COLIN McGINN

# THE PROBLEM OF PHILOSOPHY

(Received 9 March 1994)

## I

The question of the scope of human knowledge has been a longstanding preoccupation of philosophy. And that question has always had a special intensity where philosophical knowledge itself is concerned. A certain anxiety about the nature and possibility of such knowledge is endemic to the subject. The suspicion is that, in trying to do philosophy, we run up against the limits of our understanding in some deep way. Ignorance seems the natural condition of philosophical endeavour, contributing both to the charm and the frustration of the discipline (if that is the right word). Thus a tenacious tradition, cutting across the usual division between empiricists and rationalists, accepts (i) that there are nontrivial limits to our epistemic capacities and (ii) that these limits stem, at least in part, from the internal organisation of the knowing mind – its constitutive structure – as distinct from limits that result from our contingent position in the world. It is not merely that we are a tiny speck in a vast cosmos; that speck also has its own specific cognitive orientation, its own distinctive architecture. The human mind conforms to certain principles in forming concepts and beliefs and theories, originally given, and these constrain the range of knowledge to which we have access. We cannot get beyond the specific kinds of data and modes of inference that characterise our knowledge-acquiring systems – however paltry these may be. The question has been, not whether this is correct as a general thesis, but rather what the operative principles are, and where their limits fall. *How* limited are we, and what explains the extent and quality of our limits? Can we, indeed, come to understand the workings of our own epistemic

*Philosophical Studies* **76:** 133–156, 1994.

capacities? Hence the enquiries of Descartes, Locke, Leibniz, Hume, Kant, Peirce, Russell, and many others.

The most recent major theorist in this tradition, and perhaps the most explicit, is Chomsky.[1] According to him, the mind is a biologically given system, organised into discrete (though interacting) subsystems or modules, which function as special-purpose cognitive devices, variously structured and scheduled, and which confer certain epistemic powers and limits on their possessors. The language faculty is one such module: innately based and specifically structured, it comes into operation early in human life and permits the acquisition, or emergence, of an intricate cognitive system in a spectacularly short time – this being made possible by the antecedent presence of the principles of universal grammar in its initial design. As Chomsky observes, the knowledge so generated is no simpler, by any plausible objective standard, than knowledge of advanced mathematics or physics; but the human mind is so adapted that it yields this knowledge with comparative ease – somewhat as we effortlessly develop a complex physiological structure in a pre-programmed way. (Compare the ease with which our visual system converts two-dimensional arrays into three-dimensional percepts, but the difficulty we have in making even simple two-dimensional drawings on the basis of our three-dimensional visual experience.) As a corollary, however, this faculty is poorly adapted to picking up conceivable languages distinct in grammatical structure from that characteristic of human speech. Its strength is thus also its weakness; in fact, it could not be strong in one way without being weak in another.

With language as his model case Chomsky develops a general conception of human intelligence which includes the idea of endogenously fixed cognitive limits even for conscious reason. Here, too, the price of ready success in some domains is fumbling or failure in others. He says:

> The human mind is a biologically given system with certain powers and limits. As Charles Sanders Peirce argued, "Man's mind has a natural adaptation to imagining correct theories of some kinds. ... If man had not the gift of a mind adapted to his requirements, he could not have acquired any knowledge." The fact that "admissible hypotheses" are available to this specific biological system accounts for its ability to construct rich and complex explanatory theories. But the same properties of mind that provide admissible hypotheses may well exclude other successful theories as unintelligible to humans. Some theories might simply not be among the admissible hypotheses

determined by the specific properties of mind that adapt us "to imagining theories of some kinds," though these theories might be accessible to a differently organised intelligence. Or these theories might be so remote in an accessibility ordering of admissible hypotheses that they cannot be constructed under actual empirical conditions, though for a differently structured mind they might be easily accessible.[2]

Among the theories that he thinks may *not* be accessible to human intelligence, in virtue of its specific slant, Chomsky includes the correct theory of free creative action, particularly the ordinary use of language. We seem able to develop adequate theories of linguistic competence, i.e. grammars, but when it comes to actual performance our theoretical insights are meagre or nonexistent. And this is a reflection of the contingencies of our theoretical capacities, rather than an indication of objective intransigence.

Now much could be said in explication and defence of Chomsky's general position, but that is not my purpose here. I wish to start from something like his general perspective and explore some questions seemingly at some distance from Chomskyan concerns: in particular, I want to ask whether the phenomenon of philosophical perplexity might be a consequence of the kind of constitutive cognitive inaccessibility of which he speaks. Is the hardness of philosophy a result of cognitive bias? Might our difficulties here be a side-effect of our adeptness in other areas? Where does the felt profundity of philosophical questions come from? But first I shall have to make some further general remarks about the idea of cognitive limitation; for we will not be in a position to approach my main question unless we have properly taken the idea of cognitive limitation to heart.[3]

First, it is easy to see that comparable limitation theses hold with respect to other aspects of our mental life. We cannot experience every possible type of sensation, nor every emotion; neither can we desire everything that might conceivably be desired. Differently constituted minds from ours might well enjoy a different range of phenomenal, affective and conative states. There are also sensory thresholds of various kinds which fix the bounds of our perceptual acuities, and which can vary across perceivers, as well as obvious restrictions on our memory capacities and reasoning power. These limits are not in any way dictated by objective phenomena but stem rather from our species-

specific endowments. They give us a particular psychological profile, not necessarily shared by other species, actual or possible. Indeed, it is hard to know what it would be like for a psychological being not to have limitations of these kinds, since such limitations are a direct consequence of having any determinate psychology at all.[4]

