

The Sad and Sorry History of Consciousness: being, among other things, a Challenge to the ‘Consciousness-studies Community’

P.M.S. HACKER

1. Consciousness As a Mark of Modernity

The term ‘consciousness’ is a latecomer upon the stage of Western philosophy. The ancients had no such term. *Sunoida*, like its Latin equivalent *conscio*, meant the same as ‘I know together with’ or ‘I am privy, with another, to the knowledge that’. If the prefixes *sun* and *cum* functioned merely as intensifiers, then the verbs meant simply ‘I know well’ or ‘I am well aware that’. Although the ancients did indeed raise questions about the nature of our knowledge of our own perceptions and thought, and introduced the idea of an inner sense, they did not characterize the mind as the domain of consciousness. Aristotelians conceived of the mind as the array of powers that distinguish humanity from the rest of animate nature. The powers of self-movement, of perception and sensation, and of appetite, are shared with other animals. What is distinctive of humanity, and what characterizes the mind, are the powers of the intellect – of reason, and of the rational will. Knowledge of these powers is not obtained by consciousness or introspection, but by observation of their exercise in our engagement with the world around us. The mediaevals followed suit. They likewise lacked any term for consciousness, although they too indulged in reflections upon ‘inner senses’ – in the wake of Avicenna’s distinguishing, arguably to excess, five such senses.

To us, this may seem extraordinary. How could the ancients and mediaevals manage to make sense of human nature and of the nature of the human mind *without* an *explicit* concept of consciousness? After all, is not consciousness the mark of the mental? Is it not consciousness that distinguishes us from mindless nature? Is it not precisely *because* we are conscious that there is something it is like to be us, and that there is not something it is like to be an automaton?

This response is too swift. It presupposes the cogency of the early modern and contemporary *philosophical* conceptions of consciousness. If we attend carefully, we may well hear the ancients in the Elysian fields laughing at us moderns, wondering how *we* can possibly hope to make sense of human nature and of the nature of the human mind with the knotted tangle of misconceptions that *we* have woven into reflections on consciousness. For consciousness, as conceived by early modern and, rather differently, by contemporary, philosophers, is a mark, not of the mental, but of subtle and ramifying confusion. Of course, the laughter of the ancients may be a little wry – for they would have to admit that they had sowed the seeds of confusion. They had done so by their deeply misleading question: ‘How do we know our own perceptions?’ And they had made things worse by their confused answer, namely: that we do so by means of a ‘common, or general, sense’ (*koinê aisthêsis* (Aristotle), subsequently translated into Latin as *sensus communis*) or ‘an internal sense’ (*sensus interior* (Augustine)).

The English word ‘conscious’ is recorded by the OED as first occurring at the beginning of the seventeenth century, when, like the Latin ‘consciūs’, it signified sharing knowledge with another or being witness to something. In its early forms, it occurred in phrases such as ‘being conscious to another’ and ‘being conscious to something’. But sharing knowledge rapidly evolved into being privy to unshared knowledge, either about others or about oneself. So ‘to be conscious to’ quickly became a cousin to the much older expression ‘to be aware of’. The form ‘to be conscious to’ was slowly displaced by ‘to be conscious of’. ‘To be conscious of something’, of course, signified a form of knowledge. So like ‘to know’, ‘to be conscious of something’ is a factive verb – one cannot be conscious of something that does not exist or is not the case. Outside philosophy, there was no suggestion whatsoever that the objects of consciousness, i.e. that of which one can be said to be conscious, are restricted to one’s own mental operations. One could be said to be conscious of *what* one perceived, or of some *feature* of what one perceived, of one’s own or another’s deeds – both good and evil, of a pertinent fact (the lateness of the hour, the merits of a case) and of one’s own or another’s virtues or vices, and so forth. It was not until the middle of the nineteenth century that ‘consciousness’ came to be used to signify wakefulness as opposed to being unconscious. Thenceforth one could speak of losing and regaining consciousness. The common or garden notions of self-consciousness, i.e. either being excessively aware of one’s appearance (a usage now lapsed) or being embarrassingly aware that others are looking at

The Sad and Sorry History of Consciousness

one, is nineteenth-century vintage. Being class-conscious, money-conscious, or safety-conscious are twentieth century coinage.

2. The Early Modern Philosophical Conception of Consciousness

The expression 'conscious' was introduced into philosophy, almost inadvertently, by Descartes.¹ It does not appear in his work prior to the *Meditations* (1641), and even there it occurs just once. In the Third Meditation, it occurs not in relation to knowledge of one's 'thoughts' or 'operations of the mind', but in relation to awareness of the power to perpetuate one's own existence (AT VII, 49; CSM II, 34). It was only under pressure from objectors to this single remark that Descartes was forced, in his 'Replies to Objections', to elaborate his ideas on knowing our own 'thoughts'. His developed position in the *Principles* and late correspondence was unstable. The expression and attendant conception, caught on among Descartes' contemporaries and successors (Gassendi, Arnauld, La Forge) and among English philosophers (Stanley, Tillotson, Cumberland and Cudworth). But it is to Locke, almost fifty years later, that we must turn to find the most influential, fully fledged, *philosophical* concept of consciousness that was to dominate reflection on the nature of the human mind thenceforth. The attendant conception was to come to its baroque culmination (or perhaps nadir of confusion) in the writings of Kant and the post-Kantian German idealists.

