

## Ethics, Autonomy, and the Grounds of Belief *Christopher Norris*

### I

How far can or should we be held responsible for what we believe? From one point of view – call it ‘doxastic voluntarism’ – such responsibility is the *sine qua non* of any approach that would take due account of our various intellectual and ethical obligations, that is to say, our proper concern as rational agents with the business of sorting true from false or morally acceptable from unacceptable beliefs. From another, more inclined toward some form of doxastic determinism, it has seemed nothing short of self-evident that beliefs are to a large extent non-volitional, or subject to various kinds of causal or socio-cultural influence. Where the former stakes its claim on our capacity for freely-willed, autonomous choice in keeping with the dictates of moral or intellectual conscience the latter requires that we take more account of those other (heteronomous) factors that may limit or constrain the extent of our responsibility in this regard. Moreover, the determinist will then remark that there is a problem for the advocate of free will or doxastic autonomy if their exercise is taken to involve submission to overriding imperatives such as those of dedication to truth, valid inference, evidential warrant, or openness to persuasion by the best (most rational) argument. For in that case they would be subject to constraints of a different but no less binding character, namely to the norms of epistemic good conduct or respect for those same (in this sense heteronomous) standards of rational debate.

At its crudest this response takes the form of postmodernist jibes to the effect that Kant’s great watchword *sapere aude* – ‘think for yourself’, ‘let reason be your guide’ – is a plain performative contradiction, proclaiming the virtues of intellectual freedom while enjoining a strict compliance with its own demand.<sup>1</sup> It is not hard to see the confusion here between the general advice that we should strive so far as possible to exercise our powers of reflective, self-critical, conscientious thought and the specific injunction – no part of Kant’s claim – that we should think just like him. However there are real problems to be faced when the advocate of free will (or doxastic responsibility) attempts to explain how we can reconcile those values with the fact that any freedom of intellectual conscience worth having must involve a commitment to reasons or principles that will then play a crucial determining role in our various beliefs and actions.

Otherwise the notion of freedom will reduce to that of sheer randomness or unmotivated chance occurrence, as with certain, in my view misconceived arguments that adduce the indeterminate or probabilistic character of events on the subatomic (quantum) scale as evidence that moral philosophy no longer has anything to fear from the old Newtonian bugbear of iron-cast physical determinism.<sup>2</sup> Quite apart from their dubious scientific credentials such arguments clearly invite the charge of leaving no room for the exercise of a responsible freedom, as opposed to just a notional ‘freedom from’ the otherwise all-encompassing laws of physical cause and effect. Still the autonomist may be hard put to make her case against various objections that are apt to arise when considering the extent to which cultural, religious, ideological, or other such formative influences may play a predisposing or determining role even – or especially – in the case of our most deeply-held principles and beliefs.

Elsewhere, amongst followers of Wittgenstein, it is argued that the problem will simply disappear (like all such ‘metaphysical’ quandaries) if one sees that there are two different language-games involved, those of reason-based justification on the one hand and causal explanation on the other.<sup>3</sup> Or again, it can be conjured away through a ‘naturalised’ (or detranscendentalised) reading of Kant which recommends that we jettison a great deal of his outmoded metaphysical machinery but retain the basic distinction between a physical realm where causal explanations are perfectly in order and a ‘space of reasons’ where the pertinent criteria are those of valid inference, well-formed argument, or justificatory warrant.<sup>4</sup> However, these strategies are no more effective in resolving the central issue – to put it bluntly, no less of a philosophic cop-out – than the idea that both sorts of talk make sense on their own terms and therefore cannot possibly get into conflict just so long as we regard them as belonging to disparate language-games or modes of thought. For this recourse to language as a means of escape from all our philosophic perplexities is one that leaves the conceptual problem firmly in place, amounting as it does to a placid assurance that ‘everything is in order’ with our accustomed linguistic practices. Yet the free will/determinism issue is just as pressing or worrisome when couched in everyday, non-specialist terms as when subject to a full-scale ‘metaphysical’ treatment in the Kantian manner. Quite simply, such problems cannot be wished away by any amount of linguistic therapy or Wittgensteinian attempts to persuade us that they are really just a form of self-induced philosophic bewitchment. Besides, these approaches are always at risk of implicitly espousing a cultural-determinist view – and thus belying their professions of even-handedness – in so far as they entail the idea that beliefs are intelligible only within some language-game or communal ‘form of life’. For then it follows that the freedom to question or to challenge doxastic norms must *ipso facto* be limited to whatever makes sense by those same communal lights.

Outside the analytic line of descent these issues have received some very different kinds of treatment. Among them is Sartre’s existentialist idea of human consciousness as the locus of an ultimate, unqualified freedom (a ‘hole’ in being or a region of absolute ‘nothingness’) which marks the sole point of interruption in a physical universe otherwise governed by the iron laws of physical-causal determinism.<sup>5</sup> There is much of great interest and value in Sartre’s approach to these topics, as likewise in the work of those – like Merleau-Ponty – who have sought to offer a viable account of human moral autonomy while criticising Sartre for his all-or-nothing view and insisting that we recognise the practical constraints on our freedom in any given real-world context.<sup>6</sup> One strength of their analyses is the fact that they provide a good range of real or imagined test-case scenarios in order to flesh out the issues and bring us flat up against the kinds of dilemma confronted not only by human agents in various sorts of complex moral predicament but also by philosophers who seek to engage – and not evade – their more intractable aspects. Still the basic problem re-emerges very sharply in Sartre’s later acknowledgement that if the idea of freedom is to have any genuine, as opposed to merely notional content then it will need to be specified in terms that take adequate account of those numerous factors (whether physical, historical, socio-cultural, psycho-biographical, or whatever) that in practice must be seen as placing certain limits on our scope for moral autonomy.<sup>7</sup>

This allowance becomes more explicit in his politically engaged writings where it is a chief premise of Sartre’s Marxist-dialectical approach that human beings make their

own history, but not in circumstances or under conditions of their own choosing. However it is also present in his existentialist works to the extent that freedom is here thought of as exercised in a context – that of our relationship to other people under certain, often highly fraught circumstances – which itself requires allowance for just such constraints, though here of a more inter-personal than large-scale collective, classbased, or group-dynamic kind. Indeed the very notion – so crucial to Sartre’s early existentialist thinking – that freedom always entails responsibility even when manifested in selfish, morally or socially irresponsible ways is one that must likewise impose significant restrictions in that regard.

And the same applies to the issue of doxastic voluntarism, free as we are (in principle at least) to adopt any number of possible beliefs on any given topic yet constrained as we are (by factors ranging from passive indoctrination to rational conviction on the basis of empirical evidence or cogent demonstrative grounds) to believe what we do as a matter of involuntary assent. Thus the problems with Sartre’s existentialist ethic and his notion of absolute, unqualified freedom can be seen to work out as a close analogue – albeit more dramatically expressed – of those which have constantly resurfaced in the ‘other’, i.e., Anglophone or mainstream analytic tradition.

If one thing is clear, to repeat, it is that these problems cannot be resolved by any kind of linguistic therapy aimed toward talking us down from the heights of ‘metaphysical’ abstraction and leading us back to a sensible acceptance of the different language-games involved. This Wittgensteinian approach is really just a line of least resistance or a strategy adopted in order to evade what is surely among the most pressing issues in philosophy of mind, epistemology, and ethics. That is to say, it avoids facing up to the fact that we are inevitably tugged both ways between the powerful conviction that people *can and should* be held responsible for their beliefs and the knowledge – just as much a part of our basic moral-evaluative competence – that such responsibility is never exercised in total isolation from the various causes, influences, pressures of circumstance, or acculturated modes of thinking that predispose us toward one or another doxastic commitment. In the face of this dilemma it is tempting to adopt a standpoint analogous to that of some philosophers who have ‘answered’ the problem about consciousness by suggesting that it is just too difficult (or too far beyond our innate powers of conceptual grasp) to allow of any adequate scientific or indeed philosophical solution.<sup>8</sup> However, as opponents have been quick to remark, this ‘mysterian’ argument has nothing to commend it except the appeal – perhaps more evident to philosophers than scientists – of preserving a space where causal explanations necessarily run out and philosophy continues to set the rules for what counts as a valid or admissible hypothesis.<sup>9</sup> Besides, there is something intellectually disreputable about any theory that declares straight off as a matter of stipulative fiat that phenomenon x – whether quantum mechanics or human consciousness – is something that by its very nature exceeds our utmost powers of conceptual-explanatory grasp.

Such arguments are not so very different from those items of orthodox Christian doctrine that J.S. Mill denounces with splendid moral vigour when he states his refusal to believe in any God whose ways are so profoundly mysterious (for instance, in the matter of eternal damnation) as to lie beyond reach of mere human understanding.<sup>10</sup> To raise the problem of doxastic voluntarism to a high point of philosophic bafflement is in effect to concede the irrelevance or downright

uselessness of philosophy when it comes a matter of the greatest importance for our conceptions of moral, political, legal, and intellectual responsibility.