Second, we should distinguish between two potential sources or loci of cognitive limitation: one relating to the content of our mental representations, the other to the specific character of the operational system within which those contents occur. That is, there is the question of what range of concepts we can in principle deploy in our thought; and there is the separate question of the processing principles and architecture of the system that contains these concepts. Even with unlimited access to concepts, a system might be confined by what it can do with them – say, because of attentional or memory limitations. And a system might be quite impoverished conceptually but be capable of amazing feats of processing and deployment. So when considering whether a certain cognitive system is capable of a given task we need to ask both whether it can acquire mastery of the relevant concepts and whether it has the organisational resources to put these concepts to work in the necessary way. One live possibility is that the mind is not notably lacking at the level of individual concepts but that it lacks the capacity to combine these into systematic explanatory theories of some given class.

Third, if there are the kinds of cognitive predisposition Chomsky suggests, then we should be on the look-out for tendencies to mislocate the source of our epistemic triumphs and failures. Since the limits imposed by our mental organisation are not guaranteed to present themselves as such, we may find ourselves attributing blame to the wrong thing: we may assume that what comes easy to us *is* (intrinsically) easy, and that what comes hard is somehow objectively recalcitrant. Thus we might be forgiven for supposing, mistakenly, that grammar is objectively simple compared to (say) relativity theory; but, rightly considered, this is a projective fallacy, borne of our peculiar endowments and correctible by an impartial examination of the structure of the systems of knowledge in question. The ease of accessibility of a knowledge system to our cognitive capacities is no measure of its internal complexity or subtlety or profundity – still less of the ontological fibrillations proper to the

subject-matter of the system. Indeed, it is unclear, ultimately, whether there is any (useful) notion of simplicity or complexity that is quite unrelativised to the specific aptitudes of a selected cognitive faculty. That reason is flummoxed by a certain class of problems is thus no proof that those problems possess any inherent refractoriness, nor that there are no other conceivable epistemic systems that might take these problems in stride.

It may be thought that the existence of nontrivial epistemic limits is a peculiarity of a particular philosophical tradition and that other viewpoints will have less restrictive consequences. Let me then dispel this impression by surveying briefly some standard theories (or theory-sketches) about the nature of thought; we shall see that limits are actually the norm, at any rate by implication. In fact, one of the recurrent faults of the usual theories is that they tend to delimit our conceptual powers *too* narrowly. In any case, it is hard to see how any substantive theory of concepts could avoid imposing *some* limits on concept possession, since certain constitutive conditions will have to be laid down – and hence not necessarily be satisfied. And the more substantive the theory is the clearer the limits are apt to become; only vacuous theories give the impression of boundlessness – as if concepts were entirely weightless and shapeless beings. It might help in counteracting this subliming tendency (as Wittgenstein called it) to consider nonhuman thinkers, like dogs and dolphins, when reviewing the theories on offer; for deification comes harder in their case than for our own superlative species – at least for us. Here the idea of cognitive limits seems only right and proper.[5]

Three broad types of theory can be distinguished: sensory, behavioural, external. By sensory theories I mean those that base concepts on the contents of perceptual experience – Locke and Hume being the usual suspects. When concepts are construed in this way they are clearly, as those two were keen to stress, constrained by the sensory powers of the creature in question; they are just the traces left by the activity of the sense organs on the memory faculty. Abstraction and association may enlarge the mind's stock of sense-based representations, but concepts of the strictly nonsensory are ruled out. The key tenet of empiricism, indeed, is just that thought cannot transcend the experiential (hence the impossibility of metaphysics, according to the positivism

that sprang from these empiricist principles.) It is not, then, just the rationalist tradition, with its emphasis on rich innate structure, that issues in restrictions on thinkability; in fact, in its very cognitive sparingness, empiricism imposes even more pronounced limits than rationalism. The acuity and scope of the senses is the measure of conceptual power, and where creatures differ in their sensory equipment they must also differ in what they can think.

Behaviour-based theories also impose limits, at least under pretty unavoidable assumptions. The central point is simply that behavioural theories tie concepts to the bodily repertoire of the thinker and bodies have determinate structure and powers, varying from one kind to the next. Behaving bodies are natural objects in the world, finite and bounded, with limited histories and sets of dispositions. An organism's bodily characteristics fix the nature of the inputs and outputs it can handle, but these are bound to be restricted by the facts of anatomy and physiology. If concepts are to consist in the motion of bodies, then the natural limits on motor capacities become the limits of concept possession. A vivid (if controversial) illustration of the kinds of limits that can result from behavioural theories is provided by Quine's indeterminacy thesis.[6] If concepts come down to dispositions to assent to sentences in specific stimulus conditions, then (i) theoretical concepts lapse into radical indeterminacy and (ii) we cannot expect to distinguish concepts that apply under the same conditions of stimulation – as with those rabbits and their undetached parts. Concepts have content, for Quine, only in so far as they are keyed to discrete dispositions to assent, but then any putative concepts not so keyed turn out to be either cognitively inaccessible or plain impossible – despite the reality of the properties they purport to represent. Much the same can be said of Dummett-style 'manifestation' requirements on meaning, which cannot make room for any concepts that call for a 'realist' interpretation.[7] The requirement of an effective mapping from concept to behavioural capacity confines concepts to the causal powers of the body in question. Functionalist theories have much the same upshot, since the causal role of an internal state clearly depends upon the contingent make-up of the organism; and if a body fails to provide a basis for some role then the corresponding concept will not be available to the creature whose body it is. Since,

presumably, human bodies (say) do not instantiate every logically possible causal role, there are bound to be concepts that are not open to us (given that every such role corresponds to a potential concept). Just as functionalist theories impose limits on the sensations a creature may possess, so they impose limits on its conceptual powers. And so it is with any theory that equates concepts with dispositions of the body. Even Wittgenstein's much more relaxed emphasis on the connexion between meaning and acting has limitative consequences, which he did not forswear.