Descartes used the terms *conscientia*, *consci*, and *conscio* to signify a form of knowledge, namely the alleged direct knowledge we have of what is passing in our minds. What we are conscious of (which I shall call the 'objects of consciousness') are Thoughts, a term which Descartes stretched to include thinking (as ordinarily understood), sensing or perceiving (shorn of their factive force), understanding, wanting, and imagining. Because he held thinking to be the sole essential attribute of immaterial substances, he claimed that we are thinking all the time, waking or sleeping. He also held that consciousness of operations of the mind is indubitable and infallible. He argued that the mind is, as it were, transparent. For, he wrote (AT VII, 214; CSM II, 150), it is *self-evident* that one cannot have a thought and

¹ It was already used by Bacon, initially in the form 'conscient' (1612), and then in the form 'conscious' (1625) to signify being privy to knowledge about one's faults. But the concept had no role in his philosophy.

not be conscious of it – although the thoughts we have in sleep are immediately forgotten.²

It is noteworthy that his position was equivocal and indecisive. He equivocated between taking consciousness of a thought to be reflective thought about a thought ('Conversation with Burman', CSM III, 335), and elsewhere holding it to be identical with thinking ('Replies to Bourdin', CSM II, 382). A corollary of this was that he equivocated between taking thoughts to be the objects of consciousness, i.e. that *of which* one is conscious, and taking thoughts to be species of consciousness in the sense in which seeing, hearing, smelling are species of perceiving ('Replies to Hobbes', AT VII, 176; CSM II, 124: all acts of thought 'fall under the common concept' of consciousness). Above all, he had no explanation for the possibility of this extraordinary cognitive power, which, unlike *all* our other cognitive powers, is *necessarily exercised upon its objects*³, and is both *infallible* and *indubitable*. Within the confines of one's mind, this cognitive power is, as it were, godlike – omniscient. How can this be? As Thomas Reid later remarked, if one were to ask Descartes how he knew that his consciousness cannot deceive him, he could answer only that 'the constitution of our nature forces this belief upon us irresistibly' (*Essays on the Intellectual Powers of Man*, Essay VI, ch. vii).

Locke, writing almost half a century later, characterized consciousness not epistemically, in terms of indubitability and incorrigibility, but, as La Forge, Malebranche, and Cudworth had done, psychologically, comparing consciousness to an inner sense whereby we perceive that we perceive. 'Consciousness', he explained, 'is the perception of what passes in a Man's own Mind' (*Essay*, II-i-19). We attain knowledge of what passes in our minds by the exercise of an inner sense. We cannot perceive without perceiving that we perceive. Like Descartes, he held that one 'cannot think at any time, waking or sleeping without being sensible of it' (*Essay*, II-i-10). Unlike Descartes, he did not suppose that we must be thinking for the whole of our existence. Unlike Descartes, he did not limit the

² For the Cartesian investigations and reflections I am much indebted to Professor Hanoch Ben-Yami, with whom I spent five enjoyable days hunting together through Adam and Tannery and discussing the findings.

³ It may seem that if acts of thought are species of consciousness, then it is obvious that if one thinks one must be conscious that one thinks, just as if one sees, one necessarily perceives. But that is a mistaken analogy. If one sees a tree, then what one perceives is not *that one sees it*, but *the tree*. However, Descartes requires that the object of consciousness be the act of thinking, not the object of the act of thinking.

The Sad and Sorry History of Consciousness

objects of consciousness to the present or to the operations of the mind, since he held us to be conscious of our past mental operations and of our past and present actions.⁴ Consciousness is the glue that binds together the fleeting perceptions of the mind into one persisting self-consciousness, and is a necessary condition for responsibility for our actions.

The eighteenth-century debate developed from these foundations. Let me summarize, in a *Galtonian picture*, the conception of consciousness that Kant, to his misfortune, inherited from the Cartesian and empiricist tradition. In this tradition, give or take a couple of points, consciousness is

- i. The general form of Operations of the Mind, i.e. one cannot 'think' without being conscious of one's 'thinking';
- ii. An *inner sense* – by the use of which we know how things are subjectively with us;
- iii. Indubitable – one cannot doubt whatever one is conscious of;
- iv. Infallible – one cannot make a mistake about what one is conscious of.

Furthermore

- v. To think one is conscious of something does not differ from being conscious of something. So the mind is, so to speak, *transparent*, and what is in the mind is, as it were, *self-presenting*. So *mind is better known than matter*.

In addition, the objects of consciousness (what one is conscious of) are

- vi. Limited to the operations of the mind;
- vii. Temporally confined to the present;
- viii. privately 'owned' (no one else can have my pains or do my thinking);
- ix. Epistemically private – only I *really* know (because I have privileged access to) the operations of my mind.

Consequently, *the private is better known than the public*. Further,

- x. One's consciousness of what passes in one's mind requires possession of ideas or concepts of mental operations.

⁴ The latter is necessary for Locke because of the link he forged between consciousness, the concepts of a person and personal identity, and the idea of responsibility for one's actions.

These ideas or concepts have no logical relationship to behaviour, since they are applied to objects of inner sense without reference to one's behaviour. To possess them requires no more than consciousness of the ideas (Descartes)⁵, or a private ostensive definition (Locke).⁶ And finally,

- xi. Consciousness of the operations of the mind is *self-consciousness*: i.e. consciousness of how things are with one's self.

Points (viii) to (x) commit the early moderns and their followers to the intelligibility of a logically private language. I shall not discuss this fatal flaw here. Disagreements, which continued well into the nineteenth century, turned largely on the questions of whether (a) there are unconscious operations of the mind; (b) whether inner sense is contemporaneous with, or subsequent to, its objects; and (c) whether consciousness is or is not infallible.

