

Prolegomena

COPYRIGHTED MATERIAL

1

Consciousness as the Mark of the Mental

1. Consciousness as a mark of modernity

*The ancients did not
characterize the mind in
terms of consciousness*

Although the ancients raised questions about our own knowledge of our perceptions and thought, and introduced the idea of an inner sense, they had no word for consciousness and 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 observing their exercise in our engagement with the world around us. The medievals followed suit. They too lacked a term for consciousness, but they likewise indulged in reflection upon ‘inner senses’, arguably – in the wake of Avicenna’s distinguishing five such senses – to excess.

*Descartes’s introduction
of the term and
redefinition of the mind*

Descartes’s innovations with regard to the uses in philosophy of the Latin ‘conscientia’ (which had not hitherto signified consciousness at all), as well as the French ‘la conscience’, were of

capital importance.¹ For it was he who introduced the novel use of the term into the philosophical vocabulary. He invoked it in order to account for the indubitable and infallible knowledge which he held we have of our Thoughts (*cogitationes*) or Operations of the Mind. His reflections reshaped our conception of the mind and redrew the boundaries of the mental. Thenceforth consciousness, as opposed to intellect and sensitivity to reasons in thought, affection, intention and action, was treated as *the mark of the mental* and *the characteristic of the mind*.

The expressions ‘conscius’ and the French ‘conscient’, and the attendant conception of consciousness, caught on among his correspondents and successors (Gassendi, Arnauld, La Forge, Malebranche). So too ‘consciousness’ and ‘conscious’ caught on among English philosophers, churchmen and scientists (Stanley, Tillotson, Cumberland, Cudworth and Boyle). But it is to Locke that we must turn to find the most influential, fully fledged, *philosophical* conception of consciousness that, with some variations, was to dominate reflection on the nature of the human mind thenceforth. This conception was to come to its baroque culmination in the writings of Kant. In the Lockean tradition, consciousness is an *inner sense*. Unlike outer sense, it is indubitable and infallible. It is limited in its objects to the operations of the mind. The objects of consciousness are private to each subject of experience and thought. What one is thus conscious of in inner sense constitutes the subjective foundation of empirical knowledge. Because consciousness is thus confined to one’s own mental operations, it was conceived to be equivalent to self-consciousness – understood as knowledge of how things are ‘subjectively’ (‘privately’, *in foro interno*) with one’s self.

Development of the ordinary use The ordinary use of the English noun ‘consciousness’ and its cognates originates in the early seventeenth century, a mere three or four decades prior to the Cartesian introduction of a novel sense of ‘conscius’ and ‘conscient’ into philosophy in the 1640s. So it evolved side by side with the philosophical use – but, on the whole, in fortunate independence of it. For the ordinary use developed, over the next three centuries, into a valuable if specialized instrument in our toolkit of cognitive concepts. By contrast, as we shall see, philosophical usage sank deeper and deeper into quagmires of confusion and incoherence from which it has not recovered to this day.

¹ French to this day has only ‘la conscience’ to do the work of the distinct English nouns ‘consciousness’ and ‘conscience’.

Multiple centres of variation The ordinary use of ‘conscious’ evolved a number of related *centres of variation*: being conscious as opposed to unconscious; being perceptually conscious of something, or of some aspect of something, in one’s environment; being conscious of one’s feelings and inclinations; being conscious *that* as well as being conscious *of*; *conscious*, as opposed to *unconscious* mental attributes (such as belief or desire); *consciously doing* something *qua* agent, as well as *being conscious of doing* something *qua* spectator; and being *self-conscious*. These are not related as species to a genus. Nor are they different *senses* of ‘consciousness’, if that suggests that they are mere homonyms. Nor is consciousness an Aristotelian ‘focal concept’ (like *healthy*). Rather, there are multiple centres of variation, with various forms of connection between them (see fig. 1.1).

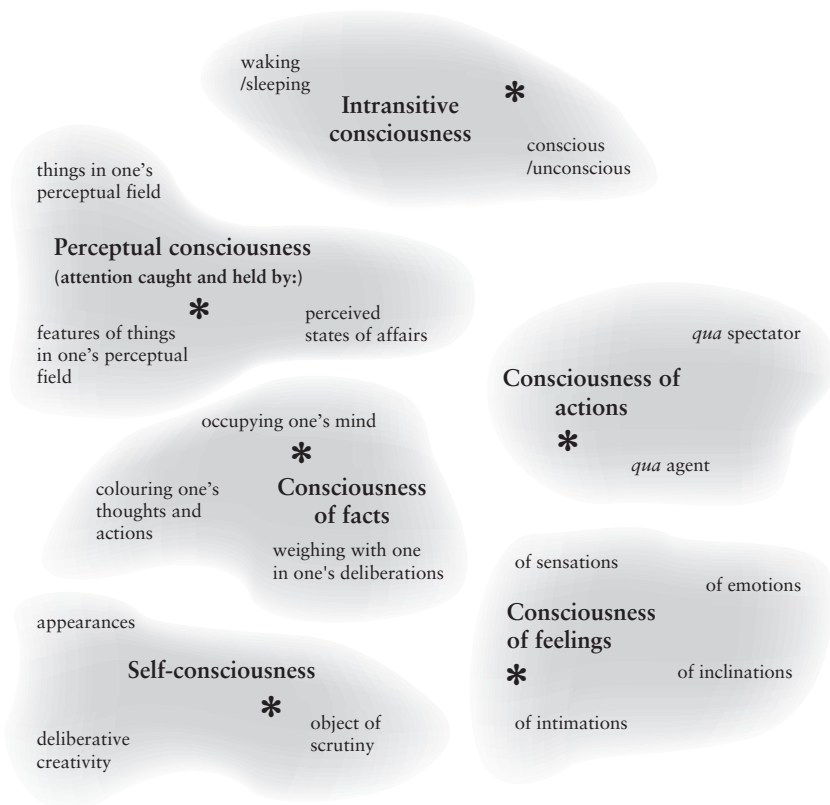


Figure 1.1 Centres of variation in the normal use of ‘consciousness’

The most important of these centres of variation are far removed from the early modern philosophical idea of an inner sense that discerns ‘operations of the mind’. They are equally far removed from the contemporary philosophical conception of *conscious experience* as possessing a unique qualitative character, of there being ‘something that it is like’ to enjoy such experience. Being perceptually conscious of something is actually a form of *cognitive receptivity* (see fig. 1.2). It is not to achieve knowledge, but to receive it (and hence is a cousin of *noticing*). The concept of *being conscious of* something belongs to the same family of concepts as *being aware of*, *noticing* and *realizing*, and is bound up with *taking cognizance* of something known. To become, and then to be, conscious of something or conscious that something is so, is either to *receive* knowledge as a result of one’s attention *being caught and held by something*, or it is for knowledge already possessed to *weigh with one*, or *on one*, in one’s deliberations, or for it to *colour one’s thought and manner of acting*. It is not to *attain* knowledge by one’s endeavours (as are discovering, discerning or detecting), but to be *given* it; or it is for knowledge *already possessed* to colour one’s thoughts, enter into one’s deliberations and modulate one’s manner of acting. *Self-consciousness*, as ordinarily used, is far removed from both *apperception* and *consciousness of one’s self*. ‘Consciousness’ and its cognates, far from signifying the general form, or ubiquitous accompaniment, of the mental, are highly specialized instruments of our language the focus of which is but rarely, and selectively, the operations of the mind.

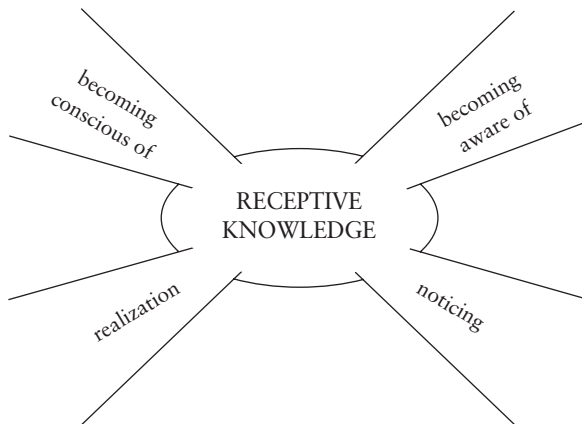


Figure 1.2 *Forms of cognitive receptivity*

Purpose of this chapter The purpose of this chapter is to clarify the ordinary concept of consciousness, and to show that consciousness is not the mark of the mind. Further, I shall show that both the early modern philosophical account of consciousness as an inner sense whereby we know what passes in our minds,² and the contemporary conception of consciousness conceived as a property of experience, namely that there is something which it is like for the subject to have it, are equally incoherent. These *philosophical* conceptions of consciousness, far from identifying the defining mark of the mental, are themselves a mark of deep and ramifying conceptual confusions.

2. The genealogy of the concept of consciousness

History of the concept: Greek The ancients had no word that can be translated as ‘consciousness’. The closest the Greeks came to our abstract noun ‘consciousness’ is *suneidesis*. The corresponding verb derives from conjoining *oida* (I know) with *sun* or *xun* (with) to yield *sunoida*: ‘I know together with’, ‘I share the knowledge that’ or, if the prefix *sun* functions merely as an intensifier, ‘I know well’, or ‘I am well aware’.³ Of course, this does not mean that they did not struggle with the same philosophical phantasms as the early moderns did and as we do. Whether that implies that they had our *philosophical* concept of consciousness, despite lacking a word for it,

² Leibniz modified the Lockean conception of consciousness. He invented the French term ‘apperception’ as a substitute for Pierre Coste’s ‘s’apercevoir de’ as a translation of Locke’s ‘perceiving one’s perceptions’. Where Locke had argued that one cannot perceive without perceiving that one perceived, Leibniz held that there are innumerable *petites apperceptions* of which we are not conscious. Kant in turn modified Leibniz’s conception of consciousness (apperception). He distinguished empirical from transcendental apperception, and held the Lockean/Leibnizean account of consciousness to be confined to empirical apperception. He agreed with Leibniz as against Locke that we can have unconscious representations, but insisted against Leibniz that it must be *possible* for us to be conscious of them. As he put it, the ‘I think’ need not accompany all my representations, but it must be capable of so doing. Nevertheless, it is arguable that Kant remained a prisoner of the incoherences of the philosophical notion of consciousness that originates with Descartes (see P. M. S. Hacker, ‘Kant’s Transcendental Deduction: a Wittgensteinian Critique’, repr. in P. M. S. Hacker, *Wittgenstein: Comparisons and Context* (Clarendon Press, Oxford, 2013)).