## II

So where have we arrived at this preliminary stage in our discussion? Not, to be sure, on the verge of suggesting any confident answer or adequate solution. Indeed it might seem that the prospect has receded even further as a result of having canvassed some views on the topic – some more-or-less qualified defences of doxastic voluntarism and the contrary (non-volitionist) stance – and having found them open to various kinds of philosophically cogent or intuitively powerful rejoinder. Now one is faced with a four-way choice between possible ways of proceeding. First there is the option that would most likely be favoured by hard-line ‘analytic’ types who see no virtue in continuing to fret over age-old metaphysical issues whose lack of any widely agreed-upon solution shows that they were ill-formed to begin with and had therefore better be shelved or re-cast in some suitably modified, e.g., linguistic or logico-semantic form. To this way of thinking the antinomies thrown up by debates about doxastic voluntarism – like those that have bedevilled the freewill/ determinism issue in philosophy of action – are such as to call for treatment in the mode of *reductio ad absurdum*, that is, as showing that they must derive from some one or more false (since deadlock-producing) premises.

However this ‘solution’ is rather like Russell’s famous but philosophically unconvincing Theory of Types which advised that we could best get over the paradoxes of classical set-theory – those of self-reference or self-inclusion – by laying it down as a rule of good mathematical and logical conduct that systems be constructed and distinctions maintained in such a way that those paradoxes simply could not arise. Still they are apt to strike home with undiminished force for anyone who tries to get her mind around Russell’s various examples of the kind, just as the problem about doxastic voluntarism won’t go away – or show up as a mere pseudo-problem engendered by false ‘metaphysical’ premises – simply through flat declarations to that effect. Nor can it be any more effectively dissolved through the second, Wittgensteinian option according to which it is only on account of our (i.e., philosophers’) chronic ‘bewitchment by language’ that these dilemmas have come to exert such a hold. For here again no amount of linguistic therapy – of patiently coaxing those philosophers down from the giddy heights of metaphysical abstraction – can be of much use when we come to reflect on the real, not illusory or language-induced problem of reconciling moral-intellectual responsibility with a due allowance for the various constraints on our own and other people’s modes of belief-formation.

The third option takes a lead from Kant in pressing those antinomies not to the point of a self-refuting *reductio* but rather to the point where they are taken to entail a very different way of conceiving the issue.<sup>11</sup> This confines our knowledge of causality to the realm of phenomenal (perceptual) experience, conceptual understanding, and physical science while conserving a strictly separate domain – or ‘space of reasons’ – for the exercise of ethical choice under no compulsion save that of the requirement to respect the universal dictates of moral law. Yet there is an obvious problem here in so far as compliance with that law is supposed to be a matter *both* of freely-willed, autonomous assent *and* of something more like a passive acquiescence in maxims or principles that brook no exception and would hence seem to leave no room for such

moral autonomy. Thus the Kantian ‘solution’ turns out to be just another variant on the same old dilemma, one that is by no means resolved – rather sharpened – by those revisionist (naturalised or ‘de-transcendentalised’) readings of Kant that have lately emerged by way of response to the defects and anomalies of old-style logical empiricism.<sup>12</sup> For such readings still have to face the choice of either reproducing that absolute Kantian distinction of realms, in which case they will hardly be rid of transcendental motifs, or else pushing right through with the naturalistic treatment of Kantian epistemology and ethics, in which case they will produce a version of Kant which cuts out some crucial load-bearing segments of his argument. Among these latter – most damagingly for the revisionist case – are just those passages that claim to establish the possibility of free-will and moral autonomy in a world that is otherwise subject to causal determinism in every last detail of every last event. Kant’s legacy is plainly visible across a wide range of latter-day responses to this problem, from John McDowell’s halfway naturalised and (in my view) deeply problematical version of Kantian epistemology to Donald Davidson’s idea of ‘anomalous monism’, itself – despite the somewhat misleading description – a dualist doctrine that dare not quite speak its name.<sup>13</sup> So there seems little hope of an answer from purported solutions of this third (Kantian or quasi-Kantian) kind, even though they are more responsive to the philosophic depth of the problem and the sheer unlikelihood that it might be laid to rest through some straightforward logical *reductio* or application of linguistic therapy in the Wittgensteinian mode.

So the question remains: what hope of an answer (a fourth-way alternative) if the best efforts of philosophy to date have produced nothing more than a series of dilemmas, deadlocked antinomies, conflicting intuitions, and conceptual dead-ends? Still things may not be as hopeless as this suggests if one just hangs onto the basic principle that whatever else philosophy may claim to do it cannot depart from certain indubitable axioms of human thought and experience. These are axioms – not just convenient working hypotheses – in so far as they serve both as a starting-point for further, more detailed and rigorous reflection and also as a check on the tendency to press toward doctrinal extremes (such as wholesale determinism or radical voluntarism) that are plainly at odds with much of what we know as a matter of self-understanding and shared experience. This tells us that Plato must have been wrong – in the grip of a false theory of mind and ethical motivation – when he argued that knowledge of the good must infallibly lead to virtuous behaviour, or that bad actions could result only from ignorance, stupidity, or misunderstanding. By the same token we are inclined to agree with Aristotle when he makes allowance for *akrasia* (weakness of will) as a complicating factor that often intervenes to prevent our following the straightforward dictates of duty or moral conscience.

So likewise with the standard objections to Kantian deontological ethics, i.e., that such thinking both under-estimates the complexity of human predicaments and motives, and over-estimates the binding character of any such abstract-universalist moral creed. Yet at the same time – and here the familiar dilemma crops up once again – we are compelled to acknowledge that there *must* be some basis for ethical values beyond this potentially all-licensing appeal to the range of qualifying clauses required for any statement concerning the scope and limits of moral responsibility. Thus it is one thing to argue that Plato and Kant got it wrong – albeit on different metaphysical grounds – when they took such a sternly uncompromising line as regards the absolute status of moral truths and the requirement that ethical reasoning

not be deflected by merely 'pathological' considerations of this sort. Yet it is another thing entirely to push so far in the opposite direction that one is left with the idea of ethical judgement as involving nothing more than a shared (even if community-wide) consensus as to what should count – in any given case – as a reasonable, decent, or morally acceptable view of the matter. In striving so hard to avoid all the problems with Kantian moral rigorism this approach runs the risk of becoming just another variant on a well-worn cultural-relativist theme, or confusing what is true, justified, or right in the way of belief with what passes as such according to our own communal practices and values. For then we are stuck with another form of determinism that is no less degrading to our basic conceptions of moral autonomy and selfhood in so far as it involves a socio-cultural rather than a causal or physicalist theory of belief-formation.

Still there is no denying the force of our conviction that we *do* have a significant measure of choice in the matter of what we believe and, moreover, that such choice is not drastically compromised or shown up as just a kind of willing self-delusion by the fact that our opting for one or another doxastic commitment can often quite plausibly be traced back to some prior influence, whatever its precise nature. After all, there is a vast (non-denumerable) range of such variously weighted influences that impinge at any moment on any individual in any given life-context and it is absurd to suppose that, even (*per impossibile*) with all the evidence to hand, one could ever predict the future course of that person's actions and beliefs. Yet as a putative solution to the free-will/determinism problem this fares no better – philosophically speaking – than the standard response to Laplace's claim that from a complete knowledge of the present state of the universe right down to its ultimate physical constituents one could in principle retrodict its entire previous history and likewise predict its entire future development.

That 'solution' consists in saying quite simply that we don't possess and could never attain such an ultimate state of knowledge, and therefore that Laplace's determinist claim is beside the point for all practical as well as genuine philosophic purposes. However this just won't do as an answer – least of all a philosophical answer – since it fails to take the crucial point that determinism might conceivably be true (as a matter of fact) quite aside from any merely contingent limits on our powers of comprehension or ability to figure out the whole, endlessly complex concatenation of causes and effects. Nor is there much comfort to be had for the hard pressed anti-determinist from the idea that science has now moved on to a stage where such claims no longer present any threat in so far as they have been superseded by developments like chaos-theory, mathematical undecidability, or (in quantum-physical terms) the uncertainty relations and limits on our powers of precise, objective measurement. For such arguments are open to the threefold charge of (1) confusing ontological with epistemological issues, (2) presupposing the truth of certain highly questionable (e.g., quantum-theoretical) conjectures, and (3) trivialising the whole debate by making it hinge on the outcome of random events – say quantum goings-on the brain – that would surely do nothing to explain our capacities for rational belief-formation or autonomous action.<sup>14</sup> For there seems little point in advancing this sort of case if the only result is to exchange one philosophically unpalatable view (hard-line psycho-physical determinism) for another, equally unwelcome idea (that rationality and free will are just illusions engendered by our post-hoc attempts to make sense of such sheerly random, unmotivated goings-on).