3. Cracks in the Facade

Such was the conception of consciousness and self-consciousness that plagued philosophy in the Cartesian/empiricist tradition. The whole structure turns on two simple and correct thoughts. First, self-ascription of *many* (but not all) psychological attributes is indubitable in the following sense. If one feels a severe pain, one cannot doubt that one is in pain. If one thinks that it is time to go, one cannot doubt that one so thinks. If one is afraid of tomorrow's examination, one cannot doubt that one is thus afraid. Secondly, in *many* cases, one cannot be mistaken. So, for example, one cannot be mistaken that one is in severe pain, or that one thinks that $2 + 2 = 4$ (and that one has not misidentified one's thought with the thought that $2 + 2 = 22$).

⁵ 'Thus it would be pointless trying to define, for someone totally blind, what it is to be white: in order to know what that is, all that is needed is to have one's eyes open and to see white. In the same way, in order to know what doubt and thought are, all one need do is to doubt or to think. That tells us all it is possible to know about them, and explains more about them than even the most precise definitions.' *The Search after Truth* (CSM II, 417f.; AT X, 524).

⁶ 'Such precise, naked appearances in the mind [viz. 'abstract general ideas'], without considering how, whence or with what others they came to be there, the understanding lays up (with names commonly annexed to them) as standards to rank real existences into sorts, as they agree with these patterns, and to denominate them accordingly.' *Essay II*, ix, 9.

The Sad and Sorry History of Consciousness

It is all too easy to follow the Cartesian tradition in supposing that if one cannot doubt things to be so with oneself, and cannot be mistaken, therefore one must know, with complete certainty that they are so. But this seemingly innocuous move is precisely where one goes wrong. For we mistake the impossibility of doubt for the necessary presence of certainty, and the impossibility of mistake for the presence of infallible knowledge.

Doubt needs reasons. The possibility of doubting something may be excluded by realization of the eliminability of all genuine alternatives in the circumstances. Here possible doubt is excluded by the available evidence. In such cases, one may typically be quite certain that things are as one takes them to be. But doubt may also be excluded by purely *logical* or *conceptual* considerations: by the fact that *it makes no sense* to doubt the kind of thing in question, or that *it makes no sense* to doubt in such circumstances. Here doubt is excluded not *de facto*, but *de jure* – because no sense has been given to the words ‘I doubt’ as a prefix to the kind of empirical proposition in question, or in the circumstances in question. To give a few familiar examples of kinds of empirical proposition other than psychological self-ascriptions: it makes no sense to doubt whether one exists (if someone said ‘I am not sure I exist’ or ‘I doubt whether I exist’ we should ask him what on earth he meant); it makes no sense, in normal circumstances, as one walks through a wood of great oak trees, to doubt whether *this* is a tree or *this* is a tree, etc.; (if someone, as he touched each great tree, said ‘I doubt whether this is a tree’, we would think him deranged – or a philosopher). When doubt is excluded *de facto*, then it makes sense to speak of certainty. But when it is logically impossible to doubt – when it makes no sense to doubt, then it equally makes no sense to be certain either. The presence of certainty does indeed exclude all doubt, but if all doubt is *logically* excluded, there is nothing for certainty to exclude. So there is no room for certainty either – the logical space, so to speak, has vanished. Similar considerations apply to the exclusion of mistake. The logical impossibility of a mistake does not imply infallible knowledge, but the exclusion of knowledge *together with error*. So it is with anything that fits the bill for a Cartesian *cogitatio*. There is no logical space for ignorance, and hence too, no logical space for knowledge, no logical space for doubt, and hence too, no logical space for certainty. The utterance ‘I know I am in pain’ is not at all akin to ‘I know he is in pain’, and although I may be certain that he is in pain, I cannot (logically *cannot*) be certain that I am in pain – for there is no possibility of doubt that might be excluded by certainty.

Why do we cleave so adamantly to the idea that we *know* with certainty that things are so with us? Because it is altogether natural to feel that if we *don't* know, then we must be *ignorant* of what we are being said not to know. And for sure, when one is in severe pain, one *is not ignorant* that one is in pain. Indeed! – But it does not follow that one knows (with certainty) that one is. It follows that one neither knows *nor is ignorant*. It is not that we don't know that things are thus-and-so with us – it is that there is no such thing as *not knowing* in these cases. But by the same token, *there is no such thing as knowing either*. The truth of the matter is that being mature language users, we can – in all the cases relevant to the early modern debate on consciousness – *say* how things are with us. Our saying so is constitutive (not inductive) evidence *for others*, for things being thus-and-so with us. Our sincere word therefore has a privileged status *for others*. Such constitutive evidence is defeasible, but if not defeated, *it stands firm*. But this does not show that we *know* that things are as we say they are – for there is *no such work* for the verb 'know' to do here (which does not mean that it cannot do *other* work here). It shows only that ignorance, *together with knowledge*, is here logically excluded.

Of course, if we assume, with the early modern tradition, that we know with certainty how things are ('subjectively') with us, then it is all too natural to ask *how* we know. Then we are strongly tempted to suppose that we do so by the exercise of a cognitive faculty. Moreover, since we can *say* how things are thus with us without any evidence, it is almost irresistible to suppose that this cognitive faculty is a form of perception – since to learn how things are by directly perceiving how they are involves no evidence either. So it seems that we know how things are with us *in foro interno* by means of an inner sense, which we then dub 'apperception', or 'introspection'. As William James put it so wrongly in 1890, introspection 'means, of course, the looking into one's own mind and reporting there what we discover' (*Principles of Psychology*, I, 185). It is by the use of this inner sense, it seems, that we perceive, or apperceive, or become conscious, of how things are with us. This inner sense is just like an outer sense, only

- (i) Without a sense organ,
- (ii) Its successful exercise is independent of observation conditions,
- (iii) It never fails us, but always yields knowledge,
- ∴(iv) We know the mind better than the material world (cp. Descartes, Brentano, Husserl).