³ See C. S. Lewis, ‘Conscience and Conscious’, in his *Studies in Words* (Cambridge University Press, Cambridge, 1960), pp. 181–213.

depends upon whether, after careful analysis, it can be shown that we do have a coherent philosophical *concept* – or whether it will become clear that we are merely floundering about in incoherent conceptual confusion.

History of the concept: Latin The Greek pattern is also exhibited by Latin, where the combination of *scio* (I know) and *cum* (with) yielded the verb *conscio*, the noun *conscientia*, and the adjective *consciūs*. These too could be used in the sense of *shared knowledge*, or of *being privy* to information about something or someone (including oneself), as well as in the thin sense of *knowing well* or *awareness*. The idea of shared knowledge, or knowledge to which one is privy, drifted into the different idea of unshared knowledge to which one is privy – a drift from being a joint witness to being a single ‘internal’ witness, in particular, a witness against oneself inasmuch as one possesses knowledge of a guilty secret about oneself. Here is the origin of our idea of a *guilty conscience*. And it is from the idea of an internal witness that the idea of *conscience* as an internal law-giver was later to evolve. Note, however, that neither *suneidesis* nor *conscientia* was employed to signify the manner in which one is (according to the Cartesian and early modern conception) held to know of whatever is ‘passing in one’s mind’ or to know (according to the contemporary conception) what it is like to have a given experience. Nor was what one was *sunoida* or *consciūs* of restricted to operations or states of one’s mind.

History of the concept: English The emergence of the English expressions ‘consciousness’, ‘being conscious of’ and ‘self-conscious’ is surprisingly late.⁴ ‘Conscious’ and its cognates occur nowhere in the writings of Shakespeare. Their earliest occurrences,

⁴ The French *la conscience* (in the sense of ‘consciousness’) evolved from the second sense of *conscientia*, namely *knowing well* or *awareness*. Leibniz, writing about Locke in French in his *Nouveaux Essais*, coined the term ‘consciosité’ (to avoid the ambiguity of *la conscience*), but it did not catch on. The German *Bewusstsein* is of even later coinage, and first appears in the early eighteenth century as ‘bewust seyn’ in the writings of Christian Wolff, as a rendering of the Cartesian use of ‘*la conscience*’ (and so quite distinct from ‘*Gewissen*’, i.e. *conscience*). ‘*Bewust*’ was derived from ‘*bewissen*’, an Early High German compound of ‘*wissen*’. It is curious that Notker Teutonicus, in the eleventh century, used the Old High German ‘*wizzantheit*’ (derived from ‘*wizzan*’ – to know, to be aware of) as a translation of the epistemic sense of the Latin ‘*conscientia*’ (i.e. *knowing well* or *being aware of*) and used ‘*giwizzani*’ to signify *conscience*. But while *giwizzani* survived as *Gewissen*, *wizzantheit* was lost by the fourteenth century. (I am indebted to Joachim Adler for this philological history.)

according to the *Oxford English Dictionary*, are at the beginning of the seventeenth century, when ‘to be conscious’, like *consciuis*, signified *being privy* to something or to some secret. It could be applied poetically to inanimate things or abstractions as sharing knowledge of, or being witness to, human actions – as in ‘the conscious time’ (Jonson, 1601), ‘the conscious groves, the scenes of his past triumphs and his loves’ (Denham, 1643), and ‘under conscious Night, Secret they finish’d’ (Milton, 1667). ‘Being conscious’, ascribed to a person, was used in the classical sense to signify sharing a secret, being privy to something with another person, as in Hobbes (*Leviathan*, 1651): ‘Where two, or more men, know of one and the same fact, they are said to be Conscious of it one to another’, or in South’s discussion of friendship (*Sermons*, 1664): ‘Nothing is to be concealed from the other self. To be a friend and to be conscious are terms equivalent.’

Sharing a secret, however, easily mutated into no more than *being privy to or witness to* something. This usage is evident already in the 1610s. The objects of *being conscious to oneself* could be facts about other people or states of affairs, or they could be facts about oneself, for example, one’s weakness (1620). One was said to be *conscious to* the patience and wisdom of another (1649), or *conscious to* a murder (1658). Gradually the suffix ‘to oneself’ was dropped, and *consciousness to something* was transformed into *consciousness of something*. Already in the 1630s we find Massinger writing ‘I am conscious of an offence’, and in the 1660s Milton was writing of ‘consciousness of highest worth’.

Ordinary use evolved
independently of
philosophical use

In blissful independence of philosophical entanglements from the 1650s onwards, the common notion of consciousness continued to evolve in the public domain. The classical sense of being privy to a secret, of being ‘in the know’, continued into the early nineteenth century. Hence we find Jane Austen writing of Mrs Morland’s ‘conscious daughter’, that is, the daughter who shared secret knowledge with another (*Northanger Abbey*, ch. 30), and of someone who ‘looked conscious’, that is, someone who, being privy to certain information, looked as if he was ‘in the know’ (ch. 18). But by the twentieth century this use had lapsed.

Outside philosophy, one use of ‘to be conscious of’ evolved in the late seventeenth and early eighteenth centuries into a first cousin of ‘to be aware of’. So, unlike the simultaneously evolving philosophical conception of consciousness, that of which one might be said to be conscious was not confined to one’s states of mind or mental

operations. One could be said to be conscious of the rain clouds on the horizon, of the lateness of the hour, of the merits of a case, of the importance of the issue under consideration. Indeed, one could be said to be conscious of the mental state of another person, as when one is conscious of the irritability of another, or of their rising anger. Even where the object of consciousness was restricted to oneself, what one could be said to be conscious of did not have to be one's mental operations or mental states. It might well be past or present facts about oneself of which one felt ashamed or guilty, hence that one *kept privy to oneself*, or of which one felt proud and hence was 'conscious of one's worth'. But even when the objects of consciousness were one's own current mental operations, the range of mental operations of which one could be said to be conscious was, on the whole, limited to things that one could be said to *feel* – as when one is conscious of butterflies in one's stomach, of one's rising anxiety or of the increasing severity of one's pain. No one (other than philosophers) would have spoken of being conscious of thinking whatever one is thinking, or of *perceiving* (= *being conscious of*) one's perceiving (as opposed to sometimes becoming and being conscious of *what* one perceives), or of being conscious of intending to do whatever one intends to do.

The old link with being privy to something, and the phrases 'conscious to oneself' and 'conscious to something', slowly faded away. Since one could be said to *be* conscious of something, one could also be said to *become* conscious of something. This had important logico-grammatical ramifications with respect to the possible objects of consciousness (by contrast with the possible objects of *noticing*, *realizing* and *being aware of*). These will be examined later.

According to the *Oxford English Dictionary*, it was not until the middle of the nineteenth century that the term 'consciousness' came to be used to signify wakefulness, as when one speaks of *regaining consciousness* or *losing consciousness* (rather than of *regaining (or losing) one's senses*). Similarly, the common conceptions (as opposed to the philosophical notion) of being *self-conscious*, that is, being overly concerned with one's appearance and dress, or being aware that the eyes of others are upon one, and being affected thereby, seem likewise to be a nineteenth-century addition. Categories of dispositional consciousness, such as class-consciousness (1903), dress-consciousness (1918), money-consciousness (1933), are twentieth-century innovations.

Deviation of ordinary from philosophical use The most striking feature of the genealogy of *consciousness* is the extent to which philosophical use deviated from common usage from its inception. This barely noticed fact should make us examine both with care. The autonomy of the philosophical use bodes ill. For it is not impossible that the philosophical use belongs to the same category of conceptual disasters as seventeenth-century *ideas* and twentieth-century *sense-data*. In 1707 Clarke wrote: ‘Consciousness, *in the most strict and exact Sense of the Word*, signifies . . . the Reflex Act by which I know that I think and that my Thoughts and Actions are my own and not Another’s’ (emphasis added). In 1785 Reid felt confident in writing: ‘Consciousness is a word used by Philosophers, to signify that immediate knowledge we have of our present thoughts and purposes, and, in general, of all the present operations of our minds.’⁵ What philosophers held to be a special philosophical sense of the word may be no more than a special philosophical muddle.

3. The analytic of consciousness

Intransitive consciousness We must distinguish first between *intransitive* and *transitive* consciousness.⁶ Being intransitively conscious is contrasted with various forms of being unconscious, for example, being comatose or anaesthetized. Consciousness is something one may lose (on fainting, when having a high fever, or being knocked out) and regain (on recovering consciousness). Being awake differs from being conscious in so far as it is contrasted with being asleep rather than with being unconscious. ‘Is A unconscious?’ and ‘Has A recovered consciousness?’ belong typically in the hospital, whereas ‘Is A asleep?’ and ‘Has A woken up?’ are more appropriate at home. Responsiveness during sleep is far greater than responsiveness during periods of unconsciousness. There are, of course, borderline cases intermediate between intransitive consciousness and

⁵ Thomas Reid, *Essays on the Intellectual Powers of Man* [1785] (Edinburgh University Press, Edinburgh, 2002), p. 24.

⁶ The grammatical nomenclature is Norman Malcolm’s, in his ‘Consciousness and Causality’, in D. M. Armstrong and N. Malcolm, *Consciousness and Causality* (Blackwell, Oxford, 1984), p. 3.

unconsciousness for which there is appropriate non-technical terminology (e.g. semi-conscious, barely conscious, groggy, dazed, sleep-walking) as well as technical nomenclature (e.g. hypnotic trance, fugue, epileptic automatism).

Unconsciousness is a state, consciousness a condition Unconsciousness is a state of a creature, though not a mental one. Consciousness is a condition for being in any occurrent mental state. A conscious state (or state of consciousness) is not a state that is conscious, any more than a happy outcome is an outcome that is happy (as opposed to an outcome that makes someone happy) or a passionate belief is a belief that is passionate (as opposed to someone's believing passionately). Nor is it necessarily a mental state of *which one is conscious* – a state of intense concentration is a state of consciousness, but not one of which one is conscious (although one may later realize how intensely one had been concentrating, since one did not notice the clock striking twelve). Rather, it is a mental state one is in *while one is conscious* (e.g. concentrating on one's work, feeling excited or elated) as opposed to a dispositional mental state (e.g. being in a depression, being cheerful, or being anxious about something, for many weeks).

