear. (Still, how much help is this? Almost everyone admits that a piece of philosophy may be *unavoidably* hard to understand because the things it is about are by their very natures hard to understand. Would Heidegger not tell us that his work is not unnecessarily difficult to read, that its difficulties are necessitated by his subject matter and the Forgetfulness of Being that has pervaded the European consciousness since Plato? I have seen scattered remarks by Sartre that indicate that he believed his own philosophical prose to be extremely clear.) But, as a general rule, analytical philosophers think that philosophical sentences should be as simple as possible and that words should be used in their everyday senses or else explicitly defined.
- Analytical philosophy, insofar as it draws on anything outside philosophy, draws on formal logic and mathematics and the physical and biological sciences (and to some measure from experimental psychology and linguistics). It does not, as a general rule, draw inspiration or material from literature or art or history or the more ‘humanistic’ parts of the human sciences. (I do not mean to imply that this is a fact that analytical philosophers should be proud of.)
- Analytical philosophy, insofar as it involves the defense of theses, places a high value on explicit argument, on pieces of text that are identified by the author as arguments and whose validity is to be judged by the rules of logic.
- Analytical philosophers have a particularly collegial relationship with the great philosophers of history. I mean this almost literally: analytical philosophers tend to regard Plato and Occam and Descartes as *colleagues*, albeit colleagues who labor under the burden of being dead. Let me

present the fictional Winifred to you. Winifred, although she is a fiction, is a typical analytical philosopher, and at the moment she is studying Kant's theory of freedom. To understand the relationship of this typical analytical philosopher to the history of philosophy, it is necessary to understand why she is studying Kant's theory of freedom. It is not because she is interested in Kant's philosophy *per se*; it is not because she is interested in the historical development of the idea of freedom; it is not because she wishes to study Hegel's theory of freedom and regards a mastery of Kant's theory of freedom as a prerequisite for that project; it is not because she is interested in contrasting the metaphysical understanding of freedom of a representative philosopher of the Enlightenment with a post-modern anti-metaphysical understanding of freedom. It is for a reason quite unlike any of these reasons. It is because Winifred is possessed of a childlike desire to understand the things to be found in the world, and one of these things is freedom. Her desire is that very desire mentioned in the famous first sentence of Aristotle's *Metaphysics*. She wants to understand freedom, and she is reading Kant because she thinks he may have something to teach her about this freedom. She would like to know whether Kant's theory of freedom is true or partly true or wholly false. If you tell her that her quest for the truth about freedom is indeed childlike, that it is not possible to undertake a project of that sort in today's post-something-or-other world, she will only smile at you. She will smile and ignore you or she will smile and treat what you have said as a thesis that is open to debate and proceed to present arguments against it, for all the world as if it were a thesis about second intentions and the two of you were opponents in a disputation in a medieval university.

I should add that although Winifred is a typical analytical philosopher, she does not represent the way every analytical philosopher approaches the history of philosophy. There are, of course, analytical philosophers who are historians of philosophy, and there are some reputable analytical philosophers who have no interest at all in the history of philosophy. But even those analytical philosophers who are historians have a tendency to approach the history of philosophy in ways that many non-analytical historians of philosophy regard as having the wrong aim, even a perverse aim. One non-analytical historian once directed the following charge at a former colleague of mine, Jonathan Bennett (who had written two books on Kant): 'Bennett is interested in Kant only insofar as he can extract from his writings concepts and arguments that are relevant to the philosophical controversies of the present day; he has no interest in Kant's philosophy itself.' Analytical philosophers (like the fictional Winifred or the real Jonathan Bennett) will reply that they know what it is to be interested in philosophical problems about space and time and knowledge *a priori* and causation, and they will say that they can understand why someone who was interested in these

things might be interested in what Kant had to say about them – but, they will say, they are less clear about what it is to be interested in something called ‘Kant’s philosophy itself’. And, they will go on to say, if there is such an interest as an interest in Kant’s philosophy itself, it is not a philosophical interest. After all, what *Kant* was interested in was philosophical problems about space and time and causation, and so on; *he* never wrote about anything called ‘Kant’s philosophy itself’. And, although Kant tells us that the philosophy of David Hume awakened him from his dogmatic slumbers, he was not interested in ‘Hume’s philosophy itself’. He was, rather, interested in what Hume’s philosophy was *about*. In this respect, at least, Kant was the reader Hume wrote for. Hume did not write for students of the philosophy of David Hume. He wrote for students of human nature and the human understanding. Analytical philosophers are (in this respect) like Kant and like Hume and like most other pre-Hegelian philosophers. When they read Plato or Occam or Descartes, their primary interest is not (not typically, at any rate) the philosophies of Plato or Occam or Descartes. Their primary interest lies rather in what the philosophies of Plato and Occam and Descartes were *about*.