The Sad and Sorry History of Consciousness

But there is no such thing as a cognitive faculty that is miraculously immune to error, and no such things as a faculty of perception that enables us to perceive without any organ of perception and the successful exercise of which is independent of circumstances of observation. ‘To perceive’, as well as ‘to see’, ‘to hear’, etc. have a use as success verbs – but there is no such thing as succeeding if there is no logical possibility of failing. (It is, of course, noteworthy that ‘to be conscious of’ is not a *success verb* – one cannot try to become or succeed in being conscious of something – although it is a *factive verb*, since what one is conscious of being so is so.)

There is indeed such a thing as introspection – but, *pace* James, it is not a form of perception and involves no ‘looking into’ one’s mind. It is a form of self-reflection, at which some people, like Proust, are better than others. It involves reflecting on one’s actions and character traits, on one’s springs of action, likes and dislikes. It is a route to self-knowledge, but also a highroad to self-deception. It is not exercised when one says that one has a headache or that one is thinking of going to London tomorrow. That a child has learnt to say ‘Mummy, my head aches’ does not show that he is becoming introspective. Nor does it show an advance in self-knowledge.

What is true is that if we are asked whether we are in pain, whether we want this or that, whether we are thinking things to be so, or thinking of something or other, we *can say so*. It is characteristic of Locke and his successors down to James, Brentano and Husserl, to confuse the ability to *say* how things are with one with the ability to *see* (by *introspection*) how things are with one. To be sure, when a human being who has mastered the use of language, has a pain, he can say so. If asked whether he is in pain, he can reply. It is tempting to think that he can say that he has a pain in his foot, because he feels, i.e. perceives the pain. But to feel pain is not a form of perception. To feel a pain in one’s foot, for one’s foot to hurt, just *is* to *have a pain* – not to have a pain and in addition to perceive it. Truthfully to say ‘My foot hurts’ is no more an expression of something one has *perceived*, *learnt* or *come to know* than is a groan of pain. Of course, one is not *ignorant* of one’s foot’s hurting either. Can one intelligibly *say* ‘I know I have a pain’? In appropriate circumstances, of course. But all it means is that I *really do* have a pain, that *it is true* that I have a pain. It does not mean that I have evidence for it, nor does it mean that I perceive it directly.

A language-user can say what he is thinking. If asked ‘A penny for your thoughts?’, he can reply. So how does he know *that* he is thinking? Is it not by introspection? No. – Let us first ask how he knows *what* he thinks. Well, he has weighed the evidence, and decided

that the weight of evidence is in favour of things being thus-and-so; so he says that things are so – that is what he has concluded is the case. If he regards it is a matter of opinion, or if he regards the evidence as not being decisive, he will affix an ‘I think’ to the sentence to indicate just that. So he says that he *thinks* things to be thus-and-so. ‘I think’ functions *here* as a qualifier signifying not a mental operation currently taking place, but as an indicator for others of the epistemic weight of the proposition to which it is affixed.⁷

Yes, but surely he *knows that he thinks what he thinks!* After all, do we not sometimes say ‘I don’t know what I think’? And if ‘I don’t know what I think’ makes sense, then surely its negation ‘I do know what I think’ makes sense too! – It is true that we sometimes say ‘I don’t know *what* I think’. But not to know what one thinks is not: to think something and not to know what it is. If I don’t know what I think about something or other, what I do is *not* ‘peer into my mind’ to find out. Rather, what I do is examine the evidence pertinent to the matter at hand, and make up my mind on the balance of evidence. ‘I don’t know what I think’ is an expression of inability to judge (‘I can’t make up my mind’, we say) – not of an introspective deficiency. It is a confession of not knowing *what to think*, which can be remedied only by looking again at the evidence.

All right; but still, we often proclaim that we don’t know what we want. And here surely what we don’t know is an operation of the mind! Don’t we then quickly introspect and then say ‘Now I know what I want’? – No. On the contrary: ‘I don’t know what I want’ signifies *inability to decide* between desiderata. And finding out what one wants is not a matter of introspectively running over one’s various desires, but rather of reflecting on the desirability characteristics of the available alternatives and choosing the most preferable. ‘Now I know what I want!’ means much the same as ‘Now I have decided’.⁸

Now, let us to return to the ancients and their confused question: When we see something or see something to be so, *how do we know that we do?* Do we perceive our seeing by sight? Or do we perceive

⁷ Of course, there are other uses of this verb. For detailed discussions, see G. Ryle, *On Thinking* (Oxford, Blackwell, 1979), A. R. White, *The Philosophy of Mind* (New York, Random House, 1969), B. Rundle, *Mind in Action* (Oxford, Clarendon Press, 1997), P. M. S. Hacker, *Wittgenstein: Meaning and Mind* (Oxford, Blackwell, 1990).

⁸ One might, provocatively, say that these uses of ‘I know’ are non-epistemic, in the sense in which ‘While you were with me, I forgot all my troubles’ is not an epistemic use of ‘forget’ – it does not signify a failure of memory. So too, ‘I know that I am in pain’ or ‘I know that I intend to go’ do not signify the successful exercise of a cognitive faculty.