Criteria for intransitive consciousness The criteria for another person's *regaining consciousness* and then *being conscious* are behavioural – namely appropriate forms of responsiveness to perceptual stimuli. We can normally see *that* a person is conscious (someone can pretend to be unconscious, but not to be conscious). However, there are and could be no criteria for saying 'I am conscious' or even 'I have regained consciousness'. That one is conscious is not evident to one by 'introspection'. Nor is it information one might acquire by having 'access' to one's consciousness (a misuse of the term 'access'). I may become and then be conscious of your regaining consciousness, but I cannot become and then be conscious of my regaining consciousness. There is no such thing as *being conscious of one's consciousness*. This is a form of words without sense. My own intransitive consciousness is not an object of possible experience for me, but a precondition for my having any experiences at all.

Transitive consciousness: dispositional and occurrent Transitive consciousness is consciousness of *something*. It may be dispositional or occurrent. A person can be said to be class-conscious, that is, conscious of his own and others' social

class (or money-conscious, or safety-conscious), if he is disposed to pay attention to the social class of others and frequently adverts to it and to his own (like Jane Austen's Sir Walter Elliot). Someone can be said to be conscious of their ignorance (like Harriet Smith) or superiority (like Mr Darcy) if they are prone to be preoccupied with their ignorance or superiority, if they tend to dwell on it and manifest this in what they do and say.

Occurrent transitive consciousness Occurrent transitive consciousness is not a disposition. It has different modes (see fig. 1.4 on p. 27):

- (i) *having one's attention caught and held* by something;
- (ii) *giving one's attention* to one's own deliberate action;
- (iii) *something's weighing with one* in one's current deliberation;
- (iv) *something's occupying one's mind and knowingly colouring* one's thoughts, feelings and manner of behaving.

It is *these* aspects of transitive consciousness that are our concern. Let us first identify the categorial post at which this concept is stationed.⁷

Becoming and being conscious of something are not acts or activities

To become and then to be conscious of something is not to perform *an act* of any kind. There is no such thing as an act of consciousness or an act of becoming conscious of something. So to become conscious of something is not to *pay* attention to it or to *give* one's attention to it. For one cannot voluntarily, deliberately or on purpose become conscious of something – whereas one can voluntarily, deliberately or on purpose pay attention to something. Hence, one cannot decide, or refuse, to be or become conscious of something, and one cannot have a reason for becoming or being conscious of something – whereas one can decide to give one's attention to something and one may have reasons for doing so. That is why, contrary to received philosophical misconceptions, *thinking about* one's Mental Operations or Thoughts is not to be conscious (or not conscious) of them, since one can voluntarily, intentionally and deliberately think about one's state of mind, and one can be asked or ordered to think about and reflect on one's mental operations. To become conscious of something is an occurrence at a given time, but it is not something one *does* – it is something that *happens* to one.

⁷ The following analysis is indebted to, and is an elaboration of, A. R. White's *Attention* (Blackwell, Oxford, 1964), ch. 4.

Neither to become nor to be conscious of something is *an activity*. One cannot be engaged in becoming conscious of something, and one cannot be interrupted in the middle of, and later resume, being conscious of it. One cannot hurry up in being conscious of something and there are no means and methods of becoming conscious of anything.

Being conscious of something is not a mental state

To be conscious of something is not to be in a *mental state*, although what one is conscious of may, sometimes, be a mental state, as when one is conscious of one's anxiety. The reason for this is perhaps the conceptual link between being conscious of something and knowing something. For to know something to be so is not to be in a mental state of any kind, but *to be able to do various things* in the light of *what* one knows, that is, of information one possesses (see chapter 4). To be in receipt of knowledge, or for knowledge already possessed to weigh with one or affect one, is not in itself to be in any particular mental state.

Locus of the concept of consciousness in our conceptual scheme

This gives us a distinct idea of consciousness. But it does not yet give us a clear one. For that we must locate the idea in the web of our conceptual scheme, and examine its reticulations. The concept of transitive consciousness lies at the confluence of the concepts of *knowledge, receptivity, realization, awareness, attention caught and held, taking cognizance of* and *being affected by* knowledge already possessed.

Verbs of cognitive receptivity

As remarked, 'to be conscious of' belongs to the same family of cognitive verbs as 'notice', 'be aware of', 'realize', which are verbs of *cognitive receptivity*. These stand in contrast to the family of verbs of *cognitive achievement*, such as 'discover', 'discern', 'detect', which may signify the successful upshot of an intentional activity, often (but not always) an actual quest for knowledge. One may try to discover, detect and discern, and if one does so successfully, one has achieved knowledge. By contrast, verbs of cognitive receptivity, in particular in their application to modes of perception, signify not forms of *achieving* knowledge, but the manner in which knowledge is *given* one – by something's striking one, dawning on one, or catching and holding one's attention. So one can neither try to become conscious of something, nor endeavour to realize or to notice (as opposed to taking note of) something. For these verbs of cognitive receptivity do not signify acts

that might be done voluntarily, intentionally or on purpose, since they do not signify acts at all (see fig. 1.3).

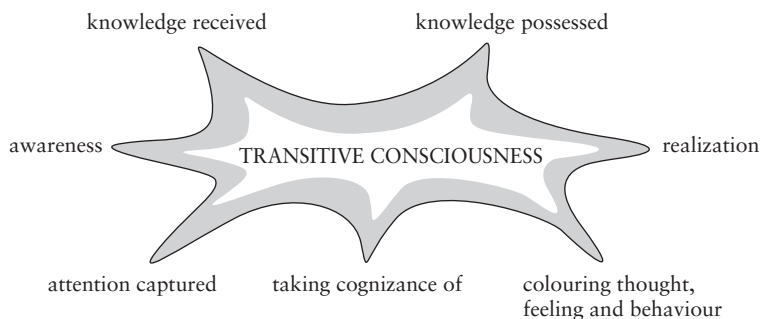


Figure 1.3 *The locus of the concept of transitive consciousness in the web of cognitive concepts*

Each of these verbs has a special role, even though they may sometimes overlap. For example, whatever one is conscious of, one is also aware of, but there is much one is perfectly aware of (since, say, one has been reliably informed) that one is not conscious of (since it is not ‘before one’s mind’, and does not occupy one). Roughly speaking, to notice something is to be struck by it, to be aware of something is for it to sink in, to realize something is for it to dawn on one, and to be conscious of something is for it to be before one’s mind. Each of these metaphorical characterizations needs to be (and can be) unpacked.

One may notice or realize something, but one may *become* aware or conscious of something. ‘To be conscious of’ is a result verb, not a success verb. It may signify the cognitive result of *becoming* perceptually conscious of something, or, in cases of non-perceptual consciousness of facts, the result of something of which one is already aware *coming before one’s mind*.

The idea of *becoming* conscious of something has immediate logical consequences marking *perceptual consciousness* off from noticing and realizing something. For one may notice something instantaneous (a flash or a bang), but what one is perceptually conscious of must be something that lasts some time. Otherwise one could not *be* perceptually conscious of it. Moreover, it must pre-exist one’s *being* conscious of it, otherwise one could not have *become* conscious of it.

*The representational
form of transitive
consciousness*

Realizing is exclusively of facts, since it is the upshot of putting two and two together. Consciousness is also *of things* (as well as of features, events and states of affairs). We may apprehend and become, and then be, conscious of Jack standing in the corner, of the ticking of the clock, of the smell of cooking, of the heat and humidity. Immediate apprehension is the normal *representational form* (even when it is not the matter) of transitive consciousness.⁸ That is, we have a marked preference for ‘consciousness of’, as opposed to ‘consciousness that’. This is no coincidence. We speak of being conscious of our ignorance, our weariness or our irritability; we may be conscious of the grief of others, of their vulnerability or of their peril; and we are conscious of impending danger, of the honour being done to us, of the importance of the situation. All these phrases can be transformed into *consciousness that* phrases: to be conscious of one’s ignorance is to be conscious (of the fact) that one is ignorant, to be conscious of the grief of another is to be conscious (of the fact) that they are grieving, and to be conscious of the impending danger is to be conscious (of the fact) that danger is impending. Why then the preference for the abstract objectual form, rather than for the factual or propositional form? Precisely because the objectual abstraction emphasizes the affinity of *consciousness of* with immediate apprehension. For what one is *conscious of* is necessarily something ‘present to the mind’, something that holds one’s attention, something that currently weighs with one in one’s deliberation, or something that colours one’s thoughts, feelings and behaviour. You may have to remind me of what I am already aware of, but you cannot remind me of what I am conscious of. Although consciousness is primarily of what is present, one can be conscious of things past too, as when one is conscious of yesterday’s victory or of the good luck one had, if these past facts are now ‘present to one’s mind’ and are affecting one’s thoughts, behaviour and manner of behaving. Moreover, one may be conscious of one’s own enduring characteristics – as when one is conscious of one’s strength or weakness, of one’s knowledge or ignorance. In such cases, one *feels* strong or weak, knowledgeable or ignorant, and one’s feeling is right. One typically

⁸ In the sense in which one may say that the representational form of knowledge is *possession*, that is, we *represent* knowledge as something we have, own, possess, can give away – this is the *picture* we use.

feels so when one is exhibiting or is about to exhibit the trait in question and realizes one is, or realizes one should refrain from, so doing.

Consciousness is polymorphous Consciousness is polymorphous (like *obeying, working, practising*).⁹ What it is to become conscious of something depends upon what it is that one has become conscious of – a sight, sound or smell, danger, weariness or a feeling of irritation. Being conscious of something may take the various forms of perceiving something – if what one perceives catches and holds one’s attention; or it may take the form of dwelling on what one is conscious of – if one is occupied with it and it colours one’s thoughts and behaviour. With some exceptions, contrary to the philosophical tradition, what one is conscious of may occur or obtain without one’s being conscious of it, that is, without its catching and holding one’s attention, and without one’s dwelling on it. Of course, perceptual verbs are not polymorphous, and consciousness is not *a form of perception*. Consciousness can be of objects of sensible perception, but it is not an *outer sense*. And it is not an *inner sense* either. This will be made clear below.

Transitive consciousness is generally a form of knowledge of its object

Consciousness of something is generally *a form of knowledge of what one is conscious of*. It may be knowledge of the presence of someone or something, as when one is conscious of Jack standing in the corner, or of the rain clouds on the horizon. Or it may be knowledge that something is so, as when one is conscious of the boredom of one’s audience, that is, conscious that they are bored, or conscious of the honour being done to one, that is, conscious that one is being honoured. Because it is a form of knowledge, what one is conscious of is so – that is, like ‘to know’, ‘to be conscious of’ is *factive*. One cannot be conscious of what is not the case. So consciousness, unlike belief, expectation, hope and fear, is *not intentional*, and its objects do not enjoy intentional *in-existence*.¹⁰ However, ironically, it is precisely when the object of consciousness is a ‘mental operation’ – in particular

⁹ The terminology is Ryle’s: see ‘Thinking and Language’, repr. in his *Collected Papers* (Hutchinson, London, 1971).