### III

The reader will perhaps have noticed a recurrent pattern in the last few paragraphs, namely the way that they each start out with a statement of the need to move beyond these vexing antinomies, only to end with a reformulation of the same basic problem in different terms. My excuse is that the problem is a tough one – among the most philosophically recalcitrant – and that any attempted solution is obliged in good conscience to register the various counterarguments or likely objections that rise up against it at every turn. There is a curious example of this in Benjamin Libet's much-discussed findings with regard to the neurophysiology of decision-making and the temporal relationship or order of dependence between brain-states and conscious mind-states.<sup>15</sup> Most controversial was the fact – as he at first claimed – that neural-imaging experiments had shown a distinct, measurable time-lag between the occurrence of chemical events in the brain that correlated with certain apparently willed or deliberate actions and the moment when subjects reported their decision to execute those same actions. From this it seemed to follow that their 'choices' of action were really no such thing but, on the contrary, epiphenomenal events that occurred only *after* the brain had entered into a certain state whose results were first manifest in overt behaviour and then became accessible to consciousness only as a kind of passive delayed effect. However Libet subsequently modified his claim by allowing that any act 'determined' by any given brain-state could always somehow be revoked or countermanded by a further, split-second intervening decision not to proceed in that particular way but to fix on some alternative outcome. Of course there is still the option, for diehard determinists, of arguing that Libet's revised claim complicates but doesn't in the least undermine his original thesis, i.e., that it is a change in brainstate and not some immaterial change of purpose or mind-set that produces the change. However this argument is open to various further objections from the voluntarist quarter. Among them is the point that it leads to a form of vicious regress – since the countermanding impulse may itself be subject to further countermanding, and so forth *ad infinitum* – and also that this process cannot be described without at some stage having recourse to an intentionalist idiom, one that involves the ascription of motivating interests, desires, or beliefs.

So there is an odd but revealing and symptomatic sense in which Libet's retreat from the first (strong-determinist) version of his thesis re-enacts precisely that complicating moment – that same split-second intervention of a contrary, action-inhibiting force – which he now incorporates, no doubt on the basis of certain empirical observations, but also (one suspects) partly in deference to our standing intuitions in that regard. For it is a demonstrable feature of all arguments in philosophy of mind, cognitive psychology, and other areas where this issue arises that any statement of the case for hard-line physical determinism will at some point involve a more-or-less covert or surreptitious appeal to the language of agency, volition, and choice. Nor is this merely, as Wittgensteinians would have it, a matter of our using different language-games in different contexts of utterance, e.g., that of causal explanation (including empirical psychology) on the one hand and reason-based, normative, or justificatory talk on the other.<sup>16</sup> Rather it is a question of our total inability to make any sense of human actions and beliefs unless by accepting – whether explicitly or (as very often in the case of determinist arguments) through various tell-tale nuances of word and

phrase – those basic voluntarist premises which between them constitute *just what it is* to understand our own as well as other people's motivating interests and concerns.

Chief among them, as I have said, and absolutely central to the issue about doxastic responsibility is our understanding that beliefs can be arrived at in a great many ways, some of which render them fit candidates for assessment in rational and moral-evaluative terms while others seem to call for explanatory treatment in an altogether different, more diagnostic mode. Of the latter kind are beliefs acquired (or absorbed) through passive exposure to various modes of causal, cultural, or ideological conditioning. In the former case we have to do with those other, more reflective or adequately reasoned forms of doxastic commitment that have taken due account of such factors, allowed for their possible distorting effect, and thereby achieved both a greater degree of self-knowledge and an outlook in matters of moral or intellectual conscience that is more reliable or truth-conducive since less in the grip of unexamined prejudice.

Of course these descriptions apply only to the twin extremes on a scale of doxastic responsibility that includes a great many intermediate degrees, that is say, cases where it is hard – maybe impossible – to distinguish the workings of causal influence from those of rational belief-formation, whether for the individual concerned or for those who seek to make intelligible sense of that individual's sayings and doings. Also there is a marked asymmetry between our readiness to fall back on causal hypotheses, i.e., imputations of rationally under-motivated thought and action in the case of other people and our much higher level of resistance to any such suggestion as regards our own most cherished or deep-laid beliefs. After all, there is something basically absurd – a kind of performative contradiction – about saying 'I believe *x* to be the case but this belief of mine is most likely a product of my upbringing, cultural background, ideological formation', or whatever. Thus causal explanations of why people think and behave as they do are much likelier to carry conviction when applied to others (especially to people remote from us in time, geographical locale, or socio-cultural background) than when applied to ourselves and those nearest to us in these respects. Nevertheless that resistance has been quite appreciably lowered, at least amongst the more educated sectors of society, as a result of various diffuse yet powerful influences such as psychoanalysis, comparative anthropology, and the whole range of present-day social sciences that have stressed the sheer multiplicity of human beliefs and value-systems.

Even if one wishes to hold the line well short of wholesale cultural relativism – including its more philosophically 'respectable', e.g., Wittgensteinian variants – still there is no denying the extent to which developments like these have altered (and expanded) our sense of what may count as a rational, or at any rate rationally intelligible item of belief. Besides, it is only from the most dogmatic of voluntarist viewpoints that philosophy could afford simply to disregard the latest findings of neurophysiology, cognitive science, evolutionary psychology, and other disciplines with an arguable bearing on the issues here addressed.

What is called for, in short, is an approach that manages to take these developments on board whilst not leaning over too far in one or the other direction. That is to say, it should avoid the kind of inertly 'culturalist' perspective wherein all beliefs and practices are deemed to make sense by their own communal lights, in which case they



must be held immune from any form of ‘outside’ criticism, but also the opposite error of supposing that standards of doxastic responsibility can be applied without regard to the cultural conditions under which those beliefs and practices emerged. In so far as it is difficult to hold this balance – to weigh the strong claims of moral-intellectual autonomy against the need for a thoughtful and sensitive reckoning with such heteronomous conditions – the difficulty is one that will surely arise in all attempts to resolve the free-will/determinism issue. So we might now seem to be back with the same old dilemma and to have shown nothing more than the unfortunate proneness of philosophy to constantly rehearse familiar problems in a range of alternative but equally futile (since non-problem-solving) ways. However this is to take an excessively sceptical or pessimistic view of the matter. What the arguments and counter-arguments bring out is not so much the ultimate impasse engendered by two conflicting (causal-explanatory and rational-justificatory) modes of thought but rather the conclusion that there *must* be some way of reconciling them – albeit one at present beyond our best powers of conceptual grasp – since they both play a strictly indispensable role in all our thinking about issues of doxastic responsibility. What should also have emerged from this discussion is the impossibility of pushing either case (i.e., for the volitional or non-volitional character of belief) to a point where it would play the other clean off the field by establishing its own irresistible claim to have finally resolved those issues.

Thus it might appear that any hope of progress must lie in the direction of a *via media* between the two extremes, that is, an approach that sensibly acknowledges the limits on our freedom of will as well as the problems that always arise with any full-fledged determinist, causal-explanatory, or anti-voluntarist account. Where this latter goes wrong is in failing to allow for all the evidence we have – not only through intuitive, first-person experience but also in the contexts of social exchange and reflection on other people’s acts and beliefs – that there does exist a margin for the free exercise of moral and intellectual conscience whatever the extent of those causal or more broadly socio-cultural constraints. Any theory is sure to be a non-starter if it has to discount the example of various heterodox thinkers or reformers whose moral-intellectual conscience has revolted against the kinds of taken-for-granted belief that defined the currency of ‘knowledge’ or ‘truth’ in their own time and place.

It is on this objection that determinist arguments must ultimately run aground, as indeed must those other, on the face of it strongly anti-determinist arguments – among them Wittgensteinian appeals to communal ‘agreement in judgement’ as the furthest one can get by way of rational or moral justification – which likewise leave no room for the values of intellectual and moral autonomy. However these antinomies cannot be resolved (as I suggested above, though prefaced by the cautionary phrase ‘. . . it might appear . . .’) through some kind of compromise or middle-ground approach that would seek to prevent them from arising in the first place by adopting a suitably qualified conception of doxastic voluntarism on the one hand and of doxastic determinism on the other. No doubt it is often best policy in various contexts – law, practical ethics, interpersonal relations, ascriptions of motive to culturally distant (including historical) agents – to work on a generalised ‘principle of humanity’ which does involve some such compromise stance as a matter of trying to see all around other people’s epistemic situations or moral predicaments and not jump straight to a dismissive or critical verdict. That is, we often make this sort of good-willed effort to maximise the imputed rational content of other people’s beliefs by achieving an

optimum balance between the claims of autonomy and those of due allowance for certain causally or socio-culturally explicable sources of error.

