PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

Splendours and miseries of the brain

S. Zeki and FRS

Phil. Trans. R. Soc. Lond. B 1999 **354**, 2053-2065
doi: 10.1098/rstb.1999.0543

THE ROYAL SOCIETY

Splendours and miseries of the brain

S. Zeki, FRS

Wellcome Department of Cognitive Neurology, University College London, London WC1E 6BT, UK

In this speculative essay, I examine two evolutionary developments underlying the enormous success of the human brain: its capacity to acquire knowledge and its variability across individuals. A feature of an efficient knowledge-acquiring system is, I believe, its capacity to abstract and to formulate ideals. Both attributes carry with them a clash between experience of the particular and what the brain has developed from experience of the many. Both therefore can lead to much disappointment in our daily lives. This disappointment is heightened by the fact that both abstraction and ideals are subject to variability in time within an individual and between individuals. Variability, which is a cherished source for evolutionary selection, can also be an isolating and individualizing feature in society. Thus the very features of the human brain which underlie our enormous evolutionary success can also be a major source of our misery.

Keywords: abstraction; idealism; knowledge; colour; variability

1. INTRODUCTION

It is commonly thought that, scientifically, the last millennium belongs more properly to the physical sciences, while the future will be dominated increasingly by the biological sciences, and especially by neurobiology. Physics and chemistry are now generally regarded as mature sciences, which have made rapid advances and major contributions, especially in the closing century of the last millennium. Through these, Man has been able to conquer and dominate his environment to a greater or lesser extent and to increase immeasurably his comfort and well-being. Physical sciences are still making major contributions. But the 1993 decision of the US Congress to reject the giant particle accelerator—the superconducting supercollider, which physicists had hoped would take them beyond quarks—suggests perhaps that society is increasingly sceptical about the value of further expensive projects that may well contribute to knowledge but, in society's implicit judgement, will perhaps not contribute commensurately to improving the condition of Man. By contrast, the expenditure of Western societies on biological sciences shows no sign of abating. This is again a reflection, I believe, of the inarticulated judgement of these societies that, now that the physical environment has been more or less understood and controlled, the time has come to devote increasing attention and resources to our biological well-being. That well-being consists of course in improving our health and genetic inheritance and in eradicating disease and physical suffering. But it also consists, above all, in making men and women happy and alleviating what Sigmund Freud called 'the psychological misery of mankind'.

In truth, the revolution brought about by the physical sciences in the last millennium was also a major biological, or at least psychological, landmark in human history. The Copernican and Galilean revolutions may have been revolutions in the physical sciences but their psychological impact on humans was not negligible, and we are still living with the consequences of these revolutions today. Demoted from his position at the centre of an undefined universe, Man had to turn increasingly to understanding his place in nature: deprived of the certainty of an omnipresent and omniscient God, who had chosen man as the instrument of His will, he had to turn increasingly to understanding the meaning of his existence. Deprived further still by the Darwinian revolution of the certainty of a singular Act of Creation by a God with a design, he has instead found himself, together with his highly evolved brain and mind, nothing more than the product of the slow and adventurous process of evolution, whose only purpose is the survival of the species, but of any species, not just the human. Unable to find unequivocal answers to the meaning of his existence, he is likely to turn increasingly inwards and seek a meaning within himself, accelerating a process started with the Copernican revolution.

The eradication of diseases, poverty and of all other agents that either shorten people's lives or incapacitate them physically or mentally to varying degrees, thus depriving them of the opportunity to contribute fully to their civilization, and in return enjoy to the maximum the fruits of its labours, does not, however, provide a solution to the problem of human unhappiness posed by Freud (1930). In his book *Civilization and its discontents*, he summarized the problem as follows. Men, he wrote, 'are beginning to perceive that all this newly won power over space and time, this conquest of the forces of nature, this fulfilment of age-old longings ... has not made them feel any happier'. He traced much of this distress to a social source—'the inadequacy of our methods of regulating human relations in the family, the community and the state'. But 'when we consider how unsuccessful our efforts to safeguard against suffering in this particular area have proved, the suspicion draws upon us that a bit of unconquerable nature lurks concealed behind this difficulty as well—in the shape of our own mental constitution'.

(a) *The study of our mental constitution*

It is the study of that mental constitution that presents a huge challenge to science, probably the most formidable that it has ever faced, and much of that challenge is neurobiological. For it seems more than likely that our mental constitution is dictated by the neural organization of our brains. How we can define mental constitution in neural terms, as indeed we must if we believe that it is an expression of neural activity, is not at all obvious today. Philosophers and others might even argue with greater or lesser conviction that there is no direct proof that our mental constitution is uniquely determined by our neural apparatus. Even in spite of this, to common sense the supposition is *bene fundata* and I shall suppose throughout this essay that underlying our mental constitution is a neurological apparatus which determines its capacities and its limits (see also Searle, this issue).