¹⁰ That is, one may believe that things are so, even though they are not – so this use of ‘believe’ is intentional. But one cannot be conscious that things are so if they are not – so ‘to be conscious that’ is not intentional and its objects do not enjoy ‘intentional in-existence’. See chapter 2.

something one feels – that, contrary to the whole philosophical tradition, consciousness, though factive, *is not a form of knowledge at all*, any more than forgetting one's troubles is a form of mnemonic deficiency. This singularity will be clarified below.

Differences between knowing and being conscious of something Although consciousness, unlike mere attention, is *generally* a form of knowledge, it is a very specific one. Whereas one can know something well, thoroughly, intimately or in detail, one cannot be conscious of something well, thoroughly, intimately or in detail. And while one can be acutely, agreeably or uncomfortably conscious of certain things, one cannot acutely, agreeably or uncomfortably know things. The reason for this is because one form knowledge may take is skill or competence – as when one knows Latin *well*. Another form of knowledge is expertise – as when one has a *thorough and detailed* knowledge of Tudor England. A further form knowledge possessed may take is acquaintance – as when one knows Jack or Jill *intimately*. But to be conscious of something is neither to possess a skill, nor to be an expert in a given domain of knowledge, nor yet to be acquainted with something or someone. One cannot be trained to become conscious of things – only trained in greater *receptivity*. There is no such thing as being *skilful* at being conscious of things – only being more *sensitive*. One can be good at learning, discovering, detecting or finding out that things are thus-and-so, but one cannot be good at becoming or being conscious of things. One can be conscious of someone without being acquainted with him, and acquainted with someone without being conscious of him. One can find out that one knows something (e.g. the dates of the monarchs of England), but one cannot find out that one is conscious of something, because one cannot *find out* that one's attention is caught by something (as opposed to finding out what has caught one's attention). One may ask 'How do you know?' but not 'How are you conscious of . . . ?' Rather one asks 'What made you conscious of . . . ?' For there are sources of knowledge (e.g. perception, reason, testimony), but no sources of what one is conscious of.

Objects of transitive consciousness Transitive consciousness may take many different kinds of objects (see fig. 1.4). What one is conscious of may be:

- (i) What one sees, hears, smells, tastes or feels – both objectually and factually (i.e. both objects (properties and relations of objects) perceived, and things being perceived to be so (as well

as events being perceived to occur and processes being perceived to go on)). This I have called ‘perceptual consciousness’.

- (ii) Facts that one has *previously learnt* and that are currently occupying one’s mind, weighing with one in one’s deliberations, or colouring one’s thoughts, behaviour or manner of behaving.
- (iii) What one is doing.
- (iv) What one is feeling, that is, a subset of traditional Mental Operations, which may be sensations, inclinations, felt dispositions to behave and, in certain circumstances, intimations (as when one feels it would be wrong to . . .).

Doubtless this crude classification can be refined. But for present purposes these distinctions suffice. Investigating them will bear fruit.

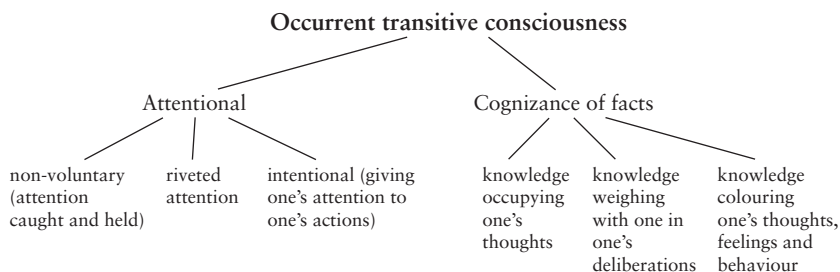


Figure 1.4 Modes of occurrent transitive consciousness

Perceptual consciousness: attention caught and held

The most prominent form of ‘consciousness of something’ in natural language is perceptual consciousness.

To become and then be conscious of something in one’s field of perception is *to have one’s attention caught and held* by something one perceives. Just as one may perceive something or perceive that something is so, so too one may be conscious of someone or something (e.g. of Jack) or conscious that something is so (e.g. that Jack is standing in the corner). That of which one is conscious is what caught one’s attention (a creature, a material thing, a sound or smell, an event or process). Its features are typically what hold one’s attention (that it is located there, its movement or manner of movement, its striking appearance and so forth).

The incompatibility of intentional attention and perceptual consciousness

The nexus with attention determines the contours of the concept of perceptual consciousness. One cannot be conscious of many things at the same time, because one cannot attend to many

things at the same time.¹¹ One cannot remain conscious of something that no longer holds one's attention (although one may be perfectly aware that things are as one was conscious of them as being). One cannot become and be conscious of something if one is *intentionally attending* to it, any more than one can involuntarily lie, discover something one already knows or detect something one has already found out. Of course, the fact that one cannot be said to become and be conscious of what one is *intentionally* attending to (since it has not caught and held one's attention) does not imply that one is *not* conscious of it, any more than the fact that one cannot be said to recognize one's wife every time one looks at her in the course of a conversation over the breakfast table means that one fails to recognize her. It means that the question of whether one is or is not conscious of what one is intentionally attending to anyway cannot arise.

Perceptual consciousness and peripheral attention This is obvious once one realizes that perceptual consciousness is commonly a matter of *peripheral attention*. I cannot be said to be either conscious or not conscious of what you are saying if I am listening attentively to you, but I may become conscious of a buzzing noise in the background. (But not all perceptual consciousness is of what one peripherally perceives. For one may become and remain conscious of a hitherto unnoticed *feature* of something one is intentionally attending to.)

Perceptual consciousness and cognitive receptivity Perceptual consciousness is not merely a matter of attention being caught and held. It is also a matter of *reception of knowledge*. Merely to have one's attention caught by something does not suffice for being conscious of that thing. For one must also *realize* what it is that has caught one's attention. One may perceive something and have one's attention caught by what one perceives, *without* being conscious of it – as when one perceives a shadow in the bushes and takes it to be a cat. Here one is neither conscious of a cat nor conscious of a shadow.

Of course, one may perceive something without its catching and holding one's attention at all – and in such cases one cannot be said to be conscious of what one perceives, although one may or may not have noticed it.

¹¹ But one can be aware of many things at the same time, since awareness of facts is not a form of attention, but of being well informed and adverting from time to time to what one knows.

To be perceptually conscious of something is not to be conscious of *perceiving it*, that is, it is not what Locke called ‘perceiving one’s perceptions’, nor is it what Leibniz and later Kant called ‘apperception’. It is, rather, to have one’s attention caught and held by *what* one perceives. Hence one cannot remain *perceptually* conscious of what one no longer perceives, just as one cannot remain conscious of something that no longer holds one’s attention. But in both cases, one may remain aware of what one was previously conscious. One can become and then be conscious of the boredom of one’s audience, of the friendliness of the company and of the spectators’ eyes upon one. These are cases of becoming conscious (*because* one comes to perceive) *that* something is so.

Self-consciousness:
1st sense *Self-consciousness*, in *one* of the senses of the English phrase, is a form of thought or awareness. It is a matter of thinking (rightly or wrongly) that others are looking at one or of being aware that they are, of this causing one to feel embarrassed and affecting the naturalness of one’s behaviour and manner. People who are self-conscious before a camera freeze, and cannot assume their normal expression. People who are self-conscious in company exaggerate their behaviour, their laughter is shrill or forced, or their shyness gets the better of them and so forth. In another sense of the phrase, to be self-conscious about one’s appearance is to be *excessively concerned* with how one will look to others, especially with regard to dress.

Consciousness
that is knowledge
possessed The cognitive receptivity of perceptual consciousness includes consciousness of perceived fact. As remarked, not all consciousness of fact involves *perceiving* things to be thus-and-so. Nor, indeed, is it always a case of cognitive *receptivity*. For it can equally well be a matter of knowledge *already possessed* coming to mind, occupying one and affecting one’s thoughts, deliberations and feelings, as well as one’s behaviour and manner of behaving. If one visits a recently widowed friend, well aware that her husband died and that she is grieving, one’s consciousness of her grief and of her recent loss does not consist in one’s attention being caught and held by something one perceives. Rather, it consists of knowledge one already possesses (things of which one is already well aware) *being before one’s mind*, *colouring* one’s thoughts and feelings, and *affecting one’s manner of behaving*. That of which one is acutely conscious *in one’s deliberations* is something that *weighs* with one and is a factor one may *take*

into account in one's decision. As noted, it is the immediacy of the influence of antecedently acquired knowledge that inclines us here towards the nominalized form 'I was conscious of her grief', rather than the more laboured 'I was conscious of the fact that she was grieving', or 'I was conscious of the honour being done to me' rather than 'I was conscious of the fact that I was being honoured'. Note that consciousness of facts incorporates realization or recollection of facts and reflection on things being as one realizes or remembers them to be. It includes a further form of self-consciousness, namely one's consciousness of one's own character traits, virtues and vices, folly or erudition, precisely to the extent that these tend to come to mind and one is prone to reflect on them.

One may be conscious of what one is doing – and this in two ways: *qua* spectator and *qua* agent. *Consciousness of one's action, qua spectator and qua agent* *Qua* spectator what one becomes and then is conscious of is typically not something one is *intentionally* doing. When one realizes with dismay that one is repeating last week's lecture, or boring one's audience, or telling a joke one has already told before, one may become embarrassingly conscious of the fact. One's attention is drawn to what one is unintentionally doing, or to an unintended consequence or side effect of what one is doing. The affinities of this form of consciousness with perceptual consciousness are patent.

One may also be conscious of what one is doing *qua* agent. One may consciously do something, for example, crack a carefully rehearsed joke at one's lecture. Here the agent knows what he is doing, and is attending to the doing of it. The agent is acting in execution of his intention, and is occupied and absorbed in carrying out his intention – as is made vivid by the common conjunction 'consciously and deliberately'. Agential consciousness is therefore altogether different from perceptual consciousness. It is not a matter of having one's attention caught and held by something – indeed, it is *deliberately giving* one's attention to something. It is an offshoot of the web of concepts of consciousness, called into being *in contrast* to spectatorial consciousness of one's action, which *is* a matter of one's attention being caught by a feature of whatever one is doing.