Third, there are externalist theories that see content as fixed by head/world relations: causal, nomic, teleological and so forth.[8] Take, as representative, the simple idea that the concepts you have are determined by your history of environmental elicitations. Then your concepts will be limited both by the nature of the impinging environment and by the capacity of your sensory transducers to respond to what is offered up to them (what you can 'interact' with). According to some theories, you simply cannot have concepts for things you have not had causal commerce with – for example, natural kinds whose instances you have not encountered. At the extreme of causal isolation, as with the brain in a vat, you cannot even have concepts of the ordinary perceptible world.[9] Similarly, if we are now not suitably hooked up to some part or aspect of the objective world, then we will not be able to form representations of that part or aspect. The danger in theories of these kinds is actually that they impose unreasonably restrictive conditions on concept possession, underestimating the creative resources of the mind; they certainly do not allow untrammelled conceptual access by sheer effort of will.

It is not that I think any of these theories of concepts is really adequate; my point is just that it is not merely an eccentricity of the tradition in which Chomsky locates himself that thought should be subject to significant limits. This is implicit even in theories that are not advanced with this kind of issue in mind; and, as I remarked, it is hard to see how a theory could be both substantive and free of limitative consequences. For what could a concept consist in that was *not* in some way bound by inherently variable and potentially absent facts? Certainly, it is scarcely plausible that every logically possible concept should be necessarily accessible by any mind capable of grasping *some* concept or other. Only

a kind of mystical thinking about concepts could occlude recognition of the virtual truism that someone might be able to think some things without being able to think all things. (Compare the question whether humans possess every conceivable motor skill in virtue of possessing some.)

## II

I hope now that the sourness of the idea of cognitive limits is recognised, however grudgingly, as simply that of real life. I want to assume, anyway, that it is reasonable to expect, on general grounds, that some areas of human enquiry or interest will be subject to problems of cognitive penetrability, perhaps as a result of our talents in other directions. Then we can ask, with this general expectation in mind, whether we can identify any such areas in particular. Let us thus conduct an impartial survey of human cognitive effort to see whether anything looks like a plausible instance of epistemic boundedness. What actual evidence is there, with respect to particular areas, that the general expectation of selective cognitive failure is being fulfilled? This kind of question could, of course, be asked about many areas of human effort, and its general motivation and evidential status are much the same across the board. For example, we suppose, on general grounds, that humans will exhibit certain areas of strength in motor activity but also certain areas of weakness or total incompetence; and we can enquire into the empirical facts of the case to determine where in particular the capacities and incapacities fall – as we can for any species. Thus we learn to walk quite naturally and everyone is pretty efficient in this department; our swimming abilities, however, are laboriously acquired and show much individual variation; and when it comes to flying, well, it shouldn't even be attempted. This pattern of motor skill is presumably innately based and irremediable (short of fanciful surgical intervention). So, comparably, let us survey our cognitive skills to see where we are strong, weak and downright inept. And let us not be put off by, or misread, the general difficulty of empirically establishing the absence of a capacity. Admittedly, it can always be maintained that the capacity

is possessed but has somehow never quite locked onto its appropriate realising conditions, but after a point this just looks like special pleading and an unwillingness to take the evidence of incapacity at face-value. There are no doubt those who stubbornly insist that humans have the ability to fly if only they would flap their arms in exactly the right way, in the right wind conditions, and with the right degree of confidence in their hearts; and it is hard to dislodge this fantasy when not every such way has been tried and tested. But we can all see that this is not the rational conclusion to draw from the evidence: it is simply that humans, unlike birds, are not naturally equipped to fly, which is the precise reason we have not yet succeeded in doing it. Similarly, I suggest, we should assess the evidence of intellectual incapacity in the same impartial way, according the relevant indications their due weight.[10] In short: no wishful thinking!

Some of what we know comes very easily. As Chomsky has long emphasised, we develop a complex competence in language with remarkably little effort; and this is best explained by supposing an innate and specific preparedness on the part of the human cognitive system. Similar hypotheses are also plausible for visual perception, face recognition, knowledge of commonsense psychology, and no doubt other areas. General evolutionary considerations, as well as ease of acquisition and uniformity across individuals, suggest the idea of unfolding innate endowments, special-purpose modules. But not everything we know comes so readily or shows the same independence of individual intelligence: some human knowledge requires conscious deliberate mental labour. Chomsky cites knowledge of physical science as an obvious example: here it appears plausible to suppose that the knowledge in question is not antecedently targeted or anticipated – rather, it is made possible by the deployment of cognitive capacities that serve some other primary purpose (I shall consider later what this might be).[11] There is no innate structure in humans that already encodes the laws of physics. Such hard-won knowledge is genuinely learned; it is not merely triggered by outside stimuli. It is to our innate cognitive capacities what ballet dancing is to our innate motor capacities: a kind of offshoot or divagation, calling for much stretching and cultivation. The will enters

essentially into its generation. While humanly possible it is not (in one good sense) humanly natural.