This, then, is analytical philosophy, or as much as I am going to say about it. What is of central importance for this conference in what I have said about analytical philosophy is that, if what I have said is right, then being an analytical philosopher does not involve commitment to any philosophical doctrine. An analytical philosopher may be a platonic realist or a nominalist, may affirm or deny the freedom of the will, may believe in or deny the existence of an immaterial soul, may make the most dogmatic claims to knowledge or may embrace a thoroughgoing scepticism. An analytical philosopher may regard metaphysics as an illusion or be the most determined and ardent defender and practitioner of metaphysics imaginable. A philosopher may take any position on any philosophical question and still be an analytical philosopher in good standing. And this generalization applies to positions about the reality of a supernatural order. An analytical philosopher may (as I do) recite the Nicene Creed with conviction every Sunday or regard the Nicene Creed as an absurd vestige of a pernicious and dying superstition: nothing in the nature of analytical philosophy lends any support whatever to either of these positions. Analytical philosophy, by its very nature, is neither the friend nor the enemy of supernaturalism. And, by the same token, analytical philosophy can be neither the friend nor the enemy of naturalism.

10 Resisting naturalism

The case of free will

Hugh J. McCann

It is no easy matter to define philosophical naturalism, but I am going to take it in a very simple and straightforward way. I shall understand naturalism as the view that the problems of philosophy are to be addressed, and their solutions framed, in terms acceptable to natural science, particularly physical science. Naturalism has both a negative and a positive thrust. The negative part is that to the extent the concepts philosophers employ have no home in the language of physics – if they are irreducibly mentalistic, say, or value laden, or peculiarly religious – those concepts have no bearing on the world of ordinary experience, which (it is implied) is the only world of which we have knowledge. Right-thinking philosophers will therefore eschew them. The positive thesis is that the concepts of the physical sciences will be found to suffice: that is, that they can and ultimately will provide a thorough and sufficient understanding of all that we experience. I think it is fair to say that naturalism represents a kind of orthodoxy within analytic philosophy. But its origins are much older, and like most orthodoxies it has seldom come close to achieving universal acceptance. One thing that impedes its success is the free will problem, which has always resisted solution in terms congenial to naturalism, and continues to do so. The reason is essentially twofold: friends of naturalism have never succeeded in providing a deterministic solution to the problem, and an indeterministic or libertarian solution violates the tenets of naturalism. In this paper, I shall chart the case for these claims, and then say just a bit about what they imply about the prospects for analytical philosophy without naturalism.

Some history

Naturalism as we know it today has its roots in the scientific revolution and the philosophical movements that emerged in that era. These tended to portray the physical universe as a grand mechanical contraption, in which all that takes place does so in accordance with a few deterministic laws. This is, of course, a long way from our present conception of the physical world, and it leaves open some questions on which contemporary naturalists have strong views. Notoriously, it is silent on the relation of the mental to the

physical: about whether psychological phenomena are merely aspects or by-products of the material world, or instead make up a separate realm, and possess their own peculiar efficacy. The image of the universe as machine is mute about other issues as well, for example, about the origin of the universe, the development of life, and, at least to some extent, the status of values.

But there were some subjects on which the scientific revolution spoke with unmistakable clarity, and what it had to say left little or no room for free will. The chief of these subjects was what counts as an appropriate explanation. For the new sciences, the explanation of physical phenomena was to be mechanistic, not teleological. A good explanation would proceed not in terms of supposed purposes for which things existed, of ends they naturally sought, or of goods their behavior achieved, but in terms of diachronic relations of cause and effect. Things moved as they did because they were impelled to do so by prior causes, whose operation was presumed to be deterministic, precisely measurable, and completely describable in mathematically formulated laws. There was, of course, the question of how far those laws would be found to reach, of how much of our experience would prove describable as just a matter of bodies in motion. But, then as now, the success of science inevitably prompted efforts to assimilate as much as possible to physical or quasi-physical models. Human action was a legitimate target for such efforts, because paradigm cases of human action are at least partly physical: I cannot move my arm without the arm moving; I cannot fire a gun without the gun firing. But any attempt to deal with human action in a mechanistic way was bound to produce a clash over free will.

The primary reason for this was, of course, the issue of determinism itself. Free will requires that the formation and execution of intention be, as we say, 'up to the agent'. That is, it requires that human decisions and actions involve the exercise of a certain sort of power – a power that appears, at least, to be contracausal. When the cue ball strikes the object ball in a game of billiards, it is not up to the object ball how it moves, precisely because the conditions in which it is placed, together with the features of the collision, determine completely what is to happen next. Only if this were false could anything be up to the object ball. The same goes for human decision and action. If I am deciding where to go after this conference, the condition of my mind does guarantee some aspects of the outcome. I will not decide to go somewhere I have never thought of, or to a place I have no reason to visit. And because these things are settled by my mental condition, they are not up to me. But my mental state seems not to guarantee everything. If more than one option presents itself to me as having something to offer, then (most of us would say) it *is* up to me which one I choose. This, we would say, is part of what it means to claim I am free, and it suggests that I am very much different from a billiard ball. It suggests, at least, that the condition of my mind does not determine what I decide – that I might have had the very same motives and beliefs, and yet have decided differently.

There is a second point on which the mechanistic view conflicts with free will. When I decide or act freely, what I do can only be made sense of by an explanation that is explicitly psychological. If, for example, I decide to visit France after this conference, I will do so because such a visit makes sense to me, and the sense it makes is what will explain my decision and subsequent action. The real import of this point is obscured in the Cartesian setting, where it gets tangled up with the question of how events conceived to occur in an immaterial soul can have material consequences. In my view, that is not the main issue. Rather, as I shall argue below, ordinary explanations of action do not proceed in terms of events, even mental events. They proceed in terms of the content of mental events, in terms of envisioned outcomes, portrayed as desirable. They explain decision and action as undertaken by the agent in pursuit of an end deemed worth having. This runs counter to the mechanistic approach, first, because mental content is very difficult to treat plausibly in naturalistic terms. Second, this kind of explanation is explicitly teleological – the very thing the proponents of the new sciences found most objectionable in traditional approaches.