Self-consciousness: 2nd sense A further strand is interwoven into the concept of agential consciousness, a strand that connects it with yet another aspect of the ordinary notion of self-

consciousness. For we say of a painter or writer that they are highly self-conscious – that they deliberate at length over their work (like Leonardo), reflect deeply upon what they are doing (like Flaubert), that what they do is not spontaneous (as Picasso often was) and intuitive (like Jackson Pollock), but carefully thought through. This notion of a self-conscious writer or artist is evidently a dispositional cousin of the concept of agential consciousness of action.

Consciousness of operations of the mind

The final class of objects of transitive consciousness consists of ‘mental operations’. It was this that obsessed post-Cartesian philosophers to the exclusion of all else. La Forge (1666) already declared that ‘conscience, ou connaissance intérieure que chacun de nous ressent immédiatement par soi-même quand il s’aperçoit de ce qu’il fait ou de ce que se passe en lui’.¹² Malebranche (1674) identified *conscience* with ‘internal sentiment’. Indeed, as we saw above, Samuel Clarke (1707) and Thomas Reid (1785) declared that *strictly speaking* consciousness is *only* of the operations of the mind. Consciousness thus conceived was ‘apperception’. We shall examine this tangle of confusions below.

Its restriction to feelings

In the natural use of ‘conscious of’, the operations of the mind of which one can intelligibly be said to be conscious are primarily feelings, in the broad sense of the term which incorporates sensations, moods, attitudes, emotions, motives and intimations. No one other than a philosopher would ever speak of being conscious of seeing, hearing, tasting or smelling something, as opposed to being conscious of what one saw, heard, tasted or smelled. No one outside philosophy would speak of being conscious of thinking, believing, knowing or remembering anything – *being able to say* that one is thinking or what one thinks is not a matter of being conscious of anything. If one were to say ‘I think that such-and-such’, and were asked whether one was conscious of thinking this – one would be bewildered. One might say ‘yes’, but only because if one said ‘no’, it might seem that one was claiming that one thought such-and-such, but was *not* conscious of so doing, that is, was ignorant of so doing – and *that* one would not want to say. To be sure, what one would probably say is, ‘What do you mean?’

¹² Louis de La Forge, *Traité de l’esprit de l’homme* (1666), repr. in *Œuvres philosophiques*, ed. Pierre Clair (Paris, 1974), p. 112.

By contrast, one may well say that one is conscious or aware of the increasing pain in one's tooth, of the tickling sensation between one's shoulder blades, of the itch in one's neck. Sensations are not objects, let alone objects we perceive. But they do *catch and hold our attention*. One may be conscious or aware of one's posture and of the disposition or movements of one's limbs. And so too, one may be conscious of one's overall bodily condition, of one's feeling of exhaustion or of well-being – if one's exhaustion or sense of well-being impress themselves upon one. We may also become conscious or aware of our affections. For we may become conscious of our increasing irritation as the speaker drones on, of our feeling of jealousy as our spouse flirts with another and of our excitement as the race we are watching reaches its climax. We can, but need not, be conscious of our moods and their changes – as when we become conscious of the deepening of our depression, or of feeling exceptionally cheerful or unusually irritable. Affective consciousness usually takes the form of realization, rather than captured attention. For it often dawns on us that we are feeling jealous or irritable, and we may then dwell on it. Consciousness of the attitudes we feel, of our likes and dislikes, our approvals and disapprovals, are likewise typically the upshot of realization, the object of which then occupies us. We can become conscious of the misgivings we feel, of our feeling that it is time to go, or of our inclination to take another drink – if these cross our mind and we dwell on them prior to resolving what to do. It is interesting that 'to be aware of' sits more comfortably here than 'to be conscious of' (see fig. 1.5 for an overview of occurrent transitive consciousness).

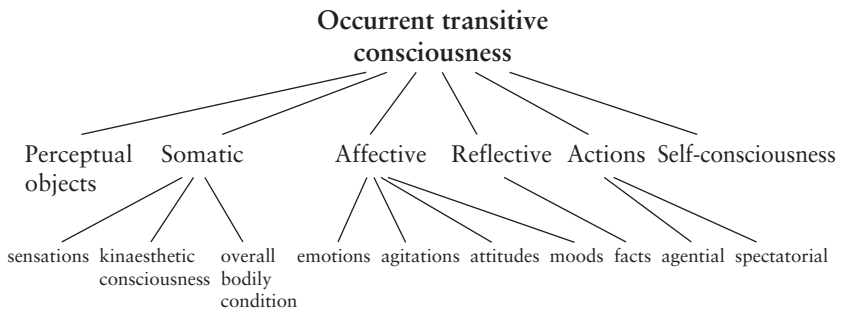


Figure 1.5 *Objects of occurrent transitive consciousness*

This aspect of the concept of transitive consciousness is a potent source of conceptual entanglement – for it is here above all that we confuse the ability *to say* how things are with us with the ability *to see*, by consciousness, apperception, introspection, or inner sense. This confusion lay at the heart of the novel early modern philosophical conception of consciousness.

4. The early modern philosophical conception of consciousness

Aristotle's questions on apperception and the general sense As noted, the ancients lacked any term for ‘conscious’ and its cognates, and they did not conceive of consciousness as the mark of having a mind. Nevertheless, it would be mistaken to suppose that none of the confusions that give rise to *our* concern with consciousness were familiar to them. Aristotle raised the question of how one perceives (or apprehends) that one sees, hears or tastes (*De Anima*, 425^b12–25). Is it by the use of the same sense-faculty as that with which one respectively sees, hears or tastes, or is it by some other sense-faculty? He queried how we distinguish white from sweet if each is perceived by the use of a different sense organ (‘On Sleep’, 455^a15–22). And he asked how we perceive a single thing as being *both* white *and* sweet, given that each quality is perceived by the use of a different sense organ (*De Anima*, 426^b8–21). How, he wondered, is the separate information all brought together to form a unified perception of a white and sweet object? His answer to this latter question was that it is by means of the *koinê aisthêsis* (later called the ‘sensus communis’), or the primary power of sense (*to prôton aisthêtikon*), the organ of which is the common or general sensorium (which Aristotle thought to be the heart and was later held to be the brain). However, these puzzling questions themselves are faulty.

The confusions underlying the questions Not surprisingly, physiological questions were here conflated with conceptual ones. How neural impulses from the separate sense organs are processed by the brain to enable us to perceive as we do involves an array of legitimate empirical questions on which neuroscientists are still working. The question of how we perceive (or apprehend) that we see, hear, smell, etc. presupposes that we do *perceive* that

we see, hear, smell, etc. whenever we do, or that our ability to *say* that we see or hear this or that rests on *apprehension* or *knowledge* that we are seeing or hearing. That is far from obvious. How we discriminate white from sweet, how we take different special sensibles to belong to one and the same object, and how we know that we are seeing or hearing, are conceptual questions of even more dubious legitimacy. Very briefly: what one cannot sensibly confuse or conflate, one cannot be said sensibly to distinguish either. Hence the question of what sense faculty is involved in distinguishing white from sweet makes no sense. The question of how, when we take and eat a lump of white sugar, we apprehend the same thing as being white, sweet, granular and cuboid presupposes the intelligibility, in these same circumstances, of sensing them as not being qualities of the same object, but as being qualities of different objects. But these presuppositions are unintelligible. We shall revert to this below.

The legacy: the general sense and the binding problem

Aristotle had opened a Pandora's box, releasing conceptual puzzles that were to occupy his successors among the Stoics, Epicureans and later the neo-Platonists for the next few centuries. Plotinus wrote of an 'inner perceptual ability' by means of which we know of our appetites (Plotinus 4.8.8.10–12). Augustine (*Confessions* vii, §17; *On Freedom of the Will* 2.2.8) held that we perceive our perception by means of an inner sense (*sensus interior*). It is the general sense (*sensus communis*), in animals and man alike, that synthesizes the information from the five external senses to form a unified perception and that enables us to perceive that we perceive. These questions were inherited from the medievals by the moderns. Descartes accepted the legitimacy of the question of how the 'information' from the different senses is synthesized to form a unified apprehension of a multiply qualified object. Indeed, he accepted a form of the Aristotelian solution that postulated a *sensus communis* to fulfil the synthesizing role. Kant endeavoured to answer the question in his account of transcendental psychology. He tried to explain the mechanisms by reference to a threefold synthesis: of apprehension, of reproduction in the imagination, and of recognition under a concept in apperception. It is this that yields self-conscious experience of unified phenomenal objects in a spatio-temporal framework. The same conceptually suspect puzzle has now transmuted into a neuroscientific question known as the 'binding problem' – namely: how does the brain bring all the 'information' from the separate

senses together to form a ‘single unified picture’? But this too is incoherent.¹³

*Origins of the early
modern philosophical
conception of
consciousness in
Descartes*

The early modern *philosophical* notion of consciousness was introduced by Descartes.¹⁴ The term does not appear in his work prior to the *Meditations* (1641), and even there it occurs just once.¹⁵ In the Third Meditation, it occurs not in relation to knowledge of one’s ‘thoughts’ or ‘operations of the mind’, but in relation to lack of awareness of power to perpetuate one’s own existence.¹⁶ 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 one’s own ‘thoughts’. He used the terms *conscientia*, *consciuis*, 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, according to Descartes, are *Thoughts*, a term which he stretched to include thinking (as ordinarily understood), sensing or perceiving (shorn of factive force), understanding, wanting and imagining.

¹³ See e.g. Francis Crick: ‘we can see how the visual parts of the brain take the picture (the visual field) apart, but we do not yet know how the brain puts it all together to provide our highly organized view of the world – that is, what we see. It seems as if the brain needs to impose some global unity on certain activities in its different parts so that the attributes of a single object – its shape, colour, movement, location, and so on – are in some way brought together without at the same time confusing them with the attributes of other objects in the visual field’ (*The Astonishing Hypothesis* (Touchstone, London, 1995), p. 22). But, of course, the brain doesn’t take what we see apart, and what we see is no picture (unless we are looking at one). The brain *makes it possible* for us to use our eyes in order to see. To do that it does not, and could not, take a picture apart and put it together again. For detailed discussion, see M. R. Bennett and P. M. S. Hacker, *Philosophical Foundations of Neuroscience* (Blackwell, Oxford, 2003), pp. 137–43.

¹⁴ The term 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.

¹⁵ I am indebted to Hanoeh Ben-Yami’s scholarship for this surprising information about Descartes.

¹⁶ René Descartes, Fourth Meditation, in *The Philosophical Writings of Descartes*, trans. S. Cottingham, R. Stoothoff and D. Murdoch (Cambridge University Press, Cambridge, 1985), vol. II, p. 34 (subsequent references ‘CSM’); *Œuvres de Descartes*, ed. C. Adam and P. Tannery, rev. edn (Vrin, Paris, 1964–70), vol. VII, p. 49 (subsequent references ‘AT’).