The first important point to make about our mental constitution is that it is the product of evolution, conferring a selective advantage on us. Some of the advantages are so obvious that they are hardly worth mentioning. But what of the disadvantages implied in Freud's phrase 'the psychological misery of mankind', which he traced to the mental constitution of Man? Could it be that the misery entailed by possessing such a constitution is the by-product of the enormous selective advantages conferred by the same constitution? In this essay, I examine two evolutionary developments underlying the enormous success of the human brain, and try to argue that, to a lesser or greater extent, both contain within them the seeds of that misery and unhappiness that Freud spoke of.[1] One relates to the brain's capacity to acquire knowledge, to abstract and to construct ideals; the other relates to its variability. The study of the former is a philosophical burden which neurobiology has to shoulder if it is to understand better the workings of the brain; in discussing it, I shall try to project my conviction that the problems that neurobiology will face in the future are those of lasting truths and ultimate values which philosophy, the discipline that Bertrand Russell (1914) tells us 'has made greater claims and achieved fewer results than any other branch of learning', has so unsuccessfully tackled in the past. In this endeavour, neurobiology, like philosophy before it but probably with greater success, will also be naturally led to probe more deeply into areas that today may seem remote from its terrain, fields such as art, aesthetics and morality. The second factor, variability, is inextricably linked with the former. It is a factor, which, if understood better, may have profound consequences for regulating human affairs in society at large, and I shall also discuss it strictly in relation to our mental constitution.

In spite of the enormous advances made by neurobiology, it is still far from clear, at the turn of the Millennium, how we can go about understanding our mental constitution. It will almost certainly depend upon the development of new technologies. In this issue, Francis Crick has alluded to some of them, presently available but still not used to maximum advantage; others will be developed in the future, either serendipitously or through the demands made by new findings. Whatever these technologies may involve, it seems inevitable that they will be dictated, in part at least, by neurobiology's increasing tendency and need to understand our mental constitution and resolve the problems associated with it. Much energy will therefore be devoted to developing new techniques to study the human brain, a process that is already under way with the increasing refinement of brain-imaging techniques. These have had a large success in the past decade but are not, at present, capable of revealing the detailed neural mechanisms in play within and between different cortical areas. Whether the resolution of these problems, assuming there to be a resolution, will improve the lot of mankind on earth is another question. But there is little doubt that it will add immeasurably to natural knowledge.

2. THE PROBLEM OF KNOWLEDGE

Among the most prominent problems that I hope neurobiology will tackle, and which it has already done in a timid way, are the nature of knowledge itself and the relationship between knowledge and belief, given the exquisite capacity of the human brain, developed through the processes of evolution, to acquire knowledge. The problem of knowledge is of course also the problem of philosophy. Without tracing its long history in philosophy, which I am in any case not competent to do, and concentrating on its neurobiological import alone, it can be defined as learning something about the nature of the knowledge that we have or are capable of having and the extent to which that knowledge is determined uniquely by the functioning of the brain. A good point at which to begin is to suppose that one of the main functions of the brain is to acquire knowledge and that the only knowledge that we have at our disposal is brain knowledge. But this ushers in formidable problems, which neurobiologists no less than physicists share with philosophers. It amounts to learning the possibilities and limitations imposed on acquiring knowledge by the detailed neural organization of the brain and what, in Immanuel Kant's term, is the '*organizing principle*' of the brain in acquiring that knowledge (Kant himself spoke of the mind rather than the brain). One can begin this by discussing the problem of constancy and the nature of the contribution that the brain makes towards resolving it, choosing as an illustrative example a subject well beloved of philosophers, namely colour. This leads naturally to the question of knowledge derived from thought alone but based upon our knowledge of particulars, which in turn leads to a discussion of abstraction and the formation of Ideals. The last two are the necessary by-products of an efficient knowledge-gathering system, though I think that a substantial price is extracted in return for this efficiency. That I should devote so much space to a discussion of the problems associated with thinking and knowledge reflects my conviction that they are both problems that neurobiology is even now able to tackle technically. I am of course aware that, in the discussion that follows, I may have read back many present-day insights into past times, and been unnecessarily generous to past philosophers. The ideas expressed by Plato, for example, are often hesitant and uncertain and his questions can be much better formulated today. Kant is often incomprehensible and sometimes mystical; he writes about the mind, not the brain, and many of his questions can be better and more simply formulated in terms of present-day knowledge. But that we are better able to formulate the problems today, and perhaps even to find solutions to them, does not alter the fact that the general quest of neurobiology in the future, and of philosophy in the past, often have a common ground.