Philosophers of the seventeenth and eighteenth centuries had few tools for attempting to analyze the relation of the mental to the physical, and so made little in the way of a concerted effort to explain away the mental dimension of behavior, or to reduce it to processes that could be viewed as acceptably physical. There was, however, great respect for the idea of determinism, at least as applied to the physical realm – enough respect that among major thinkers of the era, few were prepared to defend free will in a form that would threaten it. A notable exception is Thomas Reid, who, rejecting the mechanistic picture entirely, boldly asserts that only minds have genuine causal power (Reid 1969: I, Chapter V), that in exercising that power they are undetermined (*ibid.*: IV, Chapter VI–VIII), and that the ‘motives’ that explain behavior are not events at all but, as I just suggested, *abstracta* (*ibid.*: IV, Chapter IV). If these claims are true, and if our actions in the physical world accord with our wills, then the physical realm cannot be expected always to behave deterministically, and the explanation for what occurs in it will not always be causal. The approach of Immanuel Kant, though more radical in some ways, is far less combative. For Kant, freedom pertains only to the noumenal realm: it is a necessary presupposition of deliberation and action, but strictly a practical postulate, not an item of knowledge in the theoretic sense (Kant 1998: A 538–541, B 566–569). On the contrary: to observation action will always appear deterministic, since in the phenomenal realm the law of causation is universal. The effect of this strategy is to affirm free will, but to sequester it from the world of mechanistic causation. We may in practical contexts take ourselves to be free; but empirical observation, which alone provides legitimate knowledge, must always claim the opposite.

But, of course, Reid and Kant were swimming against the tide. The most common response to the mechanistic picture of decision and action –

manifested prominently by Locke and Hume in Britain, and by Leibniz on the Continent – was to accept it, and to postulate a weakened view of freedom that would comport with determinism. Thus Locke speaks of liberty as a power to do or forebear a particular action according to the determination of the mind (Locke 1961: Book II, Chapter XXI, Section 8), Hume gives substantially the same characterization, calling it a ‘hypothetical liberty’ given which if we choose to remain at rest, we may, and if we choose to move, we also may (Hume 1955: Section VII, Part I). For both authors, ‘liberty’ is no more than freedom of action – that is, freedom to act as we choose, or in accordance with what we will. As for the will itself, it is determined by the agent’s strongest motive, and neither philosopher accords any explicit freedom to the will as such; indeed, Locke is at pains to reject this whole idea, which he says leads to an infinite regress of willings to will (Locke 1961: Section 23). In fact, however, one can adapt the Lockean analysis of freedom to the will also, holding that decision and volition are free even though ruled by the agent’s strongest motive, provided only that a dominant motive in the opposite direction would have caused an opposite result. This is in effect the strategy of Leibniz, who, though troubled by Locke’s views on the will itself (Leibniz 1949: Book II, Chapter XXI, Section 8), finally takes an essentially similar stance, holding that the only necessity that would destroy moral responsibility must be an ‘absolute’ necessity, such that our actions would occur no matter whether we wish them or not. The conditional or ‘hypothetical necessity’ through which our motives determine our behavior is not of this kind, for in this case a different motive would cause a different outcome (Leibniz 1985: 381). Thus, Leibniz was able to assert, somewhat misleadingly, that motives incline the will without necessitating it (*ibid.*: 382). What this means, really, is that determining motives do not render our conduct *logically* necessary. And since they do not, Leibniz is prepared to consider us free. He would not deny, however, that our motives necessitate our wills hypothetically; indeed, the very thing Leibniz calls ‘hypothetical necessity’ is what Hume calls ‘hypothetical liberty’.

Some more recent history

It was against this background that analytic philosophy concerning the will and action took shape. Early analytic philosophers were little influenced by European philosophy after Kant. But they had great respect for the empiricist tradition, and shared much of its motivation, including great respect for science; indeed, at least in the beginning, the prevailing orthodoxy was that empirical observation was our sole source of all but conceptual knowledge, and that since philosophers do not engage in systematic observation, they ought to confine their endeavors to the conceptual realm. And I think it is fair to say that for about half a century, it was largely so – first with the logical positivists’ project of formulating an ideal empiricist

language for science, and later with the ordinary language movement, which sought to demonstrate that philosophical problems rested on confusion that could be alleviated through conceptual analysis. It was not to be expected, however, that this austere self-discipline would last very long – especially in light of the fact that the enterprises of both positivism and ordinary language philosophy met with little success. Inevitably, the penchant of philosophers for building their own view of the world would reassert itself, so that today little remains of analytic philosophy's once-vaunted disdain for metaphysics, or its ostensible ban on constructive theorizing.