But are there yet other areas of cognitive effort where we find things tougher going still – in which we seem chronically unable to make significant progress? Are there areas in which we simply lack the capacity to generate the kind of knowledge we desire? Signs of such an underlying incapacity would be: a stubborn lack of progress over time, both individually and across generations, with no obvious explanation in terms of objective complexity or remoteness or other exogenous factor; a subjective sense of intellectual cramp, inter-subjectively verified, where the very concepts with which to initiate and prosecute enquiry seem not to be at hand, the problems presenting an appearance of internal recalcitrance; the monotonous recurrence of the same unsatisfactory alternatives, with short-lived fashions instead of the steady elimination of unworkable theories and a growing convergence of opinion; a temptation to put all this down to the malign effects of disguised pseudo-problems. None of this would, of course, *entail* that the domain in question is one to which the human cognitive apparatus is constitutionally unsuited, but to an impartial observer it would provide *prima facie* evidence of a mismatch between the kind of theory needed and the cognitive tools being brought to the search. It would suggest the hypothesis that we are beyond the rim of human intellectual competence.

And now my point is just this: large parts of what is called 'philosophy' exemplify the above general description, so that the hypothesis of cognitive transcendence is at least a reasonable conjecture.[12] If this hypothesis were right, then the search for philosophical knowledge would be an attempt to do with our epistemic capacities what cannot be done with them. Our minds would be to philosophical truth what our bodies are to flying: wrongly designed and structured for the task in question. Let me emphasise that this is a *hypothesis*: it is to be viewed as the most plausible explanation of the data, compared to other proposed explanations, and it fits our best picture of the kind of thing the knowing organism is. Like any hypothesis of comparable scope and generality it might, of course, be mistaken; but I suggest that it is worth taking seriously and examining on its merits. After all, it simply applies to the so-called 'higher cognitive functions' what is acknowledged to be the

general condition of our various faculties, bodily and psychological. It competes, say, with the hypothesis (never to my knowledge advanced) that in fact humans do have a natural adaptation towards philosophical understanding, comparable to their innate expertise in language, but that this adaptiveness operates only during a 'sensitive period', say from five to eight years old, in which great strides would be made in philosophical inquiry if only we exposed our children to an intensive course in philosophical training during that period. We just don't get them early enough! Presumably this hypothesis, though implausible in the extreme (but why exactly is that?), is not *logically* excluded, and has never been empirically tested in any systematic way. It is at least among the range of hypotheses about human knowledge that we have learned to take seriously, at least as to its form. Well, my competing hypothesis asserts, not that we are missing a sensitive period for solving philosophical problems, but rather that the human cognitive system is just not set up for dealing with problems of this general type. This does not exclude the possibility that a differently organised intelligence might relate to philosophy as we do to physics,or indeed to language or commonsense psychology. For all I know, there are forms of intelligence out there that *do* go through a sensitive period for solving the problems of philosophy: if you miss it, you never pick up what your conspecifics take for granted – a thorough understanding of the phenomena that so perplex our earthly philosophers. According to my hypothesis, however, humans are constitutionally *in*sensitive where philosophical problems (of a certain kind) are concerned. In the rest of this paper I shall consider the prospects for this hypothesis.

## III

Let us call the hypothesis 'transcendental naturalism', TN for short, because it combines deep epistemic transcendence with the denial that what thus transcends is thereby non-natural. How well does TN account for the oddities of philosophical inquiry? I have considered this question, and allied issues, at some length in a book, and I cannot here repeat everything I say there.[13] Instead I shall try to summarise the

main points, providing what I hope will be a synoptic overview of the position. To this end, I begin with a sketch of the typical geography of philosophical debate; the suggestion will be that TN both predicts this geography and is itself superior to the sorts of position routinely adopted within it. After that I shall offer a conjecture about what it is that distinguishes (certain) philosophical problems from other problems we find cognitively amenable. All this will be highly speculative, naturally, and excessively compressed, and no doubt grievously flawed: but speculation can be audacious and risky without being irresponsible – and what else is philosophy for anyway?

When human minds interact with philosophical problems, especially those of the form 'How is X possible?', they are apt to go into one of four possible states. Either (i) they try to *domesticate* the object of puzzlement by providing a reductive or explanatory theory of it; or (ii) they declare it *irreducible* and hence not open to any levelling account; or (iii) they succumb to a *magical* story or image of what seems so puzzling; or (iv) they simply *eliminate* the source of trouble for fear of ontological embarrassment. For ease of reference, I call this pattern of responses the DIME shape. The topics on which it imprints itself, and which I have discussed in some diagnostic detail in the aforementioned book, include: consciousness and the mind-body problem, the nature and identity of the self, the foundations of meaning, the possibility of free will, the availability of a priori and empirical knowledge. In each of these areas, I claim, we can discern the same fundamental pattern of debate as the object of perplexity taxes our intellectual resources, pushing us in one direction or the other. For example, consciousness familiarly provokes the following set of philosophical reactions: attempts to explain it in naturalistic, usually physicalistic, terms; declarations that is brutely irreducible and *sui generis*; invocations of non-natural forces and relations; denials that there is really any such thing as consciousness to begin with. And much the same can be said for the other topics mentioned, despite some variation in the details. (I must here leave it to the reader to impose the DIME shape on debate about those other topics; or she can always have a look at my book). Basically what we find, quite generally, is the threat of magic or elimination in the face of the theoretical obduracy of the phenomenon that invites philosophical attention.