(a) *The problem of constancy*

For Plato, the Heraclitan doctrine of flux or constant change, the problem of that 'which forever becomes but never is' as Schopenhauer called it, constitutes the first problem for acquiring knowledge and thus the first problem of knowledge. Such an enquiry is in fact central for neurobiology, although Plato himself did not consider it in that context and did not refer to the brain. There was indeed little need for him to do so, since Ideals alone were the real thing and they belonged to the supra-sensible world, accessible through thought alone. His views as developed in *Phaedo*, *Meno*, *The Republic* and *Theaetetus* begin with an exposition that nevertheless could have been neurobiological. His starting point is that the only knowledge that is worth acquiring is knowledge about the constant and non-changing properties of all that is in this world. Herein lies a general problem for neurobiology and philosophy alike, for this knowledge has to be acquired in a world where everything is continually changing. This is the problem of constancy. Psychologists and neurobiologists generally speak about such categories as form or colour constancy. But constancy should have much wider application; it should, for example, be equally applicable to situations (situational constancy) and to more abstract concepts such as honour and justice. The latter are problems which neurobiology has not yet addressed, although it will be surprising if it does not do so within the coming century. For the present, to illustrate the general problem of the neurobiology of knowledge, it is perhaps instructive to turn to the problem of colour. Though somewhat restricted, philosophers commonly referred to colour in discussing knowledge, and asked questions such as, do colours exist in the material world and can colours be considered to be the properties of objects? These are questions to which neurobiology has provided some answers, and will no doubt provide more in the future. It illustrates in miniature form one of the great problems of both neurobiology and philosophy, that of acquiring knowledge about one unchangeable property of an object or surface when all the information reaching the individual from that object or surface is in a continual state of change. It also illustrates well, though in one restricted domain, Kant's belief, stated in his monumental *Critique of pure reason* (Kant 1781, 1787), that in enquiring into our knowledge of the world we should also enquire not only into the limitations imposed by the nature of the human mind (for which read brain), but also into the extent to which knowledge depends upon the formal contribution of the mind (for which again read brain).

(b) *The construction of colour by the brain*

It is well known that the amount of light of any given waveband reflected from a surface changes continually, depending upon the illuminant in which it is viewed. Yet the colour of a surface remains the same, even though there will be changes in shade. The stability of colours is to be sought in the capacity of the brain to undertake an operation which makes it independent of the continual change in the wavelength composition of the light reflected from a surface, and hence of servility to the constant vicissitude of things; this in turn allows the brain to obtain knowledge about a certain property of surfaces in spite of continual variations in what reaches the eye from those surfaces. The world would be a strange place if the colour of a surface changed with every change in the wavelength composition of the light reflected from it; we would no longer be able to obtain knowledge about certain properties that they have and colour would cease to be an efficient biological signalling mechanism. The mandatory involvement of the brain, the notion that it is the brain that undertakes an operation, was actually well stated by the combative Arthur Schopenhauer (1854) in a book entitled *On vision and colors: an essay*, a work that caused 'no ripples'. This was to Schopenhauer's regret and surprise, 'because a more precise knowledge and firmer conviction of the wholly subjective nature of color contributes to a more profound comprehension of the Kantian doctrine of the likewise subjective, intellectual forms of all knowledge, and so it affords a very suitable introductory course of philosophy'. The book contained what Schopenhauer, with characteristic immodesty, supposed was the first theory of colour vision. It was inspired by Goethe's theory of colour but coolly received by the latter who 'kept the manuscript longer than I had expected by taking it on his tour . . . of the Rhine' (my ellipsis). The importance of this work, unaccountably ignored by physiologists and philosophers alike, lies not in the details of the colour theory which Schopenhauer produced (he supposed that the operation was entirely undertaken in the retina). Rather, it is in the supposition that colours lie in the observer and not outside him, a fundamental insight which Schopenhauer believed neither Newton nor Goethe had understood, accusing the former especially of supposing that colour is 'inherent in these rays as a *qualitas occulta*, according to laws independent of the eye'. To Schopenhauer (1854), colour may be an immediate percept, and therefore intuitive, but 'All intuitive perception [*Anschauung*] is intellectual, for without the understanding [*Verstand*] we could never achieve intuitive perception', thus implying a formal contribution of the brain to the *Anschauung*. In the second volume of his *Die Welt als Wille und Vorstellung*,[2] Schopenhauer says that 'perception is not only the source of all knowledge, but is itself knowledge . . . it alone is the unconditionally true genuine knowledge' (my ellipsis). I therefore interpret his use of the word *Verstand* to mean the operation undertaken by the brain and leading to the intuitive perception [*Anschauung*], since Schopenhauer tells us that 'the forms underlying *Verstand*, the *Verstand* is a function of the brain'. This is of course unfair to Newton. In his book *Opticks*, Newton (1704) had written that 'For the Rays to speak properly have no Colour; in them there is nothing else than a certain power and disposition to stir up a sensation of this Colour or that', thus implicitly acknowledging the contribution of the brain. But, with hindsight, it would have been better if Newton had written that the 'Rays contain a certain power and disposition to allow the brain to stir up a sensation of this Colour or that'. Maxwell (1872) expressed this more clearly, though without reference to Schopenhauer, when he wrote 'If colour vision has any laws, it must be something in our nature that determines the form of these laws. The science of colour must therefore be regarded as being an essentially mental science'. Read in the Kantian sense, this statement is explicit in saying that the mind makes a formal contribution to the perception of colour and thus to knowledge about a certain property of surfaces which the brain interprets in terms of colour. Here again, I may be overgenerous to previous thinkers. Their failure to make the link between mind and brain is not a trivial oversight. It is that failure which, I believe, prevented them from asking better-tailored questions and hence impeded progress in the field.