But a lot of things do remain, many of which were there from the beginning – among them, ironically enough, a prevailing metaphysics: that of naturalism. For even when it tried its hardest to eschew metaphysics, analytic philosophy was guided, as any approach to philosophy must be, by a picture of the way things are, and that picture was of a world congenial to physical science. Here, however, there was a second irony, for while science had undergone dramatic change, the picture had not changed at all. The analytic movement was in part precipitated by a new scientific revolution, one dimension of which was the discovery that the behavior of subatomic particles seemed not to be deterministic after all, but was instead only probabilistic. One might have thought that this development would be welcomed by philosophers of action, as setting the stage for theory that would at once preserve free will and harmonize with the latest science. Indeed, precisely this suggestion came from some members of the scientific community.¹ But the idea did not gain great currency among philosophers, for reasons about which I shall speculate below. Rather, it was the older, mechanistic conception that dominated, and the strategies by which the notions of freedom and responsibility were to be brought into conformity with it were mostly those of the earlier era.

Things were not entirely the same. The Kantian strategy of sharply distinguishing the theoretical and practical realms and then confining notions of freedom to the latter was not much favored, especially in the early days, and it was never endorsed in the form Kant gave it. A version of this strategy did, however, gain favor among certain philosophers of the ordinary language movement – a typical example being A.I. Melden (1961). The essence of this approach is to apply the Wittgensteinian notion of a language game to discourse about human behavior. Such discourse may occur in any of several games, or patterns of discourse. Which game is going on depends on the interests of the speakers in question, and the concerns of speakers engaged in different games need not intersect. Thus, to a scientist interested in human behavior, what is observable when I raise my arm is a mere physical happening: the arm's going up. He will be interested in the inner, physiological causes of this event, and will seek to construct a deterministic explanation of it in terms of them – all of this, by Melden's lights, a perfectly legitimate endeavor. The concerns of everyday discourse, however, are altogether different. To the ordinary speaker, it was alleged,

the very same event which is my arm going up counts as a human action: my raising my arm. As such, it belongs to an equally legitimate framework of explanation, but one to which the language of deterministic causation is entirely alien. We explain actions in terms of the agent's reasons – that is, by citing his motives and intentions, which for Melden are *logically* related to action, and therefore cannot possibly cause it. And depending on how the agent was motivated – for example, on whether he was operating under duress – we may or may not decide that he could have behaved differently. But this decision belongs to the language game of action, not that of mechanistic causation. Indeed, even to raise the question whether an *action* is causally determined is to be guilty of a kind of logical howler, to mingle together incommensurate forms of discourse, thereby creating a false appearance of conflict. If we follow the proper course we will confine each language game to its peculiar sphere, and the conflict will be banished.

This maneuver is less radical, but the analogy between it and the one Kant attempted is obvious: if we cannot sequester the operations of voluntariness themselves from the arena of causality, we can at least sequester talk of what is voluntary from talk of what is caused – and so hope to reach a peaceful settlement between proponents of determinism and defenders of free will. Popular though it was for a time, however, this was not the settlement at which most analytic philosophers aimed. At least through the mid-twentieth century, most opted for what Kant had pronounced a wretched subterfuge (Kant 1956: 99) – that is, for a compatibilism which, though often more sophisticated in detail, was at bottom little changed from the heyday of the Enlightenment. A free action was an action that would have gone otherwise had the agent chosen, or perhaps tried, to do otherwise; a free decision was one that would have been different had the agent possessed adequate or sufficient motivation to decide differently. To say an agent could have chosen or done otherwise, it was urged, is simply to invoke these conditional relationships. And to hold people responsible for their deeds is to take advantage of the supposed deterministic sequence from motive to overt behavior, hoping to alter the effect by altering the cause. In one version or another, this type of compatibilism was defended early on by G.E. Moore (1912: Chapter VI), later by authors such as R.E. Hobart (1934), Moritz Schlick (1939) and A.J. Ayer (1954), and still more recently by Alvin Goldman (1970: Chapter 7) and Donald Davidson (1973).

It is fair to say, I think, that neither of these strategies met with much success. The most obvious failure was the sequestering maneuver, which ran afoul of the fact that no matter how diverse the language games of action and of physiological psychology may be, they speak of the same phenomena, and in terms that contradict. It may or may not be correct to say that when I perform the action of raising my arm, my action is identical with the event of my arm rising.² But it is true that when I raise my arm I normally

take myself to be bringing about my arm's rising, and to be doing so freely, so that its occurrence is up to me. Were it not so, I would not consider myself responsible for the event, or its consequences. By libertarian lights, however, my arm's rising will not be up to me if, as the physiologist says, that very event is caused by occurrences in my brain, which are themselves no more than a barely noticeable part of the vast causal network that constitutes the physical world, and whose present state Laplace's demon could have predicted with certainty a thousand years ago. So while players of the physiological language game and players of the human action language game may certainly agree to ignore one another, that doesn't change the fact that both groups cannot be right: in conversing among themselves, each group contradicts what the other has to say about human behavior.