The Sad and Sorry History of Consciousness

our seeing by a common or general sense? – Neither. There is no such thing as confusing seeing with hearing or tasting. If someone were to say ‘I think there is a sound coming from the bush, but I am not sure whether I see it or taste it’, we would not know what he meant. We exercise our senses and use our sense-organs in making judgements about things in our vicinity. According to the sense-qualities we apprehend, and to the sense organs we employ, we can affix an ‘I can see . . .’, ‘I hear a . . .’, ‘I can smell . . .’ to the expression of one’s perceptual judgement. These prefixes indicate the sense-faculty and sense-organ by the use of which one takes oneself to have acquired information. There is no such thing as *mistaking* sight for smell, or hearing for tasting.⁹ And if there is no room for *error*, and if there are no evidential grounds for saying ‘I see a so-and-so’ or ‘I heard a sound from over there’, then the question ‘How do you know that you see (rather than hear or taste) something or other?’ is to be *rejected*, not answered. One does not perceive that one perceives. Nor is one conscious *that one perceives*,¹⁰ although one may be *conscious of what one perceives* – if it catches and holds one’s attention. One can *say* what one perceives – but to be able to say *what one perceives* is not to *perceive that one perceives*. Roughly speaking, it is not that the ‘I think’ must accompany all my representations, as Descartes and Locke supposed. Nor is it even that it must be *possible* for the ‘I think’ to accompany all my representations, as Kant suggested. Rather, it must be possible for the ‘I say’ to accompany all my representations. Or, more perspicuously, it must be possible for me to say how things are ‘subjectively’ with me. And since I can say how things are thus with me, I can also reflect on things being so with me – which is something non language-using animals cannot do. But to reflect on things being thus-and-so with me is not the same as being conscious of thing’s being thus-and-so, any more than reflecting on Julius Caesar’s assassination is to be conscious of it.

In brief, consciousness is not an inner sense, and it is not a faculty for knowledge of the ‘inner’. Roughly speaking, anything that

⁹ Synaesthesia does not exemplify such an error, for the person who suffers from synaesthesia does not claim to hear the colours of objects, but vividly to associate sounds with colours. He does not shut his eyes and hear the colours of the flowers – indeed, there is no such thing. But when he sees the colours of the flowers, he associates sounds with them.

¹⁰ Blindsight is not an exception to this conceptual truth. It is a confusion to suppose that the blindsighted see, but are not conscious of seeing. For detailed discussion, see J. Hyman, ‘Visual experience and blindsight’ in *Investigating Psychology: Sciences of the Mind after Wittgenstein* (London, Routledge, 1991), pp. 166–200.

Descartes might, *with good reason*, wish to cite as an indubitably and infallibly known *thought (cogitatio)*, everything ‘inner’ for which *truthfulness guarantees truth*, is something of which one *cannot* oneself be either ignorant or doubtful. By the very token of the *cannot*, one cannot know or be certain about it either. Consciousness, conceived as an inner sense with operations of the mind as its objects, is not a mark of the mental, but of thoroughgoing confusion.

4. The Contemporary Philosophical Concept of Consciousness

The concept of consciousness as moulded by the early moderns plagued philosophy well into the twentieth-century. However, it did not attract much interest among most early analytical philosophers (Moore is an exception), in the Vienna Circle, or among the dominant figures in the first decades of post-war analytic philosophy. This was due partly to the rise of behaviourism, partly to a decline of interest in philosophy of mind among analytic philosophers in the inter-war years, and partly to the post-war criticisms, launched by Wittgenstein and Ryle, of the early modern conception of the mind, of mental operations and of the relationship between mind and body. The subject of consciousness was awakened from its slumbers in the 1970s as a response to functionalism.

In a seminal article ‘What is it like to be a bat?’ (*Phil. Review* 1974), Thomas Nagel laid the groundwork for the next forty years of fresh confusion about consciousness. Nagel defended three salient theses:

1. *An experience is a conscious experience if and only if there is something it is like for the subject of the experience to have that very experience.*

What it is like for an organism to have a given experience is denominated ‘the subjective character (or quality) of experience’. And this supposed consciousness – the ‘what-it’s-likeness’ (as it is now called¹¹) – of a given experience is dubbed ‘phenomenal consciousness’.

2. *A creature is conscious or has conscious experience if and only if there is something it is like for the creature to be the creature it is.*

¹¹ See *The Oxford Companion to Consciousness* (Oxford, Oxford University Press, 2009), p. 665.

The Sad and Sorry History of Consciousness

So, we all know that there is something which it is like for us to be human beings – although it is very difficult to say what it is like. On the other hand, no one (other than a bat) can even imagine what it is like to be a bat.

3. *The subjective character of the mental can be apprehended only from the point of view of the subject.*¹²

Some clarification and elaboration is needed:

- (a) Just as Descartes (and his successors) misguidedly extended the notion of Thought to include *seeming to perceive* in all its modes, *imagining*, and *wanting something*, so the new conception of Conscious Experience was misguidedly extended to include *thinking*¹³ – which is no more an ‘experience’ than wanting is a species of thought.
- (b) Each conscious experience was in due course argued to have its own qualitative character – its distinctive *phenomenal feel*.¹⁴ The individual feel of an experience was dubbed a *quale*.¹⁵ ‘The problem of explaining these phenomenal qualities’, Chalmers later declared,¹⁶ ‘is just the problem of explaining consciousness.’ For what characterizes *any* conscious experience are the distinctive *qualia* that accompany it.
- (c) It is of capital importance to realize that Nagel’s claim that ‘there is something which it is like to have a given conscious experience’ is not a statement of *similarity*. That is, to ask: ‘What is it like to walk fast?’ is not a variant upon ‘What is walking fast like, what does it resemble?’ It is not to be answered by a comparison, such as ‘Rather like running,

¹² This thesis is sketched in Nagel’s ‘Subjective and Objective’ (repr. in his *Mortal Questions* (Cambridge University Press, Cambridge, 1979), and further developed in his book *A View from Nowhere*. For critical discussion of this misbegotten notion of ‘point of view’, see Bennett and Hacker, *Philosophical Foundations of Neuroscience*, 11.2.