Yet philosophically speaking – as Kant recognised, unlike some who purport to have left all those old worries behind – there is a large problem here and one that can only be addressed in metaphysical terms, or anyway in terms that admit of no such straightforward compromise solution. For it remains the case – whether a misfortune or a stroke of good luck from the philosopher’s standpoint – that these are *antinomies* in the strict Kantian sense, rather than paralogisms (again as defined by Kant) that can be shown to result from some category-mistake or illicit transposition of concepts and categories from one to another topic-domain.<sup>17</sup> In other words they are the kinds of genuine, deep-laid problem that that are sure to confront any thinker who seriously engages with the issue of doxastic responsibility.

Not that I should wish to hold Kant up as the likeliest source of deliverance from all our perplexities in this regard. Indeed, it is very largely as a consequence of the various drastic dualisms bequeathed by Kant that philosophy has so often tended to recoil into likewise drastic oscillations from one to another doctrinal extreme on a range of epistemological and ethical issues. Among them are those between the noumenal and the phenomenal, between concepts of understanding and sensuous intuitions, and (most of all) between the realm of autonomous practical reason and that other realm of ‘heteronomous’ desires, inclinations, affections, and suchlike ‘pathological’ factors which for Kant have no legitimate bearing whatsoever on issues of moral conscience, especially when they are adduced in order to extenuate or mitigate some wrongful action that would otherwise call for outright moral censure. So if Kant poses the relevant issues with a clarity and force that are often lacking in current debate, he does so in a way that places them forever and intrinsically beyond hope of any adequate solution. This is because he conceives the ‘space of reasons’ – of our distinctively human intellectual capacities, as opposed to our mere creaturely existence within the chain of concatenated cause and effect – in those same, drastically dichotomous terms. It is a very fine distinction I am trying to draw but one that makes all the difference between an outlook which counts this dilemma as *by its very nature* unresolvable and an outlook on which it remains deeply problematic by our present-best conceptual lights but not, for that reason, to be treated as an ultimate mystery or as requiring the kind of Kantian dualist approach that has found many sublimated echoes in recent debate.

Least of all can it be justified to adopt the kind of sheerly *a priori* approach that would view any attempt to achieve a perspective atop these vexing antinomies as symptomatic of a basic failure to grasp the irreducibility of mental to physical or – in the currently preferred idiom – intentional to causal modes of describing our beliefs, attitudes, doxastic commitments, moral dispositions, and so forth. Such arguments typically issue in dogmatic claims like that of John Searle with respect to what he sees as the absolute, in-principle impossibility that the ‘strong’ programme in Artificial Intelligence might ever be successfully carried through, i.e., to the point where some silicon-based and computer-run system might be thought to manifest all the attributes of human consciousness or intentionality.<sup>18</sup>

The main problems with this kind of reasoning are firstly its neo-Cartesian assumption that there *must* be something so utterly distinctive about human mental

states that they could not conceivably be realised in any other physical form, and secondly its downright refusal to acknowledge that the brain *just is* one such physically embodied, albeit massively complex and – in our current state of scientific understanding – causally inexplicable entity. To be sure, some philosophers may choose to grasp this nettle (whether in its downright substance-dualist or updated property-dualist form) and thereby reject or at least draw the sting of the mind/brain identity thesis. Otherwise they may resort to some saving compromise like Davidson's 'anomalous monism' or the notion of mind as 'emergent from' ('supervenient upon') the physical brain-states that are found to accompany this or that item of conscious or affective experience.<sup>19</sup> However such arguments amount to no more than a handy but somewhat shifty device for avoiding the central issue about consciousness and its relation to whatever is going on in neuro-chemical terms.

Hence the ease with which opponents of Searle – hard-line physicalists like the Churchlands – can turn his case right around and object to his saying that mental states are somehow 'caused by' brain-states (that is, the sorts of state that can only exist in carbon-based, organic life-forms such as ourselves), rather than saying that mental states just are brain-states under a different, folk-psychological description.<sup>20</sup> Dragging in such redundant causal talk is an example of what they dub the 'Betty Crocker Microwave Cookbook' fallacy. This alludes to a passage where the domestic guru explains that heat is *caused by* the kinetic energy of molecules, rather than saying (correctly) that 'heat' *just is* everyday, phenomenological parlance for what physicists term 'kinetic energy of molecules'. Thus Searle's apparent concession to physicalism – his allowance that there is indeed a strict and exceptionless causal correlation between brain-states and mind-states – can none the less be seen to underwrite his claim (*contra* the advocates of strong AI) that this link exists only in creatures like ourselves with the right kind of neurophysical architecture and just the sorts of conscious or intentional experience that invariably go along with it. Once rid of this residual dualism – so the Churchlands maintain – we shall see that there is nothing unique or *sui generis* about mind, brain, or the relationship between them. Rather we shall come to treat such mentalist talk as just another item of soon-to-be-discarded since scientifically retrograde belief, along with all the other metaphysical baggage handed down by a long tradition of jointly philosophic and folk-psychological thought.

My point is not so much to take sides on this issue of Searle *versus* the Churchlands but rather to bring out the irony of a situation where Searle's hard-headed causal talk – his overt refusal of Cartesian dualism and insistence on the physical embodiment (i.e., the brain-dependence) of consciousness and intentionality – can be used to charge him with falling into precisely such a dualist trap. It seems to me that Searle is here hung up on yet another of those Kantian antinomies that have typified this whole debate and whose effect is to leave thinkers very often exposed to criticisms and objections which are all the more powerful since arising unnoticed from their own arguments. Thus Searle's attempt to carve out a space for the distinctive attributes of human (i.e., conscious, intentional and organically based) experience while at the same time rebutting any dualist charge is one that leads him to adopt a position – the mind-brain causal dependence thesis – which the Churchlands can treat (not without justification) as a form of epiphenomenalist doctrine that, so far from resolving the Cartesian dilemma, leaves it all the more firmly entrenched. For it then becomes a double mystery (1) by what remotely intelligible process mind-states could 'emerge

from' or 'supervene upon' (let alone be 'caused by') physical states of the brain, and (2), if so, how there could be any genuine – rather than notional – appeal to a realm of irreducibly conscious or intentional experience that would constitute a standing refutation of reductive physicalism in the Churchland mode.

This is why Searle's purportedly knockdown case against strong AI – a case advanced mainly on thought-experimental and hence *a priori* grounds – cannot bear anything like the requisite weight of demonstrative evidence or proof. On the one hand it runs into all the above-described conceptual and logical problems, while on the other it invites the charge of dogmatically denying what must surely be a matter for continued scientific investigation, that is, the possibility that conscious and intentional mind-states might eventually be realised in other-than-human (e.g., silicon-based) systems. Thus, according to Searle's famous 'Chinese Room' thought-experiment, this prospect is *a priori* ruled out by the fact that we could never know for sure whether the English-Chinese 'translator' inside the room (for which read: the 'conscious and intelligent' AI device) was actually translating the messages handed in or merely responding in mechanical fashion by comparing the word-shapes and sequences with those contained in his data-bank (for which read: a software programme set up to give the impression of thinking 'like us' but in fact quite devoid of any such powers).<sup>21</sup>

However, despite its intuitive force, this argument falls to the twofold objection that it holds just as much for our transaction with human subjects – as witness the perennial issue of scepticism *vis-à-vis* 'other minds' – and that it flies in the face of Searle's own argument that mind-states are causally dependent on brain-states. For if this is the case and if brains are themselves (no doubt fantastically complex) computational devices then there is just no rational motivation for Searle's claim that mind-states of a nature qualitatively identical to those experienced by human beings cannot conceivably be produced or supported by different kinds of physical system. Besides, there is something decidedly premature – given the current rate of advance in allied fields like neurophysiology and cognitive science – about any argument that claims to deduce the impossibility of further such advances on a basis of purely *a priori* reasoning and with minimal reference to what's going on in just those pertinent fields.

#### IV

As I have said, this should not for one moment be taken to suggest that philosophers had better now vacate the high ground of expert debate in these matters and give way to others (the neuroscientists and cognitive psychologists) who really know what they are talking about. If Searle's style of argument exemplifies the dangers of an attitude that grants philosophy the right to decide what shall count as relevant, admissible scientific evidence then it is equally the case that proponents of an out-and-out physicalist approach, like the Churchlands, push so far in the opposite direction as to lay themselves open to the charge of just ignoring – rather than genuinely seeking to resolve – the very real philosophic problems that arise with any such programme. Thus they reject as just a remnant of 'folk-psychology' the idea that there is a whole dimension of subjective experience – 'what it is like' to see the colour red, to suffer pain, to hear an oboe, to undergo the gamut of feelings from elation to despair – which cannot and could not be explained in physicalist terms, even were we to possess a completed science of the brain and its neuro-chemical or cognitive-psychological workings.<sup>22</sup> This is not the place for a detailed discussion of the various

arguments for and against the existence of such strictly irreducible qualia or modes of first-person perceptual, affective, or phenomenological experience which supposedly elude any possible description in the terms of a drastically reductionist approach like that espoused by the Churchlands. Sufficient to say, in the present context, that this debate is just as far from any prospect of being effectively resolved either way as are the closely related issues of free-will *versus* determinism or – my chief concern here – doxastic voluntarism *versus* the claim that beliefs are not volitional and hence not subject to moral blame or approbation. What gives them their distinctly philosophic character is the fact that they produce such deeply held yet sharply conflicting intuitions which dispose us to believe that they *must* be somehow resolvable, while their effect is to render the issue more perplexing and its solution more elusive the harder we think about it.