The phenomenon presents initial problems of possibility, which we try to dissolve with a domesticating theory, but there is always the danger that the failure of this undertaking will leave us facing magic or elimination or unwanted inexplicabilities. Free will, for instance, looks upon early inspection to be impossible, so we try to find some conception of it that permits its existence, but this conception always turns out to be dubiously reductive and distorting, leaving us with the unpalatable options of magic, elimination or quietism.[14] And so we hop unhappily from one unsatisfactory option to the next; or dig our heels (squintingly) into a position that seems that least intellectually unconscionable or the bunch.

Now TN has a view about this familiar fix: it is because the correct theory is inaccessible to the human intellect that we inflict the DIME torment on ourselves. Since we cannot get our minds around the portion of intellectual space where the correct theory lies, we are prone to dart off in inappropriate directions. Suppose that M is the right theory of the mind-body link but that we cannot, constitutionally, reach M: then we are apt to settle for some deforming domestication programme, or to say that there is just no theory to be had, or to conjure up a pseudo-explanatory magical story, or to get rid of the thing that leads to the problem. If we *could* reach M, then we would be able to accept consciousness in its undistorted form, and dismiss the usual DIME options: for we would have a proper theory of precisely the kind of thing that consciousness is. What TN counsels is that we be guided by the truth of that counterfactual, accepting the undiluted existence of something we cannot comprehend. TN says that it is because we cannot gain access to the concepts and principles required to make sense of consciousness (say) that we allow ourselves to be taken in by the DIME shape. But it is better to accept that the world contains things whose ultimate nature we cannot penetrate.

In such a case, according to the TN hypothesis, there comes to be a subject called 'philosophy', with its peculiar addiction to insoluble mysteries. Minds that were better tuned to the requisite theories would have no use for the category of the philosophical, or might perhaps include a quite different set of problems within it. Science, then, might be aptly characterised as that set of questions that does not attract the DIME options – where our cognitive faculties allow us to form the

necessary concepts and theories. The distinction between science and philosophy is, on this view, at root a reflection of the cognitive powers we happen to possess or lack, and is therefore creature-relative: it does not correspond to any interesting real division within objective reality. Conceivable creatures might invert the classification we make with these concepts, finding consciousness and free will easy to penetrate and explain scientifically, while being quite mystified by the movement of the planets or the nature of digestion. For it is not, for TN, *intrinsic* to consciousness and free will that they should occasion the kind of perplexity they do in minds like ours; such perplexity results, rather, from the *interaction* between a certain natural phenomenon and a certain type of cognitive set-up. It is not beyond the bounds of possibility that our brains would have to be made of something other than neurons in order for us to have the kinds of cognitive powers needed to solve the problems philosophy poses; at any rate, this is the *sort* of diagnosis TN offers for our philosophical retardation. Evolution selected neural tissue, suitably arranged, as the machinery for making intelligence, but that decision is surely substantive; perhaps other materials are used elsewhere in the universe, producing different sorts of intelligence from the earthly kind. The hardness of philosophy is thus an upshot of the particular way that natural selection has built our thinking organ, not an objective trait of the subject-matter of philosophical questions.[15]

## IV

So far we have seen some general motivation for TN, deriving from a certain conception of cognitive capacity, and we have considered some evidence that can be interpreted as favouring the hypothesis, though not of course conclusively. Crudely put, the idea has been this: philosophical problems are uniquely recalcitrant, chronically so, though nothing about their subject-matter entails this; but it is actually quite predictable that there should exist problems with this degree of intractability, given the most plausible view of the kind of system the human mind is; hence philosophy is a good candidate for being an instance of what is bound to be so on general grounds. But it would obviously be desirable

for TN to have some positive theory about the inner structure of our theoretical capacities from which the claimed partition among problems would follow. In the case of language, universal grammar is what plays this role: the language faculty is internally structured according to that specific set of principles, and any possible language that fails to conform to these principles will not be accessible to a faculty so structured. Linguists have theories about the form of universal grammar, so in effect they have theories about the limits of the human language faculty. Can we come up with anything comparable with respect to the human theoretical capacity? What specific characteristics of conscious reason put philosophy beyond its scope? What kinds of theories are accessible to or belief-forming mechanisms? What does it take for something to be intelligible to us?

Let me begin by rehearsing a suggestion of Chomsky's about our grasp of number theory.[16] He observes that natural human languages possess the property of 'discrete infinity' – roughly, they comprise a system of distinct basic elements that can combine to produce infinitely many complex wholes. Bee languages, by contrast, lack discreteness, being analogue systems, while other animal signal systems fail of infinitude. Presumably the property of discrete infinity arose from some specific biological adaptation, which was then exploited to generate languages as we know them. Chomsky speculates that this feature of our linguistic competence may be the basis of our ability with numbers, since the number series also exhibits discrete infinity, albeit over a distinct domain. If so, then our arithmetic faculty is a by-product of our linguistic faculty, got by abstracting from one domain to another. Pursuing Chomsky's speculation, we might go on to see the cognitive structure thus made available by this extension from language as a central element in our general ability to formulate intelligible theories of the world. That structure enables us to conceive of arbitrary domains in terms of combinatorial rules that generate a potential infinity of derived entities from a fixed set of individual elements. Thus it is that the crucial notion of *compositionality* enters our thinking, cropping up in many unrelated areas, and allowing us to generate theories in which it essentially features.

Suppose now that a representation of this abstract property were to join with the kind of spatial representation employed by our senses, notably vision, so that a kind of cooperation of faculties was initiated.[17] Then we might expect a mode of cognition that deals in discrete elements embedded within a continuous medium and capable of rule-governed processes of agglomeration. This would be suitable as a basis for representing the world of material objects in space, these being systems of combined elements, variously located, and capable of assuming indefinitely many different forms. So we can understand, at least in broad outline, how our grasp of physics might arise from grammar plus spatial representation – as arithmetic arises, according to Chomsky's speculation, from the iterative character of language. Here then we can tell the beginnings of a by-product story about our knowledge of physical science, and incidentally explain why such knowledge does not arise with the kind of spontaneity we observe elsewhere.