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 held that consciousness of operations of the mind is indubitable and infallible, and argued that the mind is, as it were, transparent. For, he wrote (CSM II, 150; AT VII, 214), it is *self-evident* that one cannot have a thought and *not* be conscious of it. Thinking is *self-presenting* – although the thoughts we have in sleep are immediately forgotten.

Descartes's 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).¹⁷ A corollary of this was that he equivocated between taking thoughts to be the objects of consciousness, that is, that of *which* one is conscious (so consciousness is an accompaniment of thought), and taking thoughts to be species (or forms) of consciousness in the sense in which seeing, hearing, smelling are species (or forms) of perceiving ('Replies to Hobbes', CSM II, 124; AT VII, 176: all acts of thought 'fall under the common concept' of consciousness). Above all, he had no explanation of the possibility of this extraordinary cognitive power, which, unlike *all* our other cognitive powers, is allegedly *necessarily exercised upon its objects*,¹⁸ and 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'.¹⁹

¹⁷ The difficulty was inherited by his successors. Arnauld, sensitive to the issue, distinguishes *reflexion virtuelle* from *reflexion actuelle*. The former, he averred, 'accompanies all our perceptions', but in addition 'there is also something explicit, which occurs when *we examine* our perceptions by means of another perception' (*On True and False Ideas* [1683] (Manchester University Press, Manchester, 1990), p. 71). The latter, he said, is not consciousness, but voluntary reflection.

¹⁸ 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 *one's seeing it*, but *the tree*. However, Descartes requires that the object of consciousness be the act of thinking, not merely *what* one is thinking.

¹⁹ Reid, *Essays on the Intellectual Powers of Man*, Essay VI, ch. 7.

Locke's psychological conception of consciousness as inner perception Locke, writing almost half a century later, characterized consciousness not epistemically, in terms of indubitability and incorrigibility, but psychologically, comparing consciousness to an 'internal sense'²⁰ whereby we perceive that we perceive (a move already made by others, such as Arnauld, La Forge and Cudworth). 'Consciousness', he explained, 'is the perception of what passes in a Man's own Mind'.²¹ We attain knowledge of what passes in our minds by the exercise of an inner sense. We cannot perceive without perceiving that we perceive.²² He did not use the term 'introspection' to name this alleged faculty of inner sense, but that should not be surprising, as the term was barely yet in currency.

Locke's introduction of self-consciousness Like Descartes, Locke held that one 'cannot think at any time, waking or sleeping without being sensible of it'. 'To suppose the Soul to think, and the Man not to perceive it is . . . to make two Persons in one Man . . . For 'tis altogether as intelligible to say, that a body is extended without parts, as that any thing *thinks without being conscious of it*, or perceiving that it does so.'²³ Unlike Descartes, he did not suppose that we must be thinking for the whole of our existence. And unlike Descartes, he did not limit the objects of consciousness to the present or to the operations of the mind. He held us to be conscious of our past mental operations and of our present as well as our past actions whenever we remember our doing and thinking whatever we did and thought. 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. It is noteworthy that consciousness has by now been individualized. One can now speak of *a* consciousness, of the same and of different consciousnesses, and of the numerical identity of a single consciousness over time. Consciousness, thus reified, *has become the mind!* According to Locke, consciousness is constitutive of the diachronic identity of a person. In a striking passage in which he repudiates the need for the same person to be the same substance, Locke invokes the novel expression self-consciousness:

²⁰ Locke, *An Essay concerning Human Understanding*, 4th edn [1700], II. i. 4.

²¹ *Ibid.*, II. i. 19.

²² *Ibid.*, II. xxvi. 9.

²³ *Ibid.*, II. i. 10; II. i. 19).

Had I the same consciousness, that I saw the Ark and *Noah's* Flood, as that I saw an overflowing of the *Thames* last Winter, or as that I write now, I could no more doubt that I, that write this now, that saw the *Thames* overflow'd last Winter, and that view'd the Flood at the general Deluge, was the same *self*, than that I that write this am the same *my self* now whilst I write . . . I being as much concern'd, and as justly accountable for any Action was done a thousand Years since, appropriated to me now by this self-consciousness, as I am for what I did the last moment.²⁴

Self-consciousness and consciousness are assimilated. Consciousness evidently encompasses all 'operations of the mind'. But because Locke conceived of personal identity as a forensic concept, and because he linked personal identity with consciousness, he included among its objects one's consciousness of one's own actions while performing them ('consciousness . . . that I write now').

Descartes held that thinking (in his broad sense of the term) is the defining essence of mental substances (minds), so he argued that one must think (engage in mental operations) all the time, otherwise one would cease to exist. Locke disagreed, denying that substances are defined by a single essential property. But he agreed that one could not think without perceiving that one thinks. Leibniz in turn disagreed with Locke, holding that there are multitudinous *petites perceptions* which we do not perceive, of which we are not conscious. But he agreed with Locke (against Arnauld) in holding consciousness to be a form of reflection (for which Reid was later to criticize him).

*A Galtonian picture
of the early modern
philosophical
conception*

tradition.²⁵

It was from these foundations that the eighteenth-century debate developed. One may summarize, in a *Galtonian picture*, the conception of consciousness that Kant, to his misfortune, inherited, via Wolff, from the Cartesian and empiricist

²⁴ *Ibid.*, II. xxvii. 16. The term 'self-consciousness' was initially a philosopher's term of art. Locke was not the first to use the expression to mean the capacity for reflexive knowledge of one's mental operations. Cudworth, in his *Treatise on Freewill* (1688) wrote: 'We are certain by inward sense that we can reflect upon ourselves and consider ourselves, which is a reduplication of life in a higher degree; for all cogitative beings as such are self-conscious'. It is interesting that Pierre Coste translated 'self-consciousness' by *conscience* and added the English term – which is indicative of the novelty of the usage.

²⁵ For detailed investigation of Kant's conception of consciousness and self-consciousness, see Hacker, 'Kant's Transcendental Deduction'.

- (i) Consciousness is the general form of Operations of the Mind, that is, one cannot 'think' without being conscious of one's 'thinking'.
- (ii) Consciousness is an *inner sense* – by the use of which we know how things are subjectively with us.
- (iii) The deliverances of consciousness are indubitable – one cannot doubt whatever one is conscious of.
- (iv) The deliverances of consciousness are infallible – one cannot make a mistake about what one is conscious of.
- (v) One can think things to be thus-and-so, and one is then unavoidably conscious of so thinking. But one cannot in turn think that one is conscious of thinking. It may sensibly seem to one that things are thus-and-so, but it cannot sensible seem to one that it sensibly seems to one that things are thus-and-so.
- (vi) Objects of consciousness are operations of the mind.
- (vii) Objects of consciousness are confined to the present.
- (viii) The objects of consciousness are privately 'owned' (no one else can have my experiences – experiences are logically private, inalienable, property).
- (ix) The objects of consciousness are epistemically private – only I *really* know (because I have privileged access to) the operations of my mind.
- (x) One's consciousness of what passes in one's mind requires possession of ideas or concepts of mental operations. These ideas or concepts have no logical relationship to behaviour, since they are applied in inner sense without reference to one's behaviour. To possess them requires no more than consciousness of the ideas (Descartes^a), or a private ostensive definition (Locke^b).
- (xi) Consciousness of the operations of the mind is *self-consciousness* – consciousness of how things are with one's self *in foro interno*.

List 1.1 A Galtonian representation of the early modern philosophical conception of consciousness

So the mind is, as it were, *transparent*, and what is in the mind is, so to speak, *self-presenting*. So *mind is better known than matter*. Consequently, *the private is better known than the public*. Points (viii) to (x) commit the early moderns and their followers to the intelligibility of a logically private language. This fatal flaw will not be discussed here. Disagreements, which continued well into the nineteenth century, turned largely on the questions of (a) whether there are unconscious operations of the mind; (b) whether inner sense is contemporaneous with, or subsequent to, its objects (Comte, Spencer, Mill); and (c) whether consciousness is or is not infallible. In the post-Kantian and German idealist debate, attention was focused on the nature of transcendental self-consciousness and its ramifications. This will not be discussed here.

5. The dialectic of consciousness I

3 presuppositions of the early modern philosophical conception

Such was the conception of consciousness and self-consciousness that plagued philosophy in the Cartesian/Lockean tradition. The whole structure turns on three simple and

correct thoughts.

First, the sincere first-person use of *many* psychological attributes is indubitable. If one feels a pain, one cannot doubt that one is in pain. If one thinks that it is time to go, one cannot doubt that one does. If one is afraid of tomorrow's examination, one cannot doubt

Notes a and b to List 1.1

^aDescartes: '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).

^bLocke: 'Such precise, naked appearances in the mind, 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).

that one is. It makes no sense to say ‘I doubt whether I am in pain’, or ‘I doubt whether I think that . . .’.

Secondly, in *many* cases, one cannot be mistaken. So, for example, one cannot be *mistaken* that one is in pain, any more than one can mistake a pain for a tickle; nor can one be mistaken that one thinks that $2 + 2 = 4$, any more than one can misidentify one’s thought that $2 + 2 = 4$ as the thought that $2 + 2 = 22$.

Thirdly, in those cases which Descartes held to be suitable as the premise of a *cogito* proof of his existence, that is, all the cases that seem to involve certainty and infallibility, *truthfulness guarantees truth*.

The logic of illusion 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, then 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 presence of certainty, and the impossibility of mistake for the presence of infallible knowledge. To clarify this we must penetrate *the logic of conceptual illusion* – the dialectic of consciousness.

The logical exclusion of doubt excludes empirical certainty too Doubt needs reasons. The possibility of doubting an empirical truth such as ‘Jack is in pain’ or ‘Jill thinks that it is time to go’ may be excluded by realization of the eliminability of all genuine alternatives in the circumstances. Here possible doubt is excluded by the available evidence. Here, it is quite certain (and one is 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 not excluded *de facto*, but *de jure*. For no sense has been given to the words ‘I doubt’ as a prefix to the empirical proposition in question, or in the circumstances in question. To give a few familiar examples: 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). Similarly, 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, for certainty can be *established* by excluding alternative possibilities. But when it is *logically impossible* to doubt an empirical proposition – when it makes *no sense* to doubt, then it makes no sense to speak of certainty either. The satisfaction of the conditions of subjective certainty does indeed exclude all doubt, but if all doubt is *logically* excluded, there is nothing for subjective certainty to exclude. There is no room for certainty – the logical space, so to speak, has vanished. Similar considerations apply to the exclusion of mistake with regard to an empirical proposition. The logical impossibility of a mistake does not imply infallible knowledge, but the exclusion of knowledge *together with error*. This is precisely how things are with regard to first-person uses of the subset of psychological verbs that satisfy Descartes's demands on *cogitationes*. It is precisely because it *makes no sense* for someone to be in pain and doubt whether he is, or to mistake his thinking that it is time to go for his thinking that Paris is the capital of France, that it makes no sense to say that he is certain, or knows infallibly and incorrigibly, how things are with him in such respects.