(c) *The formal contribution of the brain to the construction of colour*

What is that formal contribution? The fact that the colours of objects and surfaces do not change appreciably even when the wavelength composition of the light in which they are viewed (and thus the wavelength composition of the light reflected from them) changes markedly is known as colour constancy. We have a rough idea of the kind of operation that the brain must undertake to make itself independent of changes in wavelength composition in assigning colour to a surface, and thus a rough idea of its formal contribution to constructing colours. It consists in determining the ratio of light of all visible wavebands reflected from a surface and from its surrounds, the ratio remaining always the same (Land 1974). The ratios are taken by the brain and the end result of that ratio-taking process, colour, is a property of the brain, not the world outside. Through this operation the brain acquires knowledge about a certain property of objects. But this knowledge is not of the colour (since bodies have no colour) but of *reflectance*, namely the efficiency with which a surface is capable of reflecting light of different wavebands. Hence, though we perceive colour as a property of objects, colour is really the *interpretation* that the brain gives to that physical property of objects (their reflectance), an interpretation that allows it to acquire knowledge rapidly about the property of reflectance. And that interpretation can be put to good use, a favourite example being the colour of a fruit as an indication of its ripeness and therefore readiness to consume. By saying that the brain makes a formal *contribution* towards acquiring knowledge about the properties of surfaces, we are acknowledging the fact that such knowledge is also determined in part by the physics of both light and the surfaces reflecting it. It is in fact an acknowledgement of Kant's (1783) statement in his *Prolegomena* that 'The Mind does not derive its laws (*a priori*) from nature, but prescribes them to her'. But, although we know that, to construct colours, the operation consists of ratio taking within strict limits defined by the spectral sensitivity of the photoreceptors in the retina, and although we have learned a great deal about the pathways in the brain that are critical for colour vision, we are very much ignorant of the detailed neural implementation of that formal contribution, of how the cells in the critical areas of the brain undertake the ratio-taking operation. This deficiency in our knowledge is not one that I expect to be with us for long. The development of powerful human brain-imaging techniques has already allowed us to define the brain centre, the V4 complex, in which the ratio-taking operation occurs (Zeki *et al.* 1991; Bartels & Zeki 1999). The development of similar techniques in the macaque (Logothetis *et al.* 1999), coupled to single-cell recordings from cortical cells, should theoretically reveal a good deal about the detailed neural procedure that the brain uses to undertake such an operation.

That the V4 complex, a relatively small area lying in the fusiform gyrus of the brain, is critical for the construction of colour becomes obvious from the description of clinical cases of patients rendered cortically colour blind (achromatopsic) by lesions there (for a review, see Zeki 1990, 1993). Such patients are only able to see the world in shades of grey; their world is largely monochrome and, unlike normals, they are therefore not able to obtain certain kinds of knowledge about this world by interpreting the reflectance properties of surfaces in terms of colour. It is notable that the visual apparatus from the retina, up to (but not including) the critical brain area for constructing colours, can be normal in such patients (Mollon *et al.* 1980). It is as if a certain concept, that of ratio taking, must be applied to the otherwise normal visual signals that enter the brain. This gives substance to Kant's (1781) view in the *Critique of pure reason* that 'perceptions without concepts are blind'. The concept here is that of ratio taking, very strictly a brain concept and applied by it to the incoming signals. This concept of ratio taking constitutes the brain's formal contribution to knowledge about the world of colour. It is worth emphasizing that it is the transition from the intangible mind, which Kant and other philosophers wrote vaguely about, to the tangible brain that allows us to formulate many of the questions in more precise terms, not only conceptually but in terms of what constitutes the formal contribution of the brain in neural terms, that is the activity of single cells and the circuits of which they are part.

(d) *Innate intuitions*

Viewed in the context of our present knowledge, the questions posed by philosophy should often have been of profound interest to neurobiologists. That they have not managed to capture the imagination of the latter is due in part to the imprecise and hesitant way in which they were formulated; it is also due in part to the fact that the philosophical formulations have often been cast in too broad and unmanageable a form for neurobiological investigation, at least with presently available technologies. Thus the general question that exercised Kant, namely what formal contribution the mind makes to knowledge, is not readily answerable in neurobiological terms today, in spite of its profound interest, even if one were to suppose a unity of knowledge, which is doubtful. But when made narrower, the question falls well within the possibilities of neurobiological investigation. Returning to the more restricted example of colour vision, Herman von Helmholtz (1911) supposed that assigning a correct colour to a surface is done by a vague process referred to as the 'unconscious inference', which he supposed allowed one to judge what a surface would look like in white light. In this he followed Leibniz (1714), who in his *Monadologie* had written that it would be difficult to imagine that the reality behind all natural phenomena is consciously extracted by the mind. This led him to the concept of the unconscious mind, an important concept that neurobiology is beginning to tackle, but in terms of the brain not the intangible mind. We would now say that the operation that the brain undertakes to construct colour is indeed an unconscious and instantaneous process, one of which we are not normally aware, but the result is almost instantaneously available consciously as a form of knowledge. The question, neurobiological no less than philosophical, is the extent to which the (unconscious) processes involved in constructing colours are intuitive and the extent to which they are dependent upon experience. Hence the task for neurobiology here is twofold: to define the neural processes involved in the construction of colour, and to probe into the question of what neural activity leads from the unconscious process of colour construction to the conscious experience of colours.