But if any area of human enquiry required us to go radically beyond the kinds of cognitive principles thus made available, we would find ourselves bereft of the intellectual resources with which to handle that area: our cognitive equipment would not be adequate to the objective properties of what confronts us. (Compare the blankness exhibited by most animals with respect to any representational medium: they lack the idea of *reference*.) That we can understand a given domain depends, according to this story, on a more or less fortuitous match between that domain and the domains targeted by the faculties from which the relevant modes of representation were derived. In other words, cognitive accessibility is a function of similarity to the concerns of our linguistic and perceptual faculties; crucially, it turns upon the applicability of the combinatorial paradigm supplied by language.[18]

This line of thought then suggests the following conjecture: what distinguishes the two kinds of questions is the applicability or otherwise of the 'discrete infinity' mode of understanding, supplemented by the sorts of representation with which our senses operate. That is the demarcation line that separates what we can make theoretically intelligible to ourselves and what we cannot. I refer to this as the CALM conjecture: Combinatorial Atomism with Lawlike Mappings.[19] It says, roughly speaking, that we can understand what conforms to CALM

principles, and we cannot understand what does not. According to the Chomskyan speculation, as I interpret it, the CALM schema has its roots ultimately in the structure of language itself, where words combine by rules into phrases and sentences – the abstract character of this property then being detached from the particular atoms to which it originally applies. I think this explains pretty well (though rudimentarily) how we come to have geometry and arithmetic and linguistic theory and physics and even biology – the domains that are tolerably transparent to us. The question now is whether it also hints at why we do *not* have philosophy. Do our difficulties here arise from the circumstance that the phenomena of interest to us cannot be made to conform to the paradigm of a collection of elements that combine lawfully into complex wholes which depend for their properties upon those of their constituent parts? Is it a basic lack of CALM that generates philosophical perplexity?

To answer this question fully we would need to conduct a careful survey of the standard and central problems of philosophy, trying to determine whether it is a systematic breakdown of CALM that produces these problems. I have attempted this in the aforementioned book; let me here cite just one area in which the conjecture seems to carry considerable plausibility – namely, the mind-body problem. Consider how consciousness relates to its neural substrate: then the problem is simply that no intelligible generative relation can be identified. In particular, we cannot regard conscious states as complexes made up of neural elements. So one good way to formulate the mind-body problem is precisely to say that conscious states cannot be regarded as compound structures derived from the neural units that correlate with them. Sensations do not stand to neurons as sentences stand to words or as macroscopic bodies stand to molecules. Hence the usual (and reasonable) talk of radically emergent properties, of explanatory gaps, of peculiar kinds of novelty in the world. In some of its aspects the brain no doubt conforms to CALM principles, as cellular structures do generally; but there must be other aspects of it too, if it is to be capable of generating conscious states, since such states are not, evidently, merely CALM products of neural units and their relations. The principles governing the brain's operation cannot be purely combinatorial, or else it *could* not be the basis of consciousness. Thus the correct theory of the psychophysical link must

deal in properties that are not subsumable under the CALM schema. Some kind of projection from the neural basis is apparently what we must assume, but this type of projection cannot be reconstructed from within the CALM framework that controls our thinking. Ultimately, if the Chomskyan speculation is on the right lines, the problem is that in some deep way the brain is not organised after the pattern of language – its properties are not merely those of a system exhibiting combinatorial principles. Because of the basis of our theory-forming capacities we are prone to conceive of the brain on the model set by language (and perception), but this does not suffice to disclose the properties in virtue of which the brain contrives to produce conscious states. Thus theoretical reason is targeted away from what it seeks to understand.

I think what I am suggesting should become clearer if I enter another speculation, geared to a rather different conception of the psychophysical link. To many theorists a double aspect view of the mind-body relation has seemed attractive, where each aspect is not supposed to be a *product* of the other.[20] The difficulty with such a view is that the aspects seem merely juxtaposed; no intelligible relation binds them together. Now suppose we take seriously the idea that our sense of intelligibility derives from language, specifically from its combinatorial features. Then we might hypothesise that the notion of a semantic whole governs the way we conceive of joinings of other kinds – at least by analogy and extension. For instance, the subject-predicate relation is the kind of joining of elements that we find maximally transparent; so we yearn for joinings that approximate to this. The hooking together of pieces of matter is not so far off, so we do not fret unduly about that. But the concatenation of the mental and the physical looks far too much like a mere ungrammatical string, exhibiting no inner coherence: we cannot make it fit, even by analogical extension, the paradigm of a well-formed unitary sentence. Thus we are perplexed that such a juxtaposition should obtain at all, especially when we have good indirect reason to believe that it cannot be merely accidental. Our sense of intelligible linkage is set by the linguistic basis of our theory-forming capacities – that is, by the principles of syntactic and semantic combination – but the kind of linkage that connects mental and physical states cannot be subsumed under this paradigm. It is as if we cannot locate the argument-places of

the function that would connect the mental to the physical. We cannot grasp what kind of unitary whole the mental and the physical combine to generate. Thus our language-based sense of intelligibility lets us down in the present case.