The efforts of analytic philosophers to defend classical compatibilism also encountered serious obstacles. In his famous paper 'Ifs and Cans', J.L. Austin pointed out that not all *if*-clauses in English are conditional – so that even if, as G.E. Moore had maintained, to say 'I can' is as much as to say 'I can, if I choose', it would be wrong to jump to the conclusion that statements about what we can do have reference to a conditional relation between choice and action (Austin 1961). Even if such a relation is involved in freedom of *action*, moreover, it seems insufficient to capture what we have in mind by freedom of the *will*, or the ability to do otherwise. Thus, Roderick Chisholm, though not fully persuaded by Austin's rejection of compatibilism (Chisholm 1964), still thought a decisive argument was available. That an agent could have done otherwise cannot, he claimed, mean no more than that he would have done otherwise had he so chosen. For it may yet be the case – and if determinism holds, it certainly *would* be the case – that the agent could not have chosen otherwise. And if that is so, then he still could not have done otherwise (Chisholm 1997: 146–147). The essential insight here is that freedom of *the will* pertains not to the question what obstacles an agent may face in achieving the ends he wills, but rather to the etiology of willing itself. It may well be that if I choose to visit France after this conference, nothing will prevent me. But it seems empty to claim on this ground that it is up to me whether I will go there if my decision to go is settled in advance, a mere element of a uniformly causal framework in which all that takes place has been determined from the world's beginning. If my choice is the inevitable consequence of events over which I have no control, then it too is beyond my control. Peter van Inwagen has shown that this kind of argument can be made perfectly general (van Inwagen 1975). If determinism is true, then all that we do is settled by the distant past and the laws of nature, neither of which are in our power to alter. But if we cannot alter these, then we cannot alter what we do. It follows that there is no freedom; we can never do otherwise – and so, it seems, are never responsible.

There is, however, a third strategy to consider, unlike the first two in its claims about freedom, but still compatibilistic in spirit. Perhaps partly in

response to criticisms of classical compatibilism, a number of philosophers have followed Harry Frankfurt in maintaining that if by freedom is meant the legitimate possibility of doing otherwise, then responsible agency does not require free will after all (Frankfurt 1969). This, they argue, is because the conditions that rule out alternative possibilities need not enter into the actual etiology of choice or action – so that despite their presence, agency is still exercised normally. Suppose, for example, that I decide in the normal way that following this conference, I will visit France. Since no obvious compulsion is at work in my decision, we would doubtless take it as an exercise of free will. Suppose, however, that my old friend Pierre, who is bent on having me visit France, has implanted a device in my brain that would enable him to force me to decide to do so. As it turns out, Pierre doesn't have to use the device in this case, since I decide that way anyhow. But, the argument runs, Pierre also has means of knowing what I am about to decide; and had I been about to decide differently, he would simply have pushed a button on his hand-held computer, thereby forcing me to decide to visit France. This is a case, then, in which I could not have decided differently. Yet, given what actually occurred, it appears I decided responsibly.

The appeal of such examples is that if they are legitimate, we may approach issues of moral responsibility in an essentially compatibilist spirit, while avoiding implausible claims about what constitutes freedom. For even though, in our example, Pierre is poised to control my decision if need be, the decision actually occurs without his intervention. Accordingly, the argument runs, my faculties of deliberation and decision can be assumed to have operated normally. And as long as the etiology of my decision is normal, I can be held responsible for it. We need not worry about whether my choice was determined from the remote past, or about whether I could have chosen differently, if that is what constitutes free will. For, it is alleged, the example shows that this kind of freedom is not necessary for moral responsibility.

But are such examples legitimate? Several philosophers, most prominently David Widerker,³ have argued plausibly that they are not, at least from a libertarian perspective. We may assume that it takes some time for a decision to take place – though doubtless it requires very little – and let t be the moment in our example when I commenced to decide to visit France. The question is, might I have done otherwise at t ? Might I, for example, have commenced to make some other decision? The answer depends on how Pierre would have known that I was going to do so. If he would have had to wait until t – that is, until the moment I commenced to decide not to visit France – then the example fails. For in fact I could have done otherwise: I could have undertaken to decide not to visit France. I might not, of course, have completed the decision if Pierre was able to interfere quickly enough. But that does not matter. Freedom is freedom, no matter how brief its exercise. In order for the example to succeed, therefore, Pierre must have some reliable sign *prior to t* as to what I will decide – that is, there must be

some signal whose occurrence prior to t is *guaranteed* if I am about to make any other decision. But if the signal is guaranteed, then its occurrence is in essence a causally necessary condition for my deciding not to visit France. Consequently, although it is true that I could not, at t , have decided differently, the reason has little or nothing to do with Pierre. It is because a causally necessary condition of my doing otherwise – namely, the prior occurrence of the signal – was lacking. What this indicates is that Frankfurt-style examples can be constructed only by begging the question against libertarian notions of freedom. If the claim that a person could have done otherwise means what libertarians take it to mean, such examples would appear to be impossible,⁴ and if that is so then the third strategy by which analytic philosophers have sought to develop a theory of responsible action in line with deterministic principles also fails.

Indeterminism and naturalism

One cannot, of course, consider the matter closed; there is always the possibility that some future effort to bring peace between determinists and free will advocates will be found to succeed. For the present, however, the hostilities continue. Yet one may wonder why allegiance to determinism is so popular among those who would favor a naturalistic account of human behavior. It was mentioned above that part of the inspiration for the analytic movement in philosophy came from developments in science, among them the rise of quantum mechanics, according to which subatomic phenomena do not obey deterministic principles. We should note that this claim at first encountered considerable resistance among scientists. Today, however, quantum theory is solidly entrenched, and if it is true there is reason for thinking that, at least as far as indeterminacy goes, libertarian free will can in principle be brought into line with naturalism. The key requirement is that quantum-level indeterminacies, which ordinarily cancel one another out at the macro-level, be able to have impact at the level of neurological firings – that is, at the level of the physiological processes we take to correspond to mental occurrences. If this were so, then the neurological events that correspond to acts of decision and volition might after all turn out to be indeterministic, so that exercises of free will would be secured as to their physiological counterparts. Robert Kane has argued that this kind of impact is not at all implausible to imagine. His suggestion is that, especially in situations of moral temptation, the tension of the deliberative process corresponds to a chaotic situation in the brain that is able to amplify quantum-level perturbations, so that the neurophysiological sequence that corresponds to the agent's decision is rendered indeterminate (Kane 1996: 129–130). How likely it is that this particular suggestion will turn out correct is, of course, anybody's guess. But the sources Kane cites indicate that if chaos theory can be wedded to quantum theory, macro-level sensitivity to quantum-level events could turn out to be widespread in the