That truthfulness guarantees truth does not imply knowledge

It is perfectly correct that with regard to avowals of pain, confessions of one's thoughts, assertions of how things sensibly appear to one to be ('It visually seems to me . . .'), truthfulness in general guarantees truth. In such cases, the speaker's word goes (although not always indefeasibly). It is all too easy to try explain this by reference to the idea that the speaker *knows* how things are with him because he has 'privileged access' to his mind by introspection, and that is why truthfulness guarantees truth. That is mistaken. The speaker's word goes, not because he is a *witness* to his own consciousness, but because he is an articulate *agent*. I shall elaborate.

Where knowledge is logically excluded, so is ignorance

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 it is not the case that we 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. 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 so with us. And our sincere word therefore has a privileged status *for others* (it is *logically* good evidence for them). 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 work* for the verb ‘know’ to do. It shows only that ignorance, *together with knowledge*, are here logically excluded.

The illusions of inner sense, apperception and introspection

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 too it seems that we know how things are with us ‘inwardly’ 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’.²⁶ It is by the use of this inner sense, it seems, that we perceive, apperceive, introspect 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 (there is here no ‘more light, please’, no looking more closely or using a telescope);
 - (iii) it never fails us, but always yields knowledge;
- ∴ (iv) we know the mind better than the material world (cp. Descartes, Brentano, Husserl).

But there is no such thing as a cognitive faculty that is miraculously immune to error, and no such thing as a faculty of perception that enables us to perceive without any organ of perception and the

²⁶ W. James, *The Principles of Psychology* (Holt, New York, 1890), vol. 1, p. 185.

successful exercise of which is independent of circumstances of observation. 'To perceive', as well as 'to see', 'to hear', etc. have a legitimate use as success verbs – but there is no such thing as succeeding if there is no logical possibility of failing. (As noted, 'to be conscious of', although factive, is *not* a success verb – one cannot try to become or succeed in being conscious of something.)

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 high road 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.

Of the importance of not confusing the ability to say with the ability to see

What is true is that if we are asked whether we are in pain, whether we want this or that, whether we think things to be so, or are 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 normally 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, that is, 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.

Saying what one thinks

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 may have 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 takes it to be a matter of opinion, or if he takes the evidence not to be 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 which does not signify a mental operation currently taking place, but indicates (for others) the epistemic weight of the proposition to which it is affixed.²⁷

Knowing and not knowing what one thinks or wants 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 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. 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!’ amounts to the same as ‘Now I have decided’.²⁸

²⁷ Of course, there are other uses of this verb (see chapter 10).

²⁸ 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 and does not serve as a confession of epistemic fault. So too, ‘I know I am in pain’ or ‘I know I intend to go’ do not signify the upshot of a successful exercise of a cognitive faculty, and do not serve to make a cognitive claim. They serve merely to emphasize that I am indeed in pain, or to concede that I do indeed intend to go. ‘You’re in pain!’ – ‘Yes, I know’ is a joke.

To be able to say
that one perceives is
not to perceive that
one perceives

So, to return to the questionable questions of the ancients: 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 our seeing by a general sense (a *sensus communis*)? — 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 see // I can see . . .’, ‘I hear // I can hear . . .’, ‘I smell // 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. 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, in the case of proper sensibles, to be rejected, and in the case of common sensibles to be answered by citing the sense-organ and sense-faculty used. But even in the latter case (say, of feeling the shape of something with one’s fingers), one does not *perceive* that one perceives. Rather, one perceives with one’s fingers, one’s sense of touch, and *can say so*. Nor is one conscious *that one perceives*, 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*.

It must be possible
for the ‘I can say’ to
accompany all my
representations

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 with me. Therefore, 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 things being thus-and-so, any more than reflecting on Julius Caesar’s assassination is to be conscious of it. To reflect on things being thus-and-so

with me is a mental act, which I may be asked or ordered to perform, and may perform voluntarily, intentionally and deliberately. But I cannot intelligibly be asked or ordered to be conscious of things being thus-and-so with me, and to be conscious of things being thus-and-so is not an act, *a fortiori* not a voluntary or intentional one.

*The illusion of
consciousness as an
inner sense*

In brief, consciousness conceived as an inner sense is a fiction. Roughly speaking, anything that Descartes might, *with good reason*, wish to cite as an indubitably and infallibly known *act of 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 the mark of a mind, but of thoroughgoing confusion.

*Animal consciousness
contrasted with
human consciousness
and self-consciousness*

Given this confusion, the idea that consciousness is *the* mark of the mind collapses. So it should. After all, consciousness, properly understood, is characteristic of other animals than humans. All developed animals are sentient – they have the powers of sensation and perception, and are susceptible to pleasure and pain. They typically have a diurnal cycle of sleeping and waking, hence enjoy intransitive consciousness. They can have their attention caught and held by objects in their perceptual field, and so enjoy perceptual consciousness. But, of course, they are not language-users. Nor is there an 'I can say' that can accompany all their representations. They do not have an 'inner life' of reflection, recollection and articulate feeling. They are conscious, but not, in *this* sense, self-conscious beings. But it is precisely such features that characterize having a mind. Furthermore, many further attributes distinctive of creatures that *do* have a mind cannot be subsumed under the rubric of Cartesian thoughts (definitive of *Cartesian* consciousness) since these attributes are neither indubitable nor transparent. We have wide-ranging cognitive powers, but sometimes think we know something and are mistaken. We have beliefs, but sometimes deceive ourselves about what we really believe. We have mnemonic powers, but sometimes think wrongly that we remember something. Our powers of understanding are great, but we often mistakenly think we understand something. It is evident that only conscious creatures (properly so called) can be said to have a mind, but consciousness is not sufficient for having a mind.

6. The contemporary philosophical conception of consciousness

The decline and revival of interest in consciousness in the modern era

The Cartesian/empiricist conception of consciousness dominated philosophical thought concerning the mind well into the twentieth century. But among analytic philosophers of the Vienna Circle in the interwar years, and among Oxford philosophers of the postwar years, interest in consciousness waned. This was due partly to the rise of behaviourism, partly to a shift of interest away from philosophy of mind and towards philosophy of logic and language in the 1920s and 1930s, and partly to the powerful criticisms of the foundations of both Cartesianism and classical empiricism launched by Wittgenstein and Ryle in mid-century. Interest was reawakened by the emergence, first, of central state materialism in the writings of Smart, Place and Armstrong,²⁹ which identified types of mental states with types of brain states, and then by its successor, namely functionalism.

Functionalism, advanced in the USA, eschewed the identification of types of mental states with types of brain states. Philosophical functionalists hoped to explain the nature of any mental state solely by reference to its function in correlating causal inputs, behavioural outputs and its causal relations to other mental states (just as a Turing machine-table simultaneously defines the roles of all the machine states in causal terms without circularity). To be sure, any such individual mental state of a being is held to be contingently *token-identical* with whatever cortical or electro-mechanical vehicle realizes it. Functionalism seemed to offer the benefits of behaviourism (the correlation of stimuli (inputs) with behaviour (outputs)), and of materialism (the token-identity thesis), without denying the existence of internal mental states. But it construed internal mental states solely in *functional terms*. A mental state was to be defined in terms of the inputs and outputs it coordinates and its causal interaction with other internal states. This, as critics pointed out, conspicuously omitted

²⁹ U. T. Place, 'Is Consciousness a Brain Process?', *British Journal of Psychology*, 47 (1956), and J. J. C. Smart, 'Sensations and Brain Processes', *Philosophical Review*, 48 (1959); see also H. Feigl, 'The "Mental" and the "Physical"', *Minnesota Studies in the Philosophy of Science*, vol. II (University of Minnesota Press, Minneapolis, 1958); D. M. Armstrong, *A Materialist Theory of Mind* (Routledge & Kegan Paul, London, 1968).

mention of the felt character of the experiences that sentient creatures enjoy – experiences of pain or pleasure, hunger, thirst, seeing and hearing, longing, expecting, being sad or joyful. Against the functionalist background, it *seemed* that it was perfectly intelligible to suppose that there might be creatures ('zombies'), just like us in all behavioural respects, subject to the same 'inputs' and yielding the same 'outputs', and having the same causal connections between internal, non-conscious 'machine-states' – but *without enjoying any experiences whatsoever*. It was in reaction to this illusion that the new wave of interest in consciousness emerged in the 1970s with a seminal paper by Thomas Nagel.³⁰ To save us from the fear that all others might be 'zombies', to save our humanity from reductive physicalism and soulless functionalism, consciousness was appealed to as the defining feature of the mind and the characteristic mark of the mental. For, it was now argued, what was irremediably missing from functionalism was *conscious experience* (see fig. 1.6).

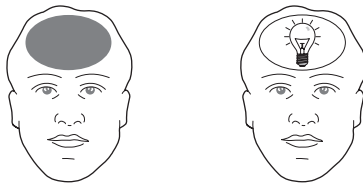


Figure 1.6 *Zombies and us. It is striking how readily the metaphor of 'light inside our heads' comes to be used here. But, if there is any light, it is certainly not inside our heads*

An experience, it was averred, is conscious if there is something which it is like for the subject of the experience to have it. For is there not something it is like to be in pain, to feel joy, to see and hear? And a subject of experience is conscious if there is something that it is like for it to be that subject. For while there is nothing it is like for a brick to be a brick, or for an ink-jet printer to be an ink-jet printer, there is surely something it is like for a cat to be a cat, for a bat to be a bat, for us to be human and indeed for me to be me. *That* is the essence of consciousness and of conscious experience. What began as a ripple in the USA in the 1970s had acquired tsunami proportions by the 2010s, when 'consciousness studies' were all the rage and 'the what-it's-likeness of experience' the slogan.

³⁰ T. Nagel, 'What is It Like to be a Bat?', *Philosophical Review*, 83 (1974).

The contemporary philosophical conception of consciousness is no less incoherent than the early modern conception. If our humanity needs saving in the face of modernity, it is from far more serious things than functionalism – which is no more than a house of cards that will collapse under the weight of conceptual criticism.