I might summarise our cognitive predicament by distinguishing two types of novelty that the world may contain. Type 1 novelty is the kind subsumable under the CALM schema: it applies to linguistic novelty and to the sort of novelty that results when material particles are arranged in various ways. It is fundamentally combinatorial, iterative, and transparent. Type 2 novelty is the kind that cannot be regarded in this way, and which therefore invites the epithet 'genuine': the emergence of consciousness from the brain is a case in point. It is just when we find ourselves reaching for ideas of type 2 novelty – with the attendant notions of radical emergence, underdetermination and irreducible duality – that we are entering philosophical territory. And a characteristic response to the consequent bafflement is an attempt (always doomed) to construe a type 2 case as really a type 1 case, as with typical domesticating projects. We have a natural drive towards the combinatorial, so we try to assimilate everything to that, frequently distorting reality in the process. We yearn for CALM, even when this is not the appropriate attitude to adopt in the circumstances. When that happens we become fixated on the DIME shape, and philosophy is the outcome.

Simply put: philosophy exists because not everything you are interested in resembles what you say and see.[21]

## V

Up to this point I have been discussing the limits of conscious reason, claiming that this particular human faculty is a poor instrument for the discovery of philosophical truth. But it does not follow from any of this that *any* epistemic system must be so bounded; it does not follow that *no* form of cognition could solve philosophical problems. This is a point that has already been made with respect to imaginary creatures, but I want now to consider some actual forms of representation that might plausibly be said to deal with the problems that elude reason.

In effect, I have been saying that our reflective belief-forming system cannot solve philosophical problems, but it is left open that we might possess some other faculty that is better equipped for the task. Some other human faculty, that is, might not be subject to the constraints that follow from the fact that conscious reason is structured by the prior faculties from which it derives. Such a faculty might, indeed, have been specifically selected to contain the kind of information relevant to answering philosophical questions, so that it is not mere luck if there is a convergence between the output of the faculty and the subject-matter of the relevant questions. So: do we contain anything that has been *built* to encode the kind of information we seek when trying to answer philosophical questions?

I want to suggest, perhaps surprisingly, that there are at least two plausible candidates for human epistemic systems that already contain the data conscious reason cannot reach. Neither is plausibly regarded as a by-product of some other faculty, with the limits attendant upon that; rather, both are expressly designed to represent what reason is not designed to represent. These are: the subconscious self-monitoring representations employed by the brain as it goes about its business; and the information contained in the genetic code. Since the latter is easier to expound in a brief space I shall focus on it. And the basic point is straightforward enough: since, as is commonly supposed, the genes work symbolically, by specifying programmes for generating organisms from the available raw materials, they must contain whatever information is necessary and sufficient for this feat of engineering. So, for example, they must somehow specify the structure and functioning of the heart, and they must supply rules for generating this organ from primitive biological components. The genes are, as it were, unconscious anatomists and physiologists, equipped with the lore pertaining thereto. But what goes for the body also goes for the mind: the genes must also contain the blueprint for constructing organisms with the (biologically based) mental properties those organisms instantiate. They must, then, represent the principles by which mental properties supervene on physical properties. They must, that is, specify instructions adequate for creating conscious states out of matter. And the same holds for other mental attributes: the genes 'know' how to construct organisms with intention-

ality, with personhood, with the capacity to make free choices, with rich systems of knowledge – just as they contain instructions for making organisms that embody innate universal grammar.[22] This requires a grip on the natural principles that constitute these attributes, as well as mastery of the trick of engineering them from living tissue. The genes represent unconsciously what creationists ascribe to the mind of God. And since God has to know the answer to the philosophical problems surrounding these attributes, so too do the genes. In fact, they have known the answers for a very long time, well before we ever formulated the questions.

Clearly this claim depends upon a robust acceptance of the idea that the genetic code constitutes a genuine semantic system, just as our belief system does. The epistemic pluralism thus presupposed comes naturally once modularity and subconscious representation have been admitted – but of course I am well aware that not everyone goes along with this. My point here is just that *if* the representationality of the genes is admitted, *then* we can see that the limits of reason need not be the limits of all human epistemic systems. The genes really need to contain the information required to generate psychological organisms, this being their task in life, but conscious reason is under no particular obligation to recapitulate that achievement. So the genes have philosophical insight built into their very job description. Hence each cell in the human body 'knows' more philosophy than shall ever be accessible to the frontal lobes.[23]

What is galling about this is that conscious knowledge is associated with intellectual pleasure, with the satisfactions of understanding. But the information contained in our genes is not hooked up to our pleasure centres in this way, so we gain no enjoyment form their relative omniscience. Consciousness is what makes knowledge pleasurable, and without it even the most profound insight has no power to scintillate. Still, we can console ourselves with the thought that we can, after all, solve philosophical problems – though with a part of ourselves we cannot reach.[24]

## NOTES

[1] See Noam Chomsky, *Reflections on Language*, London: Pantheon Books, 1975; and *Language and Problems of Knowledge*, Cambridge, MA: MIT Press, 1988. See also, Jerry Fodor, *The Modularity of Mind: An Essay on Faculty Psychology*, Cambridge, MA: MIT Press, 1983.

[2] *Reflections on Language*, pp. 155–6.

[3] Nor will my project have much appeal for those who detect no particular epistemic oddity in philosophical inquiry – for them I have no real *explanandum*. Scientistically minded philosophers will certainly not sympathise with my motivation. You have to feel that the quintessentially philosophical questions have a special profundity or refractoriness.