Philosophers have often asked about the extent to which the knowledge that we acquire, or are capable of acquiring, depends on experience and to what extent on *a priori* intuitions, which are independent of experience. Kant, like Schopenhauer after him, believed that space and time were the two *a priori* intuitions around which all knowledge is organized, although Schopenhauer added the principle of causality as another innate *a priori* intuition. To them, experience had to be read into these *a priori* intuitions. Henri Poincaré (1894, 1898) believed that the *a priori* intuitions that constituted the organizing principles through which sense data were converted into knowledge were mathematical induction, which is 'inaccessible to analytical proof and to experiment', and the intuition of continuous groups, which 'exists in our mind prior to all experience'. The notion of the *a priori* organizing principle was in fact contested by Einstein (1916*b*) who, in his obituary of Ernst Mach, complained about philosophers who have had 'a harmful effect upon scientific thinking in removing certain fundamental concepts from the domain of empiricism, where they are under our control, to the intangible heights of the *a priori* . . . This is particularly true of our concepts of time and space'.

How can we define *a priori* intuitions in neurological terms? Are there many *a priori* intuitions, relating to different kinds of knowledge, or only a restricted number of such intuitions whose operation is valid for acquiring knowledge in all domains? These are questions that neurobiology is able to provide more or less vague answers to today, but only at a certain level of enquiry, not at the grand level of universal knowledge. More precise answers will follow, I hope, in the future. One might say that the innate intuition in colour vision is the product of the neuronal wiring that allows the brain to undertake a certain formal operation on the incoming signals, comparing the amount of light of every waveband reflected from one surface with that reflected from that of surrounding surfaces. We at present have no ready answer to whether the details of that neuronal wiring are dependent upon visual experience. But experiments in another area of vision, form, have given us compelling answers about the relative roles of innate and experiential factors. Physiologists have studied in some detail what are known as orientation-selective cells, cells that respond to lines of particular orientation and are unresponsive to lines of other orientation (Hubel & Wiesel 1977). These cells are generally considered to be the physiological 'building blocks' for the perception of forms. Are these cells genetically specified or does their mature behaviour depend upon the acquisition of visual experience? This constitutes a neurological investigation, at the level of the single cell, of the question asked by the learned Mr Molyneux in John Locke's *Essay concerning human understanding*. Molyneux had asked whether a man born blind and who had learned to distinguish between a cube and a sphere by touch alone would be able, if vision were restored to him, to distinguish between the two by sight alone. Studies of patients born blind because of congenital cataracts and who had vision 'restored' to them at various stages after birth have shown that such patients are forever blighted visually (Von Senden 1932). The neuroanatomical and neurophysiological counterparts of these experiments have shown that, at birth, the pathways from retina to the brain are genetically determined and ready to function (Rakic 1977), but that early visual exposure during a critical period is essential for the later normal functioning of the visual brain and for its capacity to acquire knowledge through vision (Hubel & Wiesel 1977). Thus the innate contribution is an anatomically organized system which has to be nourished immediately after birth. And lest this powerful conclusion be thought to be restricted to vision, and to form vision at that, it is salutary to compare the conclusion reached by Hubel & Wiesel with the conclusions reached by Sigmund Freud who, in the realm of complex behaviour, traced individual histories backwards in time, and Harry Harlow who traced them forwards in time. Both established the importance of early experience for the later behaviour of an individual. Harlow & Harlow (1962) wrote, 'There is a critical period somewhere between the third and sixth months of life during which social deprivation . . . irreversibly blights the animals' capacity for social adjustment', while Hubel & Wiesel (1970) wrote that 'The effects of [visual] deprivation for the first three months of life tend to be permanent, with very limited morphological, physiological, or behavioural recovery'. Hence, we might say that a general organizing principle in the acquisition of knowledge is an innately determined neural wiring, which is left hostage to experience at critical stages after birth.