This is not – as it might well appear – just a recipe for endless equivocation or a pretext for philosophers to sit back and contemplate an ever-deepening (and action-absolving) series of conceptual quandaries. On the contrary: the main purpose of raising such issues is to keep them constantly and clearly in view when we are tempted to adopt some other way of thinking that involves less challenge to our normal, unexamined, or communally warranted habits of belief. Thus, for instance, the case for regarding our beliefs as not (or not entirely) subject to our powers of conscious, deliberative will is one that has been made – and that still needs making – at times of rampant doctrinal, religious, or political persecution. In this respect, though not in others, it may be seen as the doxastic equivalent of the moral or legal case for treating certain agents as not fully accountable for certain actions, whether in consequence of social factors, intellectual impairment, mitigating circumstance, or a whole range of causal (among them psychopathological) conditions that are felt to justify the plea of ‘diminished responsibility’.

Of course the big difference is that here we are dealing with various kinds and degrees of unfreedom, that is, of restriction on the subject’s scope for intellectual and moral autonomy, whereas in the former case – paradoxically enough – freedom of belief is upheld as a matter of socio-political right on the grounds that people are often to some extent *not* responsible for what they believe, and should hence not be subject to penalty or blame on that account. Indeed one can see the paradox emerging at full force in those two occurrences of the word ‘subject’ in my previous sentence, first with the active-autonomist sense: ‘subject (noun) = locus of freely-willed choice and rational accountability’, and second with the passive-determinist sense: ‘subject (adj.) = under some constraint or compulsion that places limits on the exercise of just those powers’. This is not the sort of problem that might be cleared up by a more precise definition of terms, nor by application of the standard Wittgensteinian therapeutic treatment. Rather, it is one that is apt to strike any thinking individual who considers the arguments on both sides not only from a philosophic standpoint but also in the wider context of debates about law, ethics, and the proper limits – if such there be – on the conscientious freedom to express ideas that go against some existing currency of values and beliefs. For these are questions that are nowhere close to being resolved and which might well be thought – at any rate by those of a sceptical, defeatist, or ‘mysterian’ mind – as lying beyond reach of any adequate solution.

While the latter view amounts to a strategy of last resort it is none the less important to remind ourselves and others of how complex and intractable these questions remain

and also how we are prone to fall into errors of judgement – even, as I have argued, miscarriages of justice – by opting too readily for one or the other standpoint. After all, it is among the chief functions of a properly trained philosophical intelligence to supply such reminders when required, as they often are in situations (like those instanced above) where social and political incentives or pressures may override the call for such fine discriminations in the matter of moral and doxastic responsibility. This is one area in which analytic philosophers have something to learn from their ‘continental’ (i.e., post-Kantian mainland-European) counterparts, whatever the degree of mutual mistrust and downright intellectual antagonism that has marked many of their dealings over the past half-century and more.<sup>23</sup> That is to say, this sense of irreducible complexity – of the objections that arise against any too quick or confident solution – is a prominent feature of much mainland-European philosophy, even (or especially) those aspects of it that analytic thinkers are apt to write off as extravagant, linguistically undisciplined, or overly speculative. I have made the point elsewhere with reference to recent interpretations of Kant and the contrast between, on the one hand, broadly ‘continental’ readings that engage deeply with problems with Kantian epistemology and ethics and, on the other hand, revisionist readings in the broadly analytic tradition that contrive to keep such complicating factors safely out of view.<sup>24</sup>

Of course this is not to suggest that the ‘linguistic turn’ in its various forms has rendered mainstream Anglophone philosophy devoid of resources for addressing such topics. One need only look to a thinker like J.L. Austin, in particular his essay ‘A Plea for Excuses’, if one wishes to find a striking example of the way that ‘ordinary-language’ analysis can illuminate questions of moral judgement or offer new ways of conceiving the free-will/determinism issue through a patient and meticulous attentiveness to nuances of verbal implication.<sup>25</sup> This is nowhere more apparent than in Austin’s reflections on the range of finely-tuned adverbial modifiers – ‘he did it “wilfully”, “deliberately”, “knowingly”, “consciously”, “intentionally”, “absent-mindedly”, “inadvertently”, “carelessly”, “accidentally”, “unwittingly”, etc. – by which we signal our intuitive grasp of just such nuances and use them to indicate varying degrees of moral culpability. All the same his critics do have a point when they suggest that there is something distinctly parochial about Austin’s brand of ‘ordinary language’ philosophy, an odd mixture of Oxonian *hauteur* and complacent acquiescence in the habits of thought – the linguistically encoded mores – of his own cultural locale.<sup>26</sup> Indeed, one result of the linguistic turn in analytic philosophy since the 1950s – and arguably since the Moore/Russell revolt against ‘idealist’ or ‘metaphysical’ excesses of whatever kind – has been to rule out any deeper engagement with such issues except in so far as they are taken to involve some category-mistake, some conceptual error, or (after Wittgenstein) some symptomatic instance of the ‘bewitchment of our intelligence by language’ For if philosophy is best, most usefully (or least harmfully) employed in clearing away or therapeutically dissolving those old – e.g., Kantian – dilemmas then the free-will issue itself becomes just another suitable case for treatment. And yet, as I have said, it is an issue that cannot be finessed by any amount of conceptual analysis or any number of placid assurances that ‘everything is order’ with our language as it stands and that philosophers must therefore be mistaken – in the grip of some (again typically Kantian) metaphysical or transcendental illusion – if they seek to raise problems where no such problems exist.

If Kant famously credited Hume with having shaken him out of his dogmatic slumbers and thereby set him on the path toward a full-scale critical reconstruction of epistemology and ethics then at present what is needed, or so I would suggest, is a similar revitalising impulse in the opposite direction. Thus analytic philosophy might very well benefit from a willingness to abandon its defensive posture and take some account of those developments in ‘continental’ thought – from Husserlian phenomenology, *via* Sartrean existentialism, to Derridean deconstruction – that have kept alive certain crucial questions (or provocative ways of framing them) beyond what counts as proper or legitimate by its own self-assured criteria. If there is one characteristic that chiefly distinguishes the ‘two traditions’ it is this greater awareness, on the continental side, of the need to pursue problematical issues – like the Kantian antinomies – to a point where they engage the genuine dilemmas of human existence, rather than supposing that these can best be kept from causing trouble through a mode of conceptual or linguistic analysis that effectively sweeps them under the carpet. Perhaps the most striking example of this latter tendency is the work of Gilbert Ryle where it is pretty much assumed that any problems of so seemingly intractable or deep-laid a character must, for that very reason, be put down to some ‘category-mistake’ or failure to perceive where thinking has been misled by its proneness to various forms of conceptual imprecision or false analogy.<sup>27</sup> Ryle is a particularly interesting case since he published a number of critical yet well-informed essays on Husserlian phenomenology during the 1930s, but later swung across to the received (analytic) view that all the talk of ‘intentionality’, ‘eidetic essences’, the ‘transcendental ego’, and so forth, was in truth just a thinly disguised version of psychologism.<sup>28</sup> Another tantalising hint of this path not taken – or abandoned after a brief reconnoitre of the alternative prospects on offer – is Austin’s passing remark to the effect that his kind of ordinary-language approach could also fairly be described as a form of ‘linguistic phenomenology’.<sup>29</sup>

However he, like Ryle, showed no inclination to pursue this idea any further, unless one construes the term ‘phenomenology’ in a scaled-down (normalised analytic) sense that would leave it quite devoid of any distinctive or substantive implications. That is to say, if the remit of phenomenological enquiry is confined to a purely descriptive account of our everyday linguistic practices – no matter how sharp-eyed, detailed, or meticulous – then it will find no room for those other, doubtless more ‘metaphysical’ sorts of question that have continued to preoccupy thinkers in the post- Kantian European line of descent.