nervous system; and if that is so, then some account along the lines Kane suggests may well be found to succeed.

Here, then, is a possibility that deserves to be explored. Yet, at least until recently, it has received rather little in the way of philosophical attention, and is still not much discussed. Part of the reason, certainly, has been the state of scientific knowledge. Until the development of chaos theory, no one had any clear idea how quantum perturbations might be magnified so as to be reflected in the physiology we think goes with decision making. And even if we view this problem as now on the way to being surmounted, we remain woefully ignorant of the precise details of the physiology in question. We know a little about the brain functions involved in deliberation and decision, but we are nowhere close to being able to say what specific brain events might correlate with an act on my part of deciding to visit France, and how they might differ from those which would correlate with my deciding to remain in Italy. Any effort along the lines Kane proposes would therefore have to be considered speculative and uncertain. But I do not think this is the whole answer, for in my experience philosophers who find the idea of a quantum mechanical treatment of free will attractive tend not to have strong naturalist sentiments anyway. Rather, they are first and foremost defenders of free will; and they find the quantum mechanical option attractive not as a step toward a naturalist theory of human decision and action, but as a partial counter to the objection that libertarian freedom is an irrational notion, one which cannot be squared with a properly scientific view of the world. By contrast with this perspective, I suspect that anyone truly dedicated to developing a naturalist theory of decision and action would find the quantum mechanical gambit rather unattractive. Why should this be?

The answer, I think, is that quantum mechanics is in itself offensive to the sentiments that underlie philosophical naturalism. Not that it doesn't fit with the definition of naturalism adopted at the beginning of this paper. Quantum mechanics is certainly physics, and Kane's views certainly represent a step in the direction of treating free will in terms of science as we know it. The problem is that in this case science as we know it runs counter to what I described earlier as the positive thesis associated with naturalism: that the concepts of physical science will ultimately provide a thorough and complete understanding of all that we experience. Even for strictly physical phenomena, quantum mechanics does not do that. Just the opposite: it is fundamentally statistical, and on the usual interpretation finds the world to be irreducibly probabilistic. If it is true, then nothing recognizable by contemporary standards as physics is going to be able to tell us why a particular atom of radium decays when it does, or why a particular photon in the classic double-slit experiment strikes the reflecting screen where it does. Is this an embarrassment for physics? Not really, in my opinion. The world is what it is, and if quantum theory describes the world correctly, then as science it must in the end be considered not just a success but a triumph.

For naturalism, however, quantum mechanics represents a defeat; if it is true, then a central thesis of naturalism is false.

This, I suggest, is what explains the hesitancy of naturalistically minded philosophers to embrace the idea of a quantum-theoretic treatment of the underlying physiology of decision and volition. Even if such a treatment were found to succeed it would be a disappointment; it would leave us without a full, scientific accounting for exercises of human agency. Worse yet, from the naturalist's perspective, it would invite philosophers of other persuasions to try to fill the explanatory gap with some sort of teleological explanation in terms of the agent's motives or intentions – something of which contemporary naturalists are no more fond than were their empiricist forebears. I shall have more to say below about such explanations, which I think are legitimate and valuable. In fact, however, I don't think the naturalist need fear that they will close the gap in our understanding of decision and action that would be left by a failure of determinism. If we are free, then no consideration we can cite will provide the kind of explanation for decision and volition that a determinist would want – that is, an explanation that closes off all alternative possibilities. All the same, a good, orthodox naturalist should find the association of free will with quantum indeterminacy an unwelcome prospect. Should it succeed, it would cement quantum theory in place more firmly than ever. And it would make the uncertainty naturalists are wont to complain about when it comes to free will part and parcel of an uncertainty that goes straight to the bottom of reality, physically understood.

Naturalism and psychological explanation

But perhaps the naturalist will be prepared to make a concession at this point. After all, he may reason, the success of quantum mechanics makes it unlikely in any case that a completely satisfying physical account of the world will be found. The discomfiture this brings to naturalism may be increased if quantum uncertainty turns out to be manifested in human decision and action, but the basic reality would be unchanged. Moreover, the really important tenet of the naturalist's creed – that the world is to be understood in terms of physics – could end up surviving such a discovery fairly nicely, just as it survived the discovery of quantum mechanics. After all, to the extent we understand quantum phenomena at all, we do so in terms of concepts that are recognizably and strictly physical. And, the argument would run, the same will eventually prove true of human behavior. Determined or not, human decision and volition will finally be understood as physical phenomena. That is, they will ultimately be reduced to physiological processes in the human brain – phenomena whose explanation, however complete or incomplete, will proceed in physical terms, the only terms appropriate to legitimate explanation.