Kant would no doubt have been deeply dissatisfied with this answer, because it does not really address properly the question of the innate *a prioris* of space and time, which he formulated in general terms as applicable to all perceptions, and into which all perceptions are read, so to speak. Of space, he wrote in section I of the *Critique* (second edition) entitled *Transcendental aesthetic*: 'Space is not an empirical concept that has been abstracted from outer experiences. For the presentation of space must already lie at the basis in order for certain sensations to be referred to something outside me (i.e. referred to something in a location of space other than the location in which I am). And it must similarly already lie at the basis in order for me to be able to present [the objects of] these sensations as outside and *alongside* one another, and hence to present them not only as different but as being in different locations. Accordingly, the presentation of space cannot be one which we take from the relations of outer appearance by means of experience; rather, only through the presentation of space is that outer experience possible in the first place' (original emphasis) (Kant 1787). In section II of the *Transcendental aesthetic*, he wrote of time: 'Time is not an empirical concept that has been abstracted from any experience. For simultaneity or succession would not even enter our perception if the presentation of time did not underlie them *a priori*. Only on the presumption of this presentation can we present this and that as being at one and the same time (simultaneously) or at different times (sequentially)' (Kant 1787).

The force of these arguments and their importance are difficult to deny, and they are questions that neurobiology should be in a position to frame, and possibly to study. It is perhaps regrettable that we have not already done so in a serious way but, once again, it is very difficult to do so in the more general and universally applicable terms in which Kant cast his argument. We can, however, illustrate the neurobiological importance of these arguments by casting the same question within a more limited sphere. With colour vision, space is of paramount importance, since the brain has to determine the amounts of light of a given waveband composition reflected from a surface located in one region of space and from its surrounds, located in a different region, the whole being located outside the subject. There are therefore two components to the space problem here. Spatial distribution of surfaces is critical for the computation of colour, and it is difficult to see how this part of the ratio-taking program could not have been given *a priori*, and implicit in the genetically determined program that makes such wiring possible. But it is not at present clear how to establish this experimentally. We have a vague insight into the second component, of referring a surface to a location outside the subject, from a single early study in colour vision, quoted by Critchley (1965). That study described how colours occupied different layers in the personal space of a congenitally blind subject to whom vision had been 'restored'. If confirmed, this would suggest that the ordering of space does depend on experience, but still does not address the question of the concept of space as being given *a priori*, since the colours were correctly located outside the patient although, unlike normals, they occupied different positions in extra-personal space.

The second Kantian *a priori*, that of time, is no less important for colour vision. In constructing colours, the brain must determine simultaneously the wavelength composition of the light reflected from a surface and that reflected from the surrounds, which implies that a timing mechanism must be built into its ratio-taking program. Moreover, when the same scene is viewed in a different illuminant, entailing a change in the wavelength composition of the light coming from all points in the field of view, the ratios have to be taken successively for the change in the illuminant to be discounted. Once again, one supposes that a timing mechanism is built into the brain's ratio-taking program.

We really have no answers to these questions today and do not even know how often the ratios are taken, even in steady conditions (see Bartels & Zeki 1999). But the general discussion I have given above is probably applicable to other aspects of visual perception, and indeed to all perception; they all involve relations of space among the constituents and the ability to locate the objects of perception outside the subject, and they all involve simultaneity and succession in time.

(e) *The modularity of knowledge*

I am well aware that philosophers might well regard the way in which I have restricted the questions that they have asked as a travesty that departs greatly from the undeniable grandeur of their formulations. In response, one can only lament the poverty of the results which this grand manner of formulating the problems has achieved in terms of understanding our brains and their mental constitutions. I restrict questions, because if we are to face the enormous challenge posed by these questions, we have to render them more manageable to experimentation. Not only is the question of a general organizing principle and its *a priori* foundations in formulating all knowledge too broad, but it also assumes a unity of knowledge. There may be grounds for such an assumption, but that unity is perhaps better sought after, not in terms of knowledge as such, but in terms of abstraction and the formulation of ideals, discussed below. Assuming there to be an organizing principle capable of transforming the sense data into knowledge, the philosophical question can be summarized like this: What is the 'organizing principle' that is necessary for acquiring all knowledge; to what extent does it 'pre-exist' within us; and to what extent does it depend on, or is nourished by, experience? In the light of present evidence, the neurobiological question cannot assume the unity of knowledge as its starting point. It must begin by asking: Is there one or are there many different kinds of organizing principle, each one tailored to the acquisition of a particular kind of knowledge (say about shapes, or colours, or geometrical relations)?, and then proceed to define the organizing principle in neurobiological terms. In terms of colour vision, this would amount to learning how the cells of the V4 complex are organized to register the amount of light of different wavebands reflected from different surfaces and compare them to obtain ratios, emphasizing especially the factors of space and time. It is in fact astonishing for those who believe in the unity of knowledge that such a relatively small part of the brain should be the seat of a relatively well-understood operation which leads to the acquisition of knowledge about a certain important property of the visual world, without affecting the capacity to acquire other kinds of knowledge. Its importance lies in showing that the kind of knowledge that the brain is capable of acquiring in one area of vision is strictly determined by the operational logic of the brain (located in a restricted part of the brain). Assuming that, through the operation of its logic, the brain makes an inference about certain physical properties of surfaces, interpreted as colour, it seems difficult to believe that it uses the same inferential method to deduce, for example, the expression on a face or the appearance of an object, properties which must rely on other kinds of operation. Indeed, the history of neurophysiology has taught us that the acquisition of knowledge is a distributed process, with different cortical areas contributing in specialized ways and according to an anatomical and functional organization tailored to the needs of acquiring different kinds of knowledge. The destruction of the colour centre in the brain (the V4 complex) deprives the subject of the ability to acquire knowledge about one property of the visible world, without necessarily compromising the ability to acquire other kinds of knowledge; similarly, a lesion in V5 compromises the ability to acquire knowledge about objects when in motion, without at the same time interfering with the ability to acquire knowledge about the same objects when they are stationary. And so the list goes on. Moreover, psychophysical experiments show that different processings systems, located in geographically distinct areas of the cerebral cortex, reach their perceptual end-points at different times. We obtain knowledge of colour 40 ms before we obtain knowledge about form and 80 ms before we obtain knowledge about motion (Moutoussis & Zeki 1997). Hence the acquisition of knowledge is a process that is distributed across both space and time. Viewed in this context and at this level, it is difficult to assume a unity of knowledge.