## V

It has been my contention throughout this essay that such issues are absolutely central to any philosophical enterprise worthy the name, and that they cannot be analysed away by some well-practised technique of conceptual or linguistic problem-control. Nor is there much benefit to be had from those recent attempts at a partial *rapprochement* – like McDowell’s semi-naturalised, detranscendentalised, Wittgenstein-influenced and studiously non- ‘metaphysical’ reading of Kant – which evade the most challenging aspects of that ‘other’ tradition while they simply reproduce all its unresolved dilemmas in a different, less overt but no less troublesome guise.<sup>30</sup> Philosophy does best in relation to intractable issues like those of free-will/determinism or doxastic responsibility by keeping the problems firmly in view, resisting any premature claim to have resolved them decisively either way, but

also holding out against the twin temptations of a drive to dissolve them through conceptual analysis and a placid assurance that they cannot arise so long as our language stays in touch with the norms of communal usage. For these counsels cannot get us very far – whether in philosophy or in thinking more clearly about questions of an ethical, legal, social, or political import – when their effect is to close off precisely the kinds of engaged and responsible thought that constitute philosophy's chief claim to attention in such matters. Thus, for instance, Kant's passages in the First and Second *Critiques* concerning the Antinomies of Pure and Practical Reason are germane to any debate about issues of doxastic and moral responsibility, whatever one may think of his proposed 'solution' and the various dilemmas to which it gave rise.

My point – to repeat – is that philosophy risks inviting the charge of triviality or downright irrelevance if it adheres too closely to the mainstream-analytic, i.e., problem-solving (or problem-dissolving) mode of address to these issues and hence fails to register the depth and extent of their bearing on our moral, social, and intellectual lives.

That Kant in some sense got it wrong about ethics – that any too rigid (that is to say, *echt*-Kantian) application of his strict universalist claims might be apt to produce morally repugnant consequences in certain situations – is a case that has been rehearsed by good many recent commentators, amongst them advocates of a communitarian approach with strong Wittgensteinian leanings.<sup>31</sup> However there are ways of getting it wrong whilst none the less posing the crucial questions in a sharply-focused and provocative form that has more to teach us than any such recourse to anodyne, philosophically and morally evasive talk about shared language-games or communal practices. The same applies, as I have said, to Sartrean existentialism and its raising of the claim for human autonomy and free will to such a high point of absolute, intransigent principle that we encounter just the kind of choice that Sartre is so good at depicting in his works of philosophy and fiction alike.<sup>32</sup> That is, we are confronted with the need either to accept his extreme voluntarist position – along with its likewise extreme demands on our allegiance in the face of strong counter-arguments – or else to frame some viable, philosophically cogent alternative that would allow for certain kinds and degrees of unfreedom while none the less resisting any form of determinist doctrine. If relations had been less strained then analytic philosophers would hardly need telling that this debate has been carried forward to instructive effect by thinkers in the post-war French tradition, from Merleau-Ponty (whose critique of Sartre raises precisely these issues) to Derrida's later writings on the ethics and politics of deconstruction.<sup>33</sup> Moreover – no doubt through its acute responsiveness to episodes in recent French history – it has achieved a far wider and deeper socio-political resonance, as likewise with those various debates within post-war German (especially Frankfurt-School) philosophy where epistemological and ethical issues are often inseparably bound up with reflection on the problems and prospects of the German federal state.<sup>34</sup>

This is *not* for one moment to go along with that other, 'continentally'-inspired variant of the two-traditions story which would have it that analytic philosophy is a narrowly technical, politically disengaged mode of discourse concerned only with footling matters of linguistic or logico-semantic exegesis. Even if – as I have suggested – that charge has some force with regard to certain developments on the



analytic side, still it is very clearly wide of the mark when applied to the kinds of conceptual clarification and teasing-out of unnoticed complications in our political as well as ethical thinking that has characterised such work at its best. Amongst the many examples that might be offered I would mention in particular Jonathan Glover's *Humanity: a moral history of the twentieth century*, a book that most impressively combines breath of historical coverage with depth of philosophical reflection and a keen sense of how our moral judgements can be educated – rendered more acute but also less prone to readymade habits of response – with the aid of such reflection. In the case of such work it becomes just a pointless labelling exercise and one with profoundly misleading implications to place it on one or the other side of the Great Analytic/Continental Rift. What emerges, rather, is the two-way relationship between philosophy's need constantly to strive for a more adequate, that is, more intellectually responsible grasp of its own operative concepts and the need that those concepts should be exercised on matters of substantive (which will often mean complex, difficult, and at times sharply divisive) ethical import.

I have made the case here that both interests could be best served by a far more open and mutually responsive attitude in each quarter, though not without the kind of productive friction that comes of their different histories, interests, and modes of development since the time of that (albeit much exaggerated) parting-of-the-ways after Kant. The main cause of such friction so far – and the reason, no doubt, for its having up to now generated more heat than light – is the belief amongst many analytic philosophers that the other lot are more in the business of creating unnecessary trouble than of solving genuine problems, and the converse belief among many continentals that analytic philosophy amounts to no more than a set of well-practised but evasive techniques for denying the existence of just those problems. The *loci classici* here would include some that I have mentioned already such as Ryle's drastic change of mind with regard to the issues raised by Husserlian transcendental phenomenology and Austin's idea that the problem of knowledge (with its main source in Kant) could best be answered by recourse to the commonsense wisdom enshrined in 'ordinary language'. It is a notion that cuts across some otherwise deep divergences of view, as for instance between the Frege-Russell claim that such language stands in need of logical analysis so as to clarify or disambiguate its surface confusions and the Wittgensteinian assurance that all such problems can be made to disappear through the straightforward appeal to whatever makes sense by our own (or other people's) communal lights.<sup>35</sup> Then again, these stereotypical conceptions can be seen emerging at full force in the encounter – the 'determined non-encounter', as Derrida mock-ruefully declared it – between Derrida and Searle on the topic of Austinian speech-act philosophy.<sup>36</sup>

They can also be traced through the history of differing responses to Kant's Antinomies of Pure and Practical reason, that is to say, the issue as to whether these should be regarded as conceptual (or linguistic) aberrations in need of coaxing down from the giddy metaphysical heights or as genuine, deep-laid problems with a crucial bearing on the scope and limits of our freedom.

Since that issue is central to the debate concerning doxastic voluntarism – the question as to whether or just how far we can be held intellectually and morally responsible for the content of our various beliefs, convictions, ideological commitments, and so forth – it is one that cannot be raised without reference

(however guarded or oblique) to the kinds of discussion carried on within the ‘other’, post-Kantian continental tradition. In other words it is a distinctly *metaphysical* issue in so far as it involves considerations beyond anything resolvable (or even discussible) on the terms laid down by analytic philosophy either in its ‘ordinary-language’ (descriptivist) or its logico-semantic (revisionist) mode. To this extent it requires both the kind of far-reaching speculative thought that has characterised philosophy in the Kantian line of descent and (as a necessary complement to that) the kind of meticulous conceptual and linguistic analysis which has typified a good deal of mainstream analytic work. Thus the question of how best to reconcile our often conflicting intuitions in this regard – our sense that beliefs are (or ought to be) purely volitional with our countervailing sense of the need to make allowance for various limiting, e.g., causal or circumstantial factors – is a question that can and should cut right across these conventional boundary-markers.

At present, as I have said, there are strong signs of this awareness in the work of thinkers like McDowell who propose a return to Kant (or to certain select topics and passages in Kant) as offering a useful way forward from the perceived impasse of analytic philosophy in the wake of old-style logical empiricism and of Quine’s root-and-branch attack on its governing assumptions.<sup>37</sup> However this proposal is often couched in terms – like those specified by P.F. Strawson in an earlier episode of ‘back-to-Kant’ thinking – which go so far toward scaling down the metaphysical or transcendental dimensions of Kantian epistemology and ethics that what remains is more like a warmed-over version of the logical-empiricist programme.<sup>38</sup> Thus it tends to work out as yet another exercise in stipulative boundary-drawing whereby Kant’s more unfortunate (metaphysically loaded) kinds of talk can be adapted to the norms of a discourse which remains well within the bounds of analytic acceptability. Such, for instance, is McDowell’s idea of a ‘space of reasons’ wherein thinking can exercise its due prerogatives – those having to do with matters of humanly-intelligible motive, meaning, or intent – as opposed to the domain of empirical enquiry which is likewise subject to the normative standards and constraints of rational inference, but not (or not directly) to those of moral-intellectual autonomy and freedom.<sup>39</sup>

According to McDowell it is possible to maintain this distinction – and thereby conserve an adequate ‘space’ for the exercise of such freedom – without falling into the kinds of vicious dualism which have plagued Kantian and much post-Kantian philosophy, e.g., those between sensuous intuitions and concepts of understanding or the promptings of mere moral ‘inclination’, no matter how well-disposed, and the absolute dictates of moral law. We can best achieve this, he thinks, by switching attention to Kant’s talk of ‘receptivity’ and ‘spontaneity’, the latter conceived as ‘only notionally separate’ (since they are bound up in a relation of strict mutual dependence) and hence as offering a means of escape from the dualist impasse.