It is, I think, on this point that we that we begin to approach a final and unbridgeable impasse between naturalism and libertarian free will. The

reason for this has to do with consciousness. From a libertarian perspective, the real value of free will lies not in the mere fact that it separates human willing from the natural causal order, but in the fact that this separation makes it possible for individual agents to exert some control over their destinies, to guide their lives according to purposes that *they* set. Paradigmatically, this is accomplished through conscious deliberation – that is, by the agent envisioning possible futures for himself, and deciding among them. And it is vital that this process be conscious. That is one reason why we never hear of anyone reaching a decision in dreamless sleep, or forming an intention while in a coma. It makes no sense to think of an entity that is not conscious envisioning a future, or finding in it any value worth pursuing. The naturalist position, by contrast, is that consciousness does not matter, that a true comprehension of the relation between an act of will and the reasons for which it occurs are to be had in terms that are physiological, not psychological. Even if the relation between reasons and willing turns out not to be fully deterministic, therefore, the naturalist can still insist that that relation will one day be replaced in our theories by relations between the physiological processes to which reason and willing are reduced. Only when this occurs, he will say, can we have a scientifically respectable account of human decision and action.

I think, however, that the difficulty here runs far deeper than the naturalist, or at least the naturalist I have depicted, takes it to do. This can be seen by considering what it is that counts as a ‘reason’ for which an action might be performed. Naturalists – and often their opponents too – are disposed to think of a reason as a mental state or event with the potential to give rise to action. Thus, the event of my experiencing a desire to visit my friend Pierre might be viewed as a reason for which I might decide to visit France. Should I proceed to make the decision, the event of my feeling the desire to see Pierre would be expected to figure prominently in the explanation of my decision – a deterministic explanation if determinism turns out to be true after all, or some kind of statistical account if it does not. In fact, however, what is treated in common life as a reason for deciding or acting cannot figure in either of these sorts of explanation. A reason for deciding has to be something I can *think* – and I cannot think an event, even the mental event of my experiencing a desire to visit Pierre. Rather, I must think the *content* of the event – something we might express in the optative mood as: *Would that I visit good old Pierre*. This, unlike the event of my thinking it, is an abstraction or *ens rationis* – a thought, rather than the thinking of it. We have already taken note of Reid’s claim that this kind of entity cannot figure in a causal explanation of an act of will. He was right about that: a cause has to be an event or state in the real world, not an abstract entity, and the same goes for anything we might point to as ‘statistically causing’ or explaining something.

When we speak of desires as explaining decision or action, then, what is at work is neither a causal nor a statistical form of explanation. What kind

of explanation is it? Obviously, a teleological one, an effort to account for what the agent does in terms of the envisioned objective at which he aims. It is appropriate to think of such explanations as encapsulating the agent's own practical reasoning. Thus, if we explain my decision to travel to France by citing my desire to visit Pierre, we treat the objective whose ostensible fitness for achievement I apprehend through my desire as the ground for my decision to pursue it. The reasoning by which I reach my decision can be summarized in a kind of Aristotelian practical syllogism:

Would that I visit Pierre.
If I travel to France, I can visit Pierre.
Therefore, I shall travel to France.

Here my decision is presented as justified by the objective my desire sets before me, in that by deciding to visit France I form an intention which, if carried out, will count as a step toward visiting my friend, a goal which will eventually be achieved by timely formation and execution of further intentions directed toward it.

There is much to be said about this kind of reasoning, but I will content myself with two points. First, this reasoning is practical not just as to content, but functionally as well. It does not end as 'practical reasoning' is sometimes held to end – that is, in a judgment as to what is best, which is then left to blind chance (or should I say blind causation?) as to whether it will lead to intention formation. Rather, this reasoning terminates in decision itself, which *is* intention formation. It is too much to go into here, but I think similar reasoning, terminating in the activity of volition, is what occurs in the execution of intention – so that both these types of active willing count also as ratiocination, or drawing a reasoned conclusion. Second, there is a clear and very satisfying logic to this kind of reasoning. This has eluded many philosophers, because they tend to assess all reasoning along epistemic lines – that is, as though it were an effort to learn more about the world, by drawing the right conclusions from independently confirmed evidence. On this understanding, there is no way the practical syllogism cited above can be judged valid or cogent. How does one deduce an intention – that is, the content of my decision – from the optation, 'Would that I visit Pierre', and a belief about how to move toward satisfying that optation? Clearly, one cannot; there are no logical principles for deducing intentions from desires and beliefs. This approach to evaluating our practical syllogism is, however, misguided. The function of intention formation is not epistemic; it is not to bring the mind into conformity with the world. Rather, it is to bring the world into conformity with the mind: to form and carry out intentions directed at changing the world to suit our preferences. Viewed in this light, our practical syllogism makes perfect logical sense. For if its conclusion – that is, my decision to travel to France – is successfully carried out, and if I am correct in believing that this will put

me in a position to visit good old Pierre, then by good old *modus ponens*, I will be that much closer to accomplishing my objective. Or, to state the point a bit more formally, if the minor premise and conclusion of the syllogism are satisfied, the major premise will progress toward being satisfied as well – exactly the relationship we should be looking for if the ‘direction of fit’ is world to mind, rather than vice versa (Kenny 1975: Chapter 5, especially 80–82).