3 pivotal theses of
current philosophical
misconceptions

Three salient theses determine the concept of consciousness advanced by contemporary philosophers and cognitive scientists:

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’. Knowing what it is like 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.*

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.*

Some clarification and elaboration is needed.

(a) Just as Descartes (and his successors) misguidedly extended the notion of Thought to include *perceiving* and *wanting something* (etc.) so the new conception of Conscious Experience is misguidedly extended to include *thinking*, *knowing*, *believing* and *understanding* (which are no more ‘experiences’ than perceiving and wanting are species or forms of thought).

(b) Each conscious experience is argued to have its own qualitative character – its distinctive *phenomenal feel*.³¹ The individual feel of an

³¹ 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* (Appleton-Century-Crofts, New York, 1932).

experience was dubbed a *quale*.³² The problem of explaining these phenomenal qualities, it is held, is the problem of explaining consciousness. For what characterizes *any* conscious experience are the distinctive *qualia* that accompany it.

(c) It is important to realize that the 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, only one foot is always on the ground’. The question is not ‘What does it resemble?’ but rather ‘What is it like *for you?*’ It concerns the subjective qualitative feel of the experience – what it feels like *for the subject*.

*The depth of mystery
or the depth of
illusion*

This conception of consciousness and of conscious experience captured the imagination of philosophers, psychologists and even cognitive neuroscientists in the USA. In due course, the confusions spread to Britain and continental Europe. It appeared to raise a whole battery of enticing and mysterious new questions for cognitive science and evolutionary theory to grapple with. What, it was wondered, is consciousness *for?* What is its evolutionary advantage? Could one not have creatures who behave just like us, only without any ‘inner light’ of consciousness – that is, without there being anything that it is like to be them? How could anything so mysterious as consciousness emerge from mere matter? Is consciousness compatible with our scientific understanding of the universe? And so forth. These are all either trivial questions or pseudo-questions.³³ But if one accepts this tempting account of the uniqueness and peculiarity of consciousness, then they seem anything but trivial or absurd – they seem deep questions at the frontiers of knowledge.

7. The dialectic of consciousness II

*The 4 temptations on
the road to illusion*

Why is it evidently so tempting to agree to this analysis of consciousness? Four factors are in play.

³² The term was borrowed from C. I. Lewis, *Mind and the World Order* (Scribner’s, New York, 1929).

³³ See Bennett and Hacker, *Philosophical Foundations of Neuroscience*, ch. 11, for detailed deconstruction of these confused questions.

First is the persuasiveness of the claim that there isn't anything which it is like to be a brick or an ink-jet printer, but there is something it is like to be a bat or a dolphin and there is certainly something it is like to be a human being. 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 what it is like for it to print a page.

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'. I shall explain this below.

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.

Finally, the appeal of mysteries, of facing the deepest and most difficult problem known to man, of being at the last frontier of knowledge, is well-nigh irresistible. In philosophy, *there are no mysteries* – only mystifications and mystery-mongering.

*3 antidotes to the
4 temptations* The temptations must be resisted, and sober analysis should take their place. I shall briefly defend three antitheses.³⁴

- (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.

*The qualitative
character of
experiences* (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 the qualitative characteristics of certain experiences.

³⁴ 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. I shall use the term 'experience' in the broad and ill-defined sense in which it is currently employed by students of consciousness.

(1b) Every experience is a *possible* grammatical subject of attitudinal predicates, for example, 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. Hence it is mistaken to hold that every experience has a qualitative character. 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 in the street or to hear the chatter in the bus feels neither pleasant nor unpleasant, and is neither repulsive nor attractive.

Experiences are identified by their object, not by their qualitative character

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

Confusions of qualia

(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 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 $3 \times 3 = 9$.

The felt character of experiences and the confusions of existential quantification

(2) It is true that one can ask someone ‘What was it like for you to V?’ (where ‘V’ signifies an experience). 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 or 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 that *it was like* for A (or for me) to V, namely . . .'. 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 is it like for you to V?' For the answer (save among the illiterati) is not 'To V *is like* wonderful', but 'To V *is* wonderful'. And the existential generalization of that is 'There is something that it is to V, namely wonderful'. It *cannot* yield the form 'There is something that it is like to V, namely wonderful'. The latter aberration is the result of a miscegenous crossing of the existential generalization of a judgement of similarity with an existential generalization of a judgement of the affective character of an experience. And the result is, strictly speaking, 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 your 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.)

Logico-grammatical constraints on what it is like to be or to do

I now turn to the third antithesis. It makes perfectly good sense to ask 'What is it like to be a soldier (a mother, an old-aged pensioner, wealthy, unemployed)?'. 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, of the typical career of a V-er, or of being in a given condition. 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.

Specification of the subject of 'What is it like to be . . . ?'

(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, that is, 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 pregnant?' to women. But sometimes it *is* necessary, for example, '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?'

Existential generalization again

As in the previous cases of 'What is it like to V?', so too here the 'like' drops out in existential generalization. If one answers 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 '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.

Constraints on the subject and object terms of 'What is it like for a . . . to be a . . . ?'

(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*.

The transgression of the 2 principles of contrast

But the form of words that we are being offered 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. 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, for example, ‘Nasty, brutish and short’ or ‘Full of hope and fear’. Nor is there any 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 obvious that the supposition that there is something it is like for me to be me is 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. But I can, of course, tell you what my life has been like.

*Reducing mountains
to molehills*

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 20 copies of one’s paper. *For only sentient creatures have social roles and experiences, enjoying some, disliking others and being indifferent to most* – a meagre result for so much noise.

8. The illusions of self-consciousness

The ordinary notions of self-consciousness It should be evident that the philosophical conception of self-consciousness not only deviates from the common or garden notions, but is also a product of philosophical confusions rooted in the notion of apperception transmitted from Locke to Leibniz and from Wolf to Kant. The ordinary notions are perfectly respectable: (a) excessive concern with one's own appearance, especially one's dress; (b) one's responses to the thought that others are looking at one; (c) deliberate, as opposed to spontaneous, creative processes.

What self-consciousness is not What self-consciousness is *not* is:

- (i) Consciousness of one's self – since there is no such thing as a 'self' thus understood.³⁵
- (ii) Apperception – since there is no such thing as perceiving one's perceptions; *a fortiori* it is not a matter of the *possibility* of perceiving one's perceptions.
- (iii) Thinking about one's 'thoughts' or 'perceptions' – since although one may indeed think about one's thinking (e.g. how muddled it is) and think about one's perceptions (e.g. how vivid they are), to do so is not to be conscious of one's thoughts or perceptions. In general, to think about something (e.g. Julius Caesar) is not to be conscious of that which one is thinking about.
- (iv) An 'I think' that is capable of accompanying all one's representations (as Kant supposed transcendental self-consciousness to be) – What may be said to be capable of accompanying all my representations is an 'I say'. But *to be able to say* does not imply being conscious of things being as one might describe them as being, only not being ignorant of one's 'representations'.

So much for philosophical confusions. Unfortunately, these have spread to the scientific domain. In psychology, self-consciousness is commonly identified with introspection traditionally construed (as in James). We need not dwell further on this. Among animal behaviourists, the idea has sprung up that the ability to recognize oneself in a

³⁵ For detailed examination of the matter, see my *Human Nature: the Categorical Framework* (Blackwell, Oxford, 2007), ch. 9, sections 1–2.

mirror is a mark of self-consciousness.³⁶ We shall discuss this misconception below (see pp. 396f.). For the moment note that this temptation is generated largely by the form of words in which this capacity that we share with chimpanzees, elephants and dolphins is described, namely ‘recognizing *oneself* in a mirror’. For it is but one short step from ‘recognizing oneself’ to ‘recognizing one’s self’. The temptation is greatly lessened if the ability is described as ‘recognizing one’s reflection in a mirror’, which is no more a siren’s song than is ‘recognizing one’s hand in a mirror’, or even just ‘recognizing one’s hand’.

Neuroscientists are subject to all these pressures, but add more of their own. Impressed by the thought that ‘the human capacity of self-perception, self-reflection and consciousness development are among the unsolved mysteries of neuroscience’, scientists in the Max Planck Institutes of Psychiatry in Munich and for Human Cognitive and Brain Sciences in Leipzig and from Charité in Berlin have been studying lucid dreams. Their supposition is that ‘during wakefulness, we are always conscious of ourselves’ – which makes it difficult to identify the ‘seat of meta-consciousness in the brain’. But lucid dreamers, it is argued, unlike normal dreamers, are conscious of dreaming while they are asleep. By examining their brain activity during sleep, it is therefore possible to identify the parts of the brain that are associated with self-consciousness. Indeed, such fMRI investigation has ‘made the neural networks of a conscious mental state visible for the first time’.³⁷

This is conceptually incoherent. First, it is wrong to suppose that when conscious (i.e. awake) we are always conscious of ourselves. This confuses the ability to say what we are doing with being conscious of doing what we are doing, either *qua* agent or *qua* spectator. Secondly, a lucid dream is a dream in which the sleeper dreams that he is dreaming, not a dream in which he is conscious that he is. For there is no such thing as being conscious of anything when one is fast asleep and dreaming. Whatever one dreams *of* is an object of one’s

³⁶ G. G. Gallop, Jr, J. R. Anderson and D. J. Shillito, ‘The Mirror Test’, repr. in M. Bekoff, C. Allen and G. M. Burghardt (eds), *The Cognitive Animal* (MIT Press, Cambridge, Mass., 2002), pp. 325–34.

³⁷ M. Dresler *et al.*, ‘Neural Correlates of Dream Lucidity Obtained from Contrasting Lucid versus Non-Lucid REM Sleep: a Combined EEG/fMRI Case Study’, *Sleep*, 35 (2012), pp. 1017–20. Reported in *ScienceDaily* (27 July 2012), at <http://www.sciencedaily.com/releases/2012/07/120727095555.htm>, accessed 1 Feb. 2013.

dream, not something that catches and holds one's attention. One does not attend to anything when one dreams – at most one might dream that one is attending to something (while in fact one is snoring away and fast asleep). Nor is anything one dreams a factor that one might take into account in one's deliberations and decisions in one's sleep, since one neither deliberates nor decides anything while one is fast asleep. Thirdly, as we have seen, self-consciousness is not consciousness of one's self, nor is it 'consciousness of one's consciousness' – for these are conceptual chimeras. They need a Theseus to answer their riddles and destroy them, not a team of neuroscientists to discover the locus of 'meta-consciousness'.