However, as I have argued at length elsewhere, when McDowell attempts to spell out the case in detail – to explain just how the autonomous-sounding claims of ‘spontaneity’ can be reconciled with those of ‘empirical constraint from the outside world’ – then it begins to look more like just another variant of that same old Kantian dilemma, one that has if anything been sharpened (not resolved or even somewhat clarified) by its recasting in these different terms.<sup>40</sup> ‘If we restrict ourselves to the standpoint of experience itself’, he suggests, then

what we find in Kant is precisely the picture I have been recommending: a picture in which reality is not located outside a boundary that encloses the conceptual sphere . . . . The fact that experience involves receptivity ensures the required constraint from outside thinking and judging. But since the deliverances of receptivity already draw on capacities that belong to spontaneity, we can coherently suppose that the constraint is rational; that is how the picture avoids the pitfall of the Given.<sup>41</sup>

One may question whether the claim of ‘coherence’ is aptly applied to so tortuous and convoluted a passage of reasoning. How ‘reality’ can possibly be thought of as exerting an external (empirical) constraint on any knowledge we can gain concerning it while all the same ‘not located outside a boundary that encloses the conceptual sphere’ is quite as problematic as anything encountered in Kant’s murkier ruminations on the topic. At any rate, if one thing is clear, it is the fact that McDowell is very far from having finally dismounted from what he calls the ‘seesaw’ – the chronic oscillation – that has been such a hallmark of epistemology from Kant to the present. Indeed, what gives his attempted solution its particular diagnostic value is the conceptual strain that emerges so vividly in passages like that cited above. Nor are these problems by any means confined to the epistemological sphere, since the issue of knowledge (of its normative claims or justificatory grounds) is one that cannot possibly be set aside in any adequate, i.e., reasoned and responsible treatment of ethical questions.

I have made this case specifically with regard to the debate about doxastic voluntarism since it is here – at the point of maximal conflict between autonomist and cultural-determinist views – that philosophy is brought up against the greatest challenge to its powers of rational arbitration. That is to say, there is little merit in any approach that fails to acknowledge the strength or intuitive force of arguments on both sides of this issue, or which treats it as merely a pseudo-dilemma brought about by our unfortunate proneness to forms of conceptual or metaphysical ‘bewitchment’. On the contrary: it is one that often arises in contexts ranging from the most rarefied levels of meta-ethical debate to other, more ‘applied’ or practical dimensions of moral philosophy and – beyond that – in the public domain where it bears directly on various concerns of a social, political, and legal character. Confronted with such difficulties thinking most frequently tends to react in one or other of the opposite ways I have outlined above. Thus it veers either toward the kind of briskly problem-solving approach that plays them down for the sake of conceptual clarity or deference to common linguistic usage, or else toward a mode of address which keeps them firmly in view though at risk of raising the resultant dilemmas to a high point of paradox and ultimate undecidability.

No doubt it would grossly simplify the issue to identify these two contrasting attitudes with the ‘analytic’ (i.e., mainstream Anglo-American) and ‘continental’ (i.e., post-Kantian mainland-European) lines of descent. Still this idea will do less harm if it is taken not, in the customary fashion, as a mark of reciprocal hostility or mutual indifference but rather as describing the tense yet productive – since in certain ways strongly antithetical – relationship between them. Such is the attitude adopted by some philosophers of a broadly ‘analytic’ persuasion when they read Sartre and register the force of certain existentialist claims even though predisposed by their background culture and intellectual training to cast a somewhat quizzical or sceptical eye on the more extreme statements of that doctrine.<sup>42</sup> With Derrida likewise one can

take his point about the irreducibility of ethical choice to rule, precedent, or formal prescription without endorsing the kind of full-fledged decisionist approach that would leave no room for the exercise of rational-deliberative thought.<sup>43</sup>

This is not to treat Derrida, or indeed Sartre, as mere *provocateurs* or intellectual gadflies whose sole claim on our interest is their knack of coming up with sharply-turned paradoxical formulations or skilfully contrived (often fictive) scenarios which may serve to administer a salutary jolt to our more routine or regimented habits of thought. To be sure, there is a prominent aspect of their writing that would seem to justify this rough characterisation, as likewise with a good deal of other work – some of it by Anglophone philosophers – which would count as ‘continental’ in terms of its distinctive thematic, stylistic, or generic attributes.<sup>44</sup> Still there is a crucial divergence of outlook between those who regard such work as *merely* a standing provocation when judged by the norms of rational, common-sense, responsible discourse and those who see in it the kind of provocation that Socrates offered when he challenged the conventional mores of his time, or that Hume put up against a whole range of orthodox philosophic and religious beliefs, or that Kant acknowledged when he credited Hume with having ‘aroused me from my dogmatic slumbers’.

It is this latter, more constructive and open-minded sort of response that holds out the best prospect of advance, not only as a matter of productive (mutually provocative) exchange between the ‘two traditions’ but also as concerns our central topic of doxastic responsibility. For if their difference can be summarised briefly yet without undue simplification it is the difference between a ‘continental’ way of proceeding that measures itself against the limits and extremities of philosophic thought and an ‘analytic’ discourse whose regulative principle – albeit with some leeway for testing or stretching its limits – is to draw such anomalies back within the compass of conceptual or linguistic normality. Such is also the tension that Derrida brings out through his early, meticulously detailed and rigorous (and in this sense properly ‘analytic’) readings of philosophers from Plato to Husserl.<sup>45</sup> Thus on the one hand he remarks that ‘a certain structuralism has always been philosophy’s most spontaneous gesture’ while on the other he acknowledges ‘the principled, essential, and structural impossibility of closing a structural phenomenology’.<sup>46</sup> Otherwise put, it is the constant oscillation between ‘hyperbole and finite structure’, or a thinking that challenges philosophy’s powers of self-assured conceptual grasp and a mode of thought – no less ‘principled and essential’ – which strives to contain that hyperbolic impulse within the bounds of established rational or logico-semantic intelligibility.

This seems to me the most fruitful way of conceiving the relationship between ‘continental’ and ‘analytic’ philosophy as that relationship has developed since Kant and, more pointedly, since the two traditions broke step over issues raised by Frege’s well-known criticisms of Husserl.<sup>47</sup> At the same time it offers a revealing approach to questions concerning the scope and limits of our moral-intellectual autonomy, whether raised (as by Sartre) in an overtly thematic existentialist mode or (as by Derrida) in terms of philosophy’s freedom – within certain specified procedural constraints – to challenge or to radically revise our understanding of canonical texts. What is involved in each case is a highly self-conscious and self-critical reflection on the character of those constraints and on the ways that such freedom can be exercised responsibly despite and against other, more orthodox (restrictive or coercive) habits of belief.

Christopher Norris  
*University of Cardiff*

---

<sup>1</sup> See for instance Jean-François Lyotard, *The Postmodern Condition: a report on knowledge*, trans. Geoff Bennington and Brian Massumi (Manchester: Manchester University Press, 1984); also – for a critique of such thinking – Christopher Norris, *The Truth About Postmodernism* (Oxford: Blackwell, 1993).

<sup>2</sup> See especially G.E.M. Anscombe, 'Causality and Determinism', in Ernest Sosa (ed.), *Causation and Conditionals* (Oxford: Oxford University Press, 1975), pp. 63-81. I put the case against such extrapolations from the micro- to the macrophysical domain in Norris, *Quantum Theory and the Flight from Realism: philosophical responses to quantum mechanics* (London: Routledge, 2000), esp. pp. 134-64.

<sup>3</sup> Ludwig Wittgenstein, *Philosophical Investigations*, trans. G.E.M. Anscombe (Oxford: Blackwell, 1951).

<sup>4</sup> See John McDowell, *Mind and World* (Cambridge, MA: Harvard University Press, 1994); also Robert Brandom, *Making It Explicit: reasoning, representing, and discursive commitment* (Harvard U.P., 1994); Wilfrid Sellars, 'Empiricism and the Philosophy of Mind', in Herbert Feigl and Michael Scriven (eds.), *Minnesota Studies in the Philosophy of Science*, Vol. 1 (Minneapolis: University of Minnesota Press, 1956), pp. 253- 329.

<sup>5</sup> Jean-Paul Sartre, *Being and Nothingness: an essay on phenomenological ontology*, trans. Hazel E. Barnes (London: Methuen, 1966).

<sup>6</sup> See especially Maurice Merleau-Ponty, *Adventures of the Dialectic*, trans. Joseph Bien (Evanston, IL: Northwestern University Press, 1973) and *The Prose of the World*, trans. John O'Neill (Northwestern U.P., 1974); also Jon Stewart (ed.), *The Debate Between Sartre and Merleau-Ponty* (Northwestern U.P., 1998).