The examples given above naturally raise the Cartesian question in a reverse sense: can we say that what we have learned about colour vision is applicable outside this narrow though important field? Today's answer is that it is not at all clear how we can do so, given the apparent modularity of the knowledge-acquiring system of the brain, at least at an elementary level. It is far more difficult to deal neurobiologically with the acquisition of knowledge in mathematics and geometry, areas emphasized heavily by philosophers and mathematicians. But the question remains highly interesting: what formal contribution does the brain make to the acquisition of knowledge through, and in, these other fields? Allied to this is another problem which, as far as I know, neurobiology has not addressed at all. Are there any formulations in mathematics and geometry that are not dependent upon the organization of our brains? Put otherwise, if our brains had a different organization would we have been able to devise mathematical theories that are substantially different from the ones that we have today? And would such a different neural organization be capable of formulating the theories that our brains have? René Descartes supposed that the one certain way of attaining knowledge was through mathematics. For him, the method of mathematics resolved around the twin methods of intuition and deduction. Is that intuition, a neurobiologist might well ask, anything more than the intuition of a brain, or of brains, with a certain kind of neural organization, and does that neural organization not determine as well the kind of deductions that we are capable of? There was of course also the problem of whether the secure way of obtaining knowledge in mathematics could be applied outside mathematics, to grander philosophical questions of what does and does not exist in this world; in other words, to the problem of knowledge at large. If we had had greater insights into the brain's capacity, in neural terms, to formulate mathematical theories, could we apply those insights to its capacity to construct colours? It is interesting here to note the arguments of Roger Penrose. For him, there is nothing any more special about mathematical understanding, compared to any other form of understanding, implying that the same qualitative brain processes are involved in all forms of understanding. But while admitting that there might be mathematical truths 'that remain, in principle, inaccessible to human reason and insight', Penrose (1994) takes a distinctly Platonic, as opposed to Kantian, view and believes that the mathematical world is not a product of our thinking. Its 'existence rests on the profound, timeless, and universal nature of these concepts, and on the fact that their laws are independent of those who discover them ... The natural numbers were there before there were human beings, or indeed any other creature here on earth, and they will remain after all life has perished' (my ellipsis), which is clearly not something that can be said of colour.

(f) *Knowledge by pure thought*

If, as seems likely, the acquisition of knowledge is a distributed cortical process, with different kinds of knowledge requiring different kinds of neural machinery, then where can we look for that unity which is characteristic of all knowledge, assuming there to be one? I suggest that one possibility may lie in our capacity to abstract and formulate ideals, characteristics that are common to all knowledge and indeed the defining features of an efficient knowledge-acquiring system. We can then suppose that, even though the brain uses different kinds of neural machinery for acquiring different kinds of knowledge, it nevertheless applies a further and similar neural process repetitively on different kinds of knowledge, either in the same or in different cortical areas. Indeed, given the impressive structural uniformity of large parts of the cerebral cortex, neurobiologists have addressed the question (though without finding a satisfactory solution) of whether there is any common process that the brain applies repetitively in all cortical areas, that is across all knowledge-acquiring systems of the cortex. It is interesting to speculate that these two brain capacities may be the result of a similar kind of neural process, repetitively applied. I introduce these two abilities by discussing the problem of knowledge arrived at by pure thought. The course of the argument I take is perhaps different from the one pursued by philosophy.