[4] Only the wilder flights of behaviourist psychology could seem to impose no constraints whatever on the kind of mental profile a creature exemplifies. No doubt this idea of indefinite plasticity and unboundedness carried millenarian connotations, which added to its appeal. The irony, of course, is that a purely S-R organism would be incapable of learning anything of interest (as Chomsky has long stressed). What I am assuming is (in effect) that there is something called human nature – as distinct from dog nature and cat nature and bat nature and gnat nature. And I am extending that conviction into our intellectual parts. (I suspect that the tendency to deny this has part of its roots in the metaphysically absurd notion of the 'bare particular' – the featureless underlying reality that carries its properties extrinsically, like a suit of clothes. The *tabula rasa* image is the bare particular in its mental version.)

[5] The tendency to find, or impose, a sharp cut-off point between human and animal minds fuels the idea of boundlessness for the human case. Once a continuity has been admitted the evident limits of animal minds will suggest a comparable position for human minds. (It is quite amazing that philosophers who pride themselves on their biological naturalism should also wish to draw a *cordon sanitaire* around human cognitive capacity.)

[6] W.V. Quine, *Word and Object*, Cambridge, MA: MIT Press, 1960, chapter 2. (I am not suggesting Quine would see his position in this way.)

[7] Michael Dummett, 'What is a Theory of Meaning? (II)', in *Truth and Meaning*, eds. Gareth Evans and John McDowell, Oxford: Clarendon Press, 1976.

[8] For a tortuous discussion, see Colin McGinn, *Mental Content*, Oxford: Blackwell, 1991.

[9] See Hilary Putnam, *Reason, Truth and History*, Cambridge: Cambridge University Press, 1981, chapter 1.

[10] I notice that intellectuals will readily concede the point for motor capacity but optimistically hold out for intellectual unboundedness; I wonder whether sports people have the opposite prejudice . . .

[11] See Chomsky, *Language and Problems of Knowledge*, chapter 5.

[12] I shall henceforth simply say 'philosophy' to cover the subset of problems I am

interested in; my position is not, of course, that whatever people discuss in *departments* of philosophy is subject to cognitive closure.

[13] Colin McGinn, *Problems in Philosophy: The Limits of Enquiry*, Oxford: Blackwell, 1993. The present paper is best taken as an abbreviated version of that book, designed to focus on the larger forestry; arboreal detail, at least down to trunk morphology, can be found in the longer version. I do not regard this paper as self-sufficient.

[14] As I use the term, a *domesticating* theory is always to some degree distorting or deflationary – an attempt to trim the phenomenon of what is essential to it and what makes it so puzzling to us. But it is part of the TN doctrine to insist that any real phenomenon is subject to some true theory or other: this theory, though, which may be beyond our capacities, is such that *were* one to grasp it one would not have any sense of distortion or deflation – for it would be fully adequate to the phenomenon. Such adequate theories, when they can be grasped, as are the bread and butter of successful science. A domesticating theory-attempt, by contrast, always carries a powerful odour of revisionism.

[15] This remains the case even if, for some deep nomological reason, the only naturally possible way to make intelligence is by using neurons; for that is still a point about mental architecture, not about the objective world one is trying to penetrate. Indeed, strictly speaking, the point holds even if philosophy is uniquely hard for *any* logically possible mind: nothing follows from that fact alone about the queerness of the philosophical subject-matter itself. (It is instructive to consider this question with respect to the infinite.)

[16] See Chomsky, *Language and Problems of Knowledge*, pp. 167ff.

[17] Chomsky comes very close to suggesting this in *Language and Problems of Knowledge*, esp. pp. 183–5. (I am grateful to Carol Rovane for reminding me of this passage.)

[18] General remark: by-product accounts of some particular cognitive attainment are empty unless backed by some specific derivational story. You have to be able to indicate *how* the basic faculty gave rise to the secondary one. It is mere irresponsible hand-waving to suggest, for example, that philosophical knowledge is possible because it is (somehow!) a by-product of 'human intelligence'.

[19] There is much more on this in McGinn, *Problems in Philosophy*, where the CALM conjecture is tested in application to the issues of consciousness, the self, meaning, free will, the a priori, and empirical knowledge.

[20] See Thomas Nagel, *The View from Nowhere*, Oxford: Oxford University Press, 1986, chapter 3.

[21] Of course, I am keenly aware that I have not established such a sweeping claim in this paper. A fuller attempt is made in my *Problems in Philosophy*. There is, as we know, a well-established tradition of making large meta-philosophical announcements without doing all the necessary spade-work; and such exuberance can serve a useful purpose, if taken in the spirit intended.

[22] We need not take the word 'know' literally here; Chomsky's technical term 'cognize' will serve equally well: see Chomsky, *Rules and Representations*, New York: Columbia

University Press, 1980, p. 70. It is hard to see how the thesis of genetic cognizing could provoke dismay in anyone who was already comfortable with unconscious cognizing of other sorts. And, for what is it worth, the current stabs at explaining reference causally or teleologically can be pretty obviously applied to putative terms of the genetic code. The genetic instructions have compliance-conditions which are fulfilled if and only if the embryological process produces an organism of the kind the genes specify.

[23] From a larger perspective, the frontal lobes are mere epistemic parvenus, recently installed to give us the benefits of planning and flexibility; they are not the very origin of all that is representational (though they are wont to arrogate this privilege to themselves). For a discussion of the peculiarities of frontal-lobe representation, see my *Problems in Philosophy*, chapter 8.

[24] I am grateful to Carol Rovane for her comments on an earlier version of this paper, and to Galen Strawson for suggesting the title.

*Department of Philosophy*
*Rutgers, The State University of New Jersey*
*New Brunswick, NJ 08903*
*USA*