<sup>7</sup> Sartre, *Critique of Dialectical Reason*, Vol. 1 (*Theory of Practical Ensembles*), trans. A. Sheridan-Smith (London: New Left Books, 1976); Vol. 2 (*The Intelligibility of History*), trans. Quintin Hoare (London: Verso, 1991).

<sup>8</sup> See for instance Colin McGinn, *The Mysterious Flame: conscious minds in a material world* (New York: Basic Books, 1999).

<sup>9</sup> See especially Paul M. Churchland and Patricia S. Churchland, *On the Contrary: critical essays, 1987-1997* (Cambridge, MA: MIT Press, 1998).

<sup>10</sup> J.S. Mill, *An Examination of Sir William Hamilton's Philosophy, and of the Principal Philosophical Questions Raised in his Writings* (London: Longmans, Green and Dyer, 1878).

<sup>11</sup> Immanuel Kant, *Critique of Pure Reason*, trans. N. Kemp Smith (London: Macmillan, 1954).

<sup>12</sup> See for instance McDowell, *Mind and World* (op. cit.).

<sup>13</sup> Christopher Norris, 'McDowell on Kant: redrawing the bounds of sense' and 'The Limits of Naturalism: further thoughts on McDowell's *Mind and World*', in *Minding the Gap: epistemology and philosophy of science in the two traditions* (Amherst, MA: University of Massachusetts Press, 2000), pp. 172-96 and 197-230.

<sup>14</sup> See Note 2, above.

<sup>15</sup> Benjamin Libet, *Mind Time: the temporal factor in consciousness* (Cambridge, MA: Harvard University Press, 2004); also Libet, Anthony Freeman and Keith Sutherland (eds.), *The Volitional Brain: towards a neuroscience of free will* (New York: Imprint Academic, 2000).

<sup>16</sup> See especially Wittgenstein, *Lectures and Conversations on Aesthetics, Psychology and Religious Belief*, ed. C. Barrett (Oxford: Blackwell, 1966); also *Last Writings on the Philosophy of Psychology*, Vols. 1 and 2, eds. G.H. von Wright and Heikki Nyman (Oxford: Blackwell, 1990 and 1993).

<sup>17</sup> Kant, 'The Paralogisms of Pure Reason', in *Critique of Pure Reason* (op. cit.), pp. 328-68.

<sup>18</sup> See John R. Searle, *Intentionality: an essay in the philosophy of mind* (Cambridge: Cambridge University Press, 1983) and *Minds, Brains and Science* (Cambridge, MA: Harvard University Press, 1984).

<sup>19</sup> See for instance Donald Davidson, *Essays on Actions and Events* (Oxford: Clarendon Press, 1980); also Jaegwon Kim (ed.), *Supervenience* (Aldershot: Ashgate, 2002).

<sup>20</sup> See Note 9, above.

<sup>21</sup> John Preston and Mark Bishop (eds.), *View Into the Chinese Room: new essays on Searle and artificial intelligence* (Oxford: Clarendon Press, 2002).

<sup>22</sup> For further discussion, see Joseph Levine, *Purple Haze: the puzzle of consciousness* (Oxford: Oxford University Press, 2002); Quentin Smith and Aleksandar Jokic (eds.), *Consciousness: new philosophical perspectives* (Oxford U.P., 2003).

<sup>23</sup> See also Norris, *Minding the Gap* (op. cit.).

<sup>24</sup> See Note 13, above.

<sup>25</sup> J.L. Austin, 'A Plea for Excuses', in *Philosophical Papers* (Oxford: Oxford University Press, 1961), pp. 123-52.

<sup>26</sup> See for instance Keith Graham, *J.L. Austin: a critique of ordinary language philosophy* (Hassocks: Harvester Press, 1977).

<sup>27</sup> Gilbert Ryle, *Dilemmas* (Cambridge: Cambridge University Press, 1954).

<sup>28</sup> Ryle, 'Phenomenology', 'Review of Martin Farber, *The Foundations of Phenomenology*', and 'Phenomenology versus *The Concept of Mind*', in Ryle, *Collected Papers*, Vol. 1 (London: Hutchinson, 1971), pp. 167-78, 215-24 & 179-96.

<sup>29</sup> See Austin, *Philosophical Papers* (op. cit.), p. 182.

<sup>30</sup> See Notes 4 and 13, above.

<sup>31</sup> For an early and highly influential statement of this view see Peter Winch, *The Idea of a Social Science and Its Relation to Philosophy* (London: Routledge & Kegan Paul, 1958) and *Trying to Make Sense* (Oxford: Blackwell, 1987).

<sup>32</sup> See Note 5, above.

<sup>33</sup> See Note 6, above; also Jacques Derrida, *Aporias: dying – awaiting (one another at) the 'limits of truth'*, trans. Thomas Dutoit (Stanford, CA: Stanford University Press, 1993); 'Force of Law: the "mystical foundation of authority"', trans. Mary Quaintance, *Cardoso Law Review*, Vol. XI (1990), pp. 999-1045; 'At This Very Moment in This Work Here I Am', trans. Ruben Berezdivin, in Robert Bernasconi and Simon Critchley (eds.), *Re-Reading Levinas* (Bloomington, Ind.: Indiana University Press, 1991), pp. 11-40

<sup>34</sup> See especially Jürgen Habermas, *Justification and Application: remarks on discourse ethics*, trans. C.P. Cronin (Cambridge, MA: MIT Press, 1993) and *Between Facts and Norms: contributions to a discourse theory of law and democracy*, trans. W. Rehg (MIT Press, 1996); also J. Bohmann and W. Rehg (eds.), *Deliberative Democracy: essays on reason and politics* (MIT Press, 1997).

<sup>35</sup> Gottlob Frege, 'On Sense and Reference', in Peter Geach and Max Black (eds.), *Translations from the Philosophical Writings of Gottlob Frege* (Oxford: Blackwell, 1952), pp. 56-78; Bertrand Russell, 'On Denoting', *Mind*, Vol. XIV (1905), pp. 479-93; Wittgenstein, *Philosophical Investigations* (op. cit.).

<sup>36</sup> See Jacques Derrida, 'Signature Event Context', *Glyph*, Vol. 1 (Baltimore: Johns Hopkins University Press, 1975), pp. 172- 97; John R. Searle, 'Reiterating the Differences', *ibid*, pp. 198-208; Derrida, 'Limited Inc abc', *Glyph*, Vol. 2 (1977), pp. 75-176; also Derrida, 'Afterword: toward an ethic of conversation', in Gerald Graff (ed.), *Limited Inc* (Evanston, IL: Northwestern University Press, 1989), pp. 111-54.

<sup>37</sup> McDowell, *Mind and World* (op. cit.); W.V. Quine, 'Two Dogmas of Empiricism', in *From a Logical Point of View*, 2nd edn. (Cambridge, MA: Harvard University Press, 1961), pp. 20-46.

<sup>38</sup> See P.F. Strawson, *Individuals: an essay in descriptive metaphysics* (London: Methuen, 1959) and *The Bounds of Sense: an essay on Kant's Critique of Pure Reason* (Methuen, 1966).

<sup>39</sup> McDowell, *Mind and World* (op. cit.) and other references under Note 4, above.

<sup>40</sup> See Note 13, above.

<sup>41</sup> McDowell, *Mind and World* (op. cit.), p. 41.

<sup>42</sup> See especially Gregory McCulloch, *Using Sartre: an analytical introduction to early Sartrean themes* (London: Routledge, 1994).

<sup>43</sup> See Note 33, above.

<sup>44</sup> For further discussion, see Christopher Norris, *Minding the Gap* (op. cit.).

<sup>45</sup> See especially Derrida, '*Speech and Phenomena*' and *Other Essays on Husserl's Theory of Signs*, trans. David B. Allison (Evanston, IL: Northwestern University Press, 1973); *Writing and Difference*, trans. Alan Bass (London: Routledge & Kegan Paul, 1978); *Dissemination*, trans. Barbara Johnson (London: Athlone Press, 1981).

<sup>46</sup> Derrida, *Writing and Difference* (op. cit.), p. 162.

<sup>47</sup> See Gottlob Frege, review of Edmund Husserl's *Philosophie der Arithmetik*, translated by E.-H. W. Kluge, *Mind*, Vol. LXXXI (1972), pp. 321-37; Gilbert Ryle, 'Phenomenology', 'Review of Martin Farber, *The Foundations of Phenomenology*', and 'Phenomenology versus *The Concept of Mind*' (Note

---

28, above); also Derrida, "Genesis and Structure" and Phenomenology', in *Writing and Difference* (op. cit.), pp. 154-68 and *La probl me de la gen se dans la philosophie de Husserl* (Paris: Presses Universitaires de France, 1990).