¹³ See J. R. Searle, *Mysteries of Consciousness* (London, Granta Books, 1997), p. 201.

¹⁴ The notion of ‘raw feels’, subservient to a very similar muddled thought, was introduced much earlier by the behaviourist psychologist E. C. Tolman in his *Purposive Behaviour in Animals and Men* (New York, Appleton-Century-Crofts, 1932).

¹⁵ The term was borrowed from C. I. Lewis, *Mind and the World Order* (1929) (New York, Dover, 1956).

¹⁶ D. Chalmers, *The Nature of Consciousness* (Oxford, Oxford University Press, 1996), p. ix.

only one foot is always on the ground'. The question is not: What does it resemble? It concerns the subjective qualitative feel of the experience: what it feels like *for the subject*.

This novel analysis of consciousness, this attempt to save us from reductive physicalism or soulless functionalism, caught on like hot cross buns. It also made it possible for philosophers to hang on to the white coat-tails of cognitive neuroscientists. Consciousness studies became the all the rage.¹⁷ Conferences proliferated, new journals were founded, a stream of articles and books on consciousness rapidly turned to a flood. A common article of faith among the self-styled 'consciousness studies community'¹⁸ is that consciousness is essentially (some would grandly say 'metaphysically') characterized by reference to there being something that it is like to be a conscious creature, and that experience or 'phenomenal consciousness' is to be explained by reference to the fact that there is something that it is like to have it.

Once one has gone down this *cul-de-sac*, then a flood of apparently deep problems follow. What is consciousness for? What is the evolutionary advantage of consciousness? Why aren't there any 'zombies'? How can such a strange phenomenon as consciousness emerge from mere matter? How can one bridge the 'explanatory gap'¹⁹ between neural activity and conscious experience? And so forth. I shall not try to answer these misconceived questions here –

¹⁷ Apparently a Google Scholar search in 2006 yielded over 600,000 books and articles with the word 'consciousness' in its title.

¹⁸ See M. Velmans and S. Schneider, *The Blackwell Companion to Consciousness*, Introduction, p. 1.

¹⁹ Curiously, this muddled idea is ascribed, as a grand new insight, to contemporary members of the consciousness studies community, in particular to J. Levine. But it is at least as old as Leibniz (see note 22 below), and was beautifully stated in the nineteenth century by Huxley and Tyndall. Huxley exclaimed 'How it is that anything so remarkable as a state of consciousness comes about as a result of irritating nervous tissue, is just as unaccountable as the appearance of Djin when Aladdin rubbed his lamp' (*Lessons in Elementary Psychology* (1866), p. 210). Tyndall remarked 'The passage from the physics of the brain to the corresponding facts of consciousness is unthinkable. Granted that a definite thought and a definite molecular action in the brain occur simultaneously, we do not possess an intellectual organ, nor apparently any rudiment of an organ, which would enable us to pass by a process of reasoning, from one to the other' (*Fragments of Science*, 5th ed. p. 420). It is striking that similar despair has been expressed in recent years by C. McGinn, who inferred, from the fact that *he* could not answer the question, that it is beyond the powers of the human mind to do

The Sad and Sorry History of Consciousness

I have already done so, together with my colleague Max Bennett, in *Philosophical Foundations of Neuroscience*.²⁰ What I should like to do is to make clear why the contemporary philosophical conception of consciousness that is embraced by the ‘consciousness studies community’ is incoherent – and to throw down a gage to members of that ‘community’.

5. A Challenge to the Consciousness Studies Community

Why is it evidently so tempting to agree to this analysis of consciousness? I believe that four factors are in play. First is the persuasiveness of the claim that, as Davies and Humphreys declared, there isn’t anything which it is like to be a brick, or an ink-jet printer, but ‘there is, presumably, something it is like to be a bat or a dolphin and there is certainly something it is like to be a human being.’²¹ For initially one is inclined to agree to this misconceived rhetorical statement. After all, you can ask someone what it was like for him to be a soldier, and you cannot ask an ink-jet printer anything. The second factor to benumb our linguistic sensibility is the relative unfamiliarity of the phrase ‘there is something which it is like to’, which involves second-level quantification over properties coupled with an unrecognized misuse of the interrogative phrase ‘what is it like’. The third operative factor is the appeal of the idea of ‘saving our humanity’ – of providing a bulwark against the rising tides of reductionism and functionalism.²² Finally, the appeal of mysteries, of facing the

so. For detailed critical scrutiny, see M. R. Bennett and P. M. S. Hacker, *Philosophical Foundations of Neuroscience* (Oxford, Blackwell, 2003), 11.3.

²⁰ Of course, I am not suggesting that there are not numerous empirical problems about the various forms of consciousness. We should like to understand, not what consciousness is for, but rather what sleep is for. It is of interest to know the neural mechanisms involved in perceptual consciousness (i.e. of having one’s attention caught by something in one’s field of perception). It is important to discover how the brain maintains intransitive consciousness. And so on. My point is merely that the so-called ‘hard problem’ of consciousness, and the battery of related questions often cited by philosophers are merely conceptual confusions masquerading as empirical questions.

²¹ M. Davies and G. W. Humphries, eds. *Consciousness* (Oxford, Blackwell, 1993), p. 9.

²² Leibniz nicely observed: ‘supposing that there were a machine so constructed as to think, feel and have perception, we could conceive of it

deepest and most difficult problem known to man, of being at the Last Frontier of knowledge, is well-nigh irresistible. But in philosophy, *there are no mysteries* – only mystifications and mystery-mongering.