Brief as they are, these points begin to illustrate the advantages of realizing that when, in common life, we speak of the reasons that explain a decision or action, we refer to what are in fact abstractions, not states or events. To view the operations of voluntary willing in this way is to see them as ratiocinative processes, wherein responsible agents guide their conduct toward consciously chosen ends. It is far from clear, however, that any of this can be done justice on a naturalist approach. From that perspective, my desire to visit Pierre would likely be viewed not as content but as the thinking of the content – that is, as a psychological event, to be understood in terms congenial to science. In recent years, that has meant submitting the event to functional analysis. That is, my desiring to visit Pierre would be treated as though what is essential to it *qua* mental is not the content of which I am conscious when it occurs, but rather the way it interacts with the rest of what goes on in my head when I deliberate over where to go after this conference. Having completed the functional analysis, the naturalist would next point out that what goes on in my head must after all be physiological, so that the desire-event may be identified with whatever brain process is found empirically to correlate to it, no mention of conscious content being thereafter necessary. Or, it might be held that any conscious goings on associated with it are at best supervenient, and so count only as epiphenomena. And, of course, the same treatment can be given for my belief as to how I may go about visiting Pierre, and my decision to travel to France to see him. Finally, it will be pointed out that since all of these events are at bottom physiological, the explanatory relations that join them, whatever they are, must be physiological as well. Thus, the naturalist program is fulfilled.

I think it is obvious, however, that this approach to understanding decision making will not satisfy defenders of libertarian freedom. For even if the explanatory relations of which it speaks turn out not to be deterministic, they leave conscious content completely out of account. Yet, as we have seen, there is no point to libertarian freedom, except as a precondition for agents to play a meaningful, autonomous and responsible role in carving out their personal destinies – something that is possible only on the supposition that they can entertain and select objectives. The libertarian will argue that it is vacuous to think someone could prosecute such an enterprise except by working with conscious content – vacuous to hold, in effect, that a zombie could as truthfully be said to make decisions as you or I. Moreover, he will say, it is a mistake to think the content of our desire and belief states is

merely epiphenomenal. The simple truth is that citing the desires and beliefs out of which a decision is made does provide understanding. A complete account of the operations of the will has to explain why it does so, not ignore or pass over the fact.

What all of this comes to is that, in the context of science as we know it today, the really important disagreement between the naturalist and the libertarian is not about indeterminism, but about the reality of consciousness, and the role it plays in behavior. And I do not see how this disagreement can be overcome. To do so would require two things. First, it would take a functional account not just of mental states and events – of the thinking of content – but of content itself. And of course there have been efforts along these lines. I cannot consider them here, but they face significant obstacles. Mental items such as desires, beliefs and intentions are not concrete entities but *abstracta*. They are, that is, proposition-like, and not to be identified with any kind of sign or physical representation, even in the brain itself, of which they might be held to constitute the meaning or significance. Furthermore, entities like this share logical relations, not functional ones, and it is in terms of these relations that they are explanatory. Thus, although the mental events of which they are the content must certainly play a functional role in the etiology of decision and action, the explanatory role of the content cannot reduce to the functional role of the events. Indeed, one could go further: there is no obvious reason why the functional role of the events could not turn out to be grounded in the explanatory relations among their contents, rather than the other way around.

But even if these points can successfully be addressed, there remains a final and I think insuperable barrier separating the libertarian from the naturalist: the fact of consciousness itself. The simple truth is that there is, in the phrase now popular, something it is like to be an agent, and a sensible libertarianism must maintain that were this not so, free will would be an empty shell. Whatever else might be said about mental content, it is only *as presented to the agent* that it makes possible the significant exercise of moral autonomy. Yet this final thing, consciousness itself, is the one thing naturalism cannot accommodate. To include it in our theories – to include, that is, what I actually experience when I desire to visit Pierre – is to admit an entirely new sort of datum, a kind that cannot be represented in the vocabulary of physics. And to admit that kind of datum is to give up being a naturalist.

A final note

If the foregoing arguments are correct, a libertarian account of free will cannot be a naturalistic one. Is it, then, the case that libertarian views of the will are not to be considered analytic philosophy, in the truest sense of the term? It is certainly possible to make such a claim. One need only characterize analytic philosophy narrowly enough that it turns out to

exclude any view of the will that is not strongly naturalistic. And there may have been a time – in the days of the Vienna Circle, perhaps – when such a characterization would have had a certain persuasiveness. But those days are long gone. The radical claims of logical positivism and the ordinary language movement have today very few defenders, and the movement that began with figures like Frege, Russell and Wittgenstein has long since broadened its scope to recover much of the historical heritage it once pretended to leave behind. There remain, to be sure, prevailing orthodoxies, naturalism among them. But the tradition of discourse that today goes by the name ‘analytic philosophy’ has by now come to encompass many perspectives, and is easily tolerant enough that one can comfortably consider oneself both a defender of libertarian freedom and an analytic philosopher. After all, I am one myself.

Notes

1 See, for example, Compton (1935).

2 I have argued elsewhere that it is not correct. See McCann (1998, Chapter 3).

3 Widerker (1995). See also Kane (1996: 142–143).

4 For further discussion of this issue see Widerker and McKenna (2003).