I believe that the temptation must be resisted, and sober analysis should take its place. I shall, very briefly, defend three antitheses.²³

- (1) Experiences are not in general individuated by reference to what it feels like to have them but by reference to what they are experiences of. Most experiences have no qualitative character whatsoever – they are qualitatively neutral.
- (2) There is not *something which it is like* to have an experience.
- (3) There is not *something which it is like* to be a human being or, for that matter, a bat.

Let me explain.

1a. It is true that being in severe pain is awful, that smelling the scent of roses is pleasant, that the sight of mutilated bodies is horrifying. These are qualitative characters of certain experiences.

1b. Every experience is a *possible* subject of attitudinal predicates, e.g. of being pleasant or unpleasant, interesting or boring, attractive or repulsive. But it is false that every experience is *an actual* subject of such an attitudinal predicate. With respect to most experiences the question ‘What did it feel like to . . . ?’ or ‘What was it like to . . . ?’ is correctly answered by ‘It did not feel like anything in particular’ and ‘It was altogether indifferent’. To see the lamp posts as one walks down the street or to hear the chatter in the class room feels neither pleasant nor unpleasant, and is neither repulsive nor attractive.

1c. Experiences, which may indeed be the subject of the same attitudinal predicate, are not essentially distinguished by reference to it,

as enlarged and yet preserving the same proportions, so that we might enter into it as into a mill. And this granted, we should only find on visiting it, pieces which push one against another, but never anything by which to explain a perception’ (*Monadology*, §17). Here is the mystery and irreducibility of consciousness. It can be updated by replacing ‘pieces that push’ with ‘neurons that fire’. The confusion remains the same.

²³ For more detailed treatment, see P. M. S. Hacker ‘Is there anything it is like to be a bat?’ in *Philosophy* 77 (2002), pp. 157–74, and M. R. Bennett and P. M. S. Hacker, *Philosophical Foundations of Neuroscience* (Oxford, Blackwell, 2003) pp. 237–351. Space prevents a discussion of so called subjectivity – a theme riven with incoherence. It is anatomized in *Philosophical Foundations*, pp. 294–302. I shall use the term ‘experience’ in the broad and ill-defined sense in which it is currently employed by students of consciousness.

The Sad and Sorry History of Consciousness

but by their object. Smelling lilac may be just as pleasant as smelling roses, but the experiences differ despite sharing the same qualitative character. What distinguishes the experiences is not what it feels like to have them, but what they are experiences *of*.

1d. A persistent mistake among defenders of *qualia* is to confuse and conflate the qualities of what one experiences (e.g. the colour of the violets, the scent of the roses, the taste of the apple) with the qualities of the experiences (delightful, enjoyable, pleasant, revolting). *A perceptible quality is not a quality of a perception.* The colours of visibilia are not qualities of seeing them, but qualities of what one sees. The seeing of a red rose is not red, and the hearing of a bang is not loud, although it may be frightening.

1e. It is altogether misguided to stretch the term ‘experience’ to include thinking. But be that as it may, what differentiates thinking that $2 + 2 = 4$ from thinking that $3 + 3 = 6$ is not what it feels like to think thus but rather is *what is thought*. Even if, as Chalmers might suggest,²⁴ a binary whiff is associated with $2 + 2 = 4$, and a tertiary whiff with $3 + 3 = 6$, that is not what individuates the thinkings, as is obvious when one remembers that the tertiary whiff might become associated with the thought that $3 \times 3 = 9$. Or is the first whiff an additional whiff and the second a multiplicative whiff?

2. It is true that one can ask someone: ‘What was it like for you to V?’ (where ‘V’ signifies an ‘experience’). Remember that this is not a request for a comparison, but for a description of the *felt character* of the experience. One may answer: ‘It was quite agreeable (unpleasant, charming, repulsive, fascinating, boring) to V’. Then, if we wish to indulge in second-level quantification, we may say ‘There was something that it *was* for A (or for me) to V, namely: *quite agreeable* (unpleasant, charming, etc.)’. What we *cannot* intelligibly say is: ‘There was something *it was like* for A to V, namely quite agreeable’. That is, existential generalization requires the dropping of the ‘like’ – for the experience was not *like quite agreeable*, it *was* quite agreeable. This should be obvious from consideration of the answer to the question: ‘What was it like for you to V?’ For the answer (save among the illiterati) is not ‘To V was like wonderful’, but ‘To V was wonderful’. And the existential generalization of that cannot yield the form ‘There is something which it is like to V, namely wonderful’. The latter aberration is the result of a miscegenous crossing of the

²⁴ Chalmers claims that his ‘experience’ of thinking about lions has a leonine whiff about it, so ‘what it is like to think of a lion is subtly different from what it is like to think of the Eiffel tower’ (see *The Conscious Mind*, p. 10).

existential generalization of a judgement of similarity with an existential generalization of a judgement of the affective character of an experience. The result is latent nonsense – which has now been rendered patent.

So, (i) it is simply ill-formed nonsense to suggest that a conscious experience is an experience such that *there is something it is like to have it*. (ii) Most experiences are qualitatively (affectively) characterless – they have no ‘qualitative (attitudinal) character’ at all. (If anyone were to ask us such questions as ‘What is it like to see the buttons on my shirt?’, ‘What is it like to hear Jack say “and”?’ or ‘What is it like to feel the arm of the armchair?’, we should be very puzzled at the questions, since such perceptual experiences are obviously qualitatively neutral in normal circumstances.)

Let us now turn to the third antithesis. It makes perfectly good sense to ask ‘What is it like to be a doctor (a mother, an old-age pensioner, ill)?’. This is a request for a description of the pros and cons of a certain social role, or of being a V-er or of being in a certain condition. Such questions demand a specification of the qualitative character of the life of an X or the typical career of a V-er or of the condition of being F. That is precisely why this form of words was misguidedly chosen by modern consciousness students to explain what it is to be a conscious creature. Hence the statement: ‘there is, presumably, something it is like to be a bat or a dolphin and there is certainly something it is like to be a human being.’ But this statement is quite mistaken.

3a. Let me explain why, from the point of view of English grammar and of the devices of second-level quantification, there isn’t anything it is like to be a bat, or to be a dolphin, and there certainly isn’t anything it is like to be human. Sometimes there is no need, in a question of the form ‘What is it like to be an X?’, to specify the subject class, i.e. to specify what it is like *for whom* to be an X. For it is often evident from the context. ‘What is it like to be a doctor?’ is restricted to adult human beings, ‘What is it like to be a mother?’ to women. But sometimes it *is* necessary, e.g. ‘What is it like *for a woman* (as opposed to a man) to be a soldier?’ or ‘What is it like *for a teenager* (as opposed to someone older) to be the champion at Wimbledon?’ And often the question is personal, as in ‘What was it like *for you* to be a soldier in the Second World War?’

As in the previous cases of ‘What is it like to V?’, so too here the ‘like’ drops out in existential generalization. If you answer the question ‘What is it like for a teenager to win at Wimbledon?’ by saying ‘It is quite overwhelming’, then the existential generalization is not ‘There is something which it *is like* for a teenager . . .’, but rather

The Sad and Sorry History of Consciousness

‘There is something that it *is* for a teenager to win at Wimbledon, namely: quite overwhelming’. But this ineradicable flaw is not the worst of the ensuing nonsense.

3b. We can licitly ask ‘What is it like for a Y – for a man, a woman, a soldier, a sailor, etc. – to be an X?’ We can also licitly ask ‘What is it like for you to be an X?’ Note the general form of these questions. (i) The subject term ‘Y’ differs from the object term ‘X’. (ii) Where the subject term is specified by a phrase of the form ‘for a Y’, then a principle of contrast is involved. We ask what it is like *for a Y*, as opposed to *a Z*, to be an X. (iii) There is a second principle of contrast involved in questions of the form ‘What is it like for a Y to be an X?’, namely with regard to the ‘X’. For we want to know what it is like for a Y *to be an X*, as opposed to *being a Z*.

But the form of words that we are being offered by the consciousness studies community is ‘What is it like for an X to be an X?’ The subject term is reiterated. But questions of the form: ‘What is it like for a doctor to be a doctor?’ are awry. One cannot ask ‘What is it like for a doctor to be a doctor *as opposed* to someone else who is not a doctor being a doctor?’ for that makes no sense. Someone who is *not* a doctor cannot also *be* a doctor – although he may *become* one. The interpolated phrase ‘for a doctor’ is illicit here, and adds nothing to the simpler question ‘What is it like to be a doctor?’ – which is a simple request for a description of the role, hardships and satisfactions, typical experiences and episodes in the life of a doctor. *A fortiori*, questions such as ‘What is it like for a human being to be a human being?’, ‘What is it like for a bat to be a bat?’ and ‘What is it like for me to be me?’ are nonsense. For, they violate the condition of non-reiteration, and they transgress the two contrast principles. Gods and avatars apart, nothing other than a human being can be a human being; a human being cannot be anything other than a human being, for if a human being ceases to be a human being he thereby ceases to exist;²⁵ and it makes no sense to suppose that I might be someone else or that someone else might be me. So the pivotal question ‘What is it like for a human being to be a human being (or ‘for a bat to be a bat’)?’ collapses into the question ‘What it is like to be a human being (or ‘to be a bat’)?’ But now it is not clear what *this* question means — unless it amounts to no more than ‘What is human life like?’ If that *is* what it means – then although it is nebulous, there is no difficulty in answering it, e.g. ‘Nasty, brutish and short’ or ‘Full of hope and fear’. Nor is there any

²⁵ When Kafka turned Gregor Samsa into a beetle, which was Gregor Samsa, he transgressed the bounds of sense.

difficulty in answering the question ‘What is the life of a bat like?’ – any decent zoologist who studies bats can readily tell us. It is even more glaringly obvious that the supposition that there is something it is like for me to be me is sheer nonsense, for it is logically impossible (there is no such thing) for me to be anyone other than myself. Not only do *I* not know what it is like for me to be me – *there is nothing to know*. I do not know what it is like for me to be a human being either – for this is a form of words without any sense. It is not as if I might have been a cat or a dog, or a ink-jet printer. But I can, of course, tell you what my life has been like.

So, does anything come out of the mystification? Well, yes. What comes out is the following. One can ask a human being what it is like for him to fulfil the various roles he fulfils or to do the various things he does – and he can normally tell one. One cannot ask a brick what it is like for it to fill a hole in the wall or an ink-jet printer what it is like for it to run off twenty copies of one’s paper. *For only sentient creatures have roles and have experiences, enjoying some, disliking others, and being indifferent to most.* – A meagre result for so much noise.

As far as I can see, these arguments are watertight. If that is correct, then the larger part of the multitudinous *philosophical* writings of the consciousness studies community, and a considerable number of neuroscientific writings, rest on fundamental conceptual confusions. So I herewith issue a formal challenge to the consciousness studies community: either show that these arguments are flawed, or retire from the field, admitting that consciousness studies *as the members of this community represent them* are sheer nonsense.²⁶

St John’s College, University of Oxford
peter.hacker@sjc.ox.ac.uk

²⁶ I am grateful to Hanoch Ben-Yami, Hanjo Glock, Hans Oberdiek and Herman Philipse for their comments on earlier drafts of this paper.