The example of colour vision is one derived from the sensible world, which has played a leading role in the formulation of philosophic doctrines about knowledge. Indeed the German word *Anschauung* refers etymologically to visual sensations only, but Kant extended its usage to apply to all perception (see Miller 1984). Plato and his successors supposed that some of the ideas we have could not have been derived from the sensible world; notable among these were mathematics and geometry. This led him to suppose that there is another world of knowledge, one that is accessible only to the intellect, a view shared by G. W. Leibniz and Christian Wolff. The Leibniz–Wolff formulation was strongly contested by Kant (1787), who wrote that 'the philosophy of Leibniz and Wolff, by considering the distinction between what is sensible and what is intellectual as a merely logical one, has imposed an entirely wrong point of view on all investigations about the nature and origin of our cognitions'. Kant emphasized the extent to which perception acting on the innate *a priori* intuitions is the source of all knowledge and supposed, more subtly, that it is possible to arrive at general truths about entities in this world, which we cannot experience but which depend on experience, and therefore to acquire knowledge by the process of pure thought alone. It is indeed remarkable how extensive a use of visual perception and imagery some of the great physicists and mathematicians, like Henri Poincaré, Ludwig Boltzmann and Albert Einstein, have made in formulating their theories (Miller 1984). But what does pure thought consist of, and what are its limits? And what, in neurological terms, constitutes the process of thinking? It is, in a sense, astonishing that neurobiologists have not addressed these questions formally, although the capacity to do so technically is there, at least in a rudimentary way, and even though its general importance was outlined many years ago by Einstein (1916*a*), who, in an article entitled 'Physics and reality', wrote: 'The whole of science is nothing more than a refinement of everyday thinking. It is for this reason that the critical thinking of the physicist cannot possibly be restricted to the examination of the concepts of his own specific field. *He cannot proceed without considering critically a much more difficult problem, the problem of analyzing the nature of everyday thinking*' (my italics). In fact, the question about knowledge in the absence of perception, knowledge derived from thinking, runs throughout philosophical speculation, from Plato through Leibniz, Locke, Berkeley and Hume to Kant and Schopenhauer. It can be phrased as follows. Is it possible to arrive at knowledge by thought alone, and what kind of knowledge would that constitute? In fact, we have to go beyond and ask: are the 'unconscious inferences' (Leibniz and Helmholtz), which lead, for example, to the construction of colour, subjected to a further thought process through which alone we can obtain knowledge?

In the example of colour vision, ratio taking is but one formal contribution that the brain makes towards the construction of colour, and that to the construction of colour in the abstract alone (see § 2(g)(i)). Ludwig Wittgenstein (1977), in his somewhat repetitive and tedious book entitled *Remarks on colour* ('That which I am writing about so tediously, may be obvious to someone whose mind is less decrepit'), nevertheless raised interesting questions about the contribution that the brain makes, through language, to the concept of colour. Why, for example, do we speak of a 'dark red light' but not of a 'black-red light'? Why can we not imagine a grey-hot? Why can't we think of it as a lesser degree of white hot? Why do we often speak of white as not coloured? Why is transparent white impossible? And so on. He wanted, he said, not so much a theory of colour but rather of the logic of colour concepts, and found it difficult to understand logically the concept of three or four primary colours (just as I do). Moreover, 'If there were a theory of colour harmony, perhaps it would begin by dividing the colours into different groups and forbidding certain mixtures or combinations and allowing others; and, as in harmony, its rules would be given no justification'. Although it is obvious that the brain makes an important contribution to colour concepts, through the neurological processes underlying both thinking and language, neurobiology has made no inroads into the relationship between the construction of colour by the cortex and the formal contribution the brain makes to the concept of colour through language, even though it is quite clear that a specific area of the brain is critically involved in naming colours, damage to which impairs the ability of the patient to name colours (Lewandowsky 1908). We also have not addressed the broader question of whether concept formation in general depends on the same or different kinds of neural processes, given that knowledge itself appears to be cortically distributed and that the kind of concept we may formulate in the world of colour may be substantially different from the one we formulate in the recognition of facial expressions or in mathematics.

(g) *The unity of knowledge*

Implicit in the thinking of philosophers, then, is a unity of knowledge. This stands in contrast to the discoveries of neurobiology, which have shown that the acquisition of knowledge, not only of particulars but also of particular categories, is a function of different parts of the cerebral cortex, with the consequence that the acquisition of knowledge is a cortically distributed process. But this demonstration does not exhaust our search. Pure thought can be applied to any of a number of perceptions and, if knowledge can be derived by pure thought alone (manifestly so in the world of physics, mathematics and colour concepts), one can well ask whether there is any similarity in the kind of thought process that is applied to different domains of knowledge. Two factors that come to mind are abstraction and the formulation of ideals.

(i) *Abstraction*

Abstraction is a difficult term to define, but there is little doubt that our brains are heavily involved in the process and that the brain's system of knowledge is heavily dependent on it. I shall here define it in two ways, while conceding that neither definition is satisfactory. Each implies a different neurological process, though with a common substratum.

Neurologically, I use the term abstraction to mean the capacity of the brain to find some property or relation that is common to many particulars, thus making the brain independent of the particular. We do not know what neural processes lead to it. Whatever they may be, there is little doubt that abstraction is a critical step in the efficient acquisition of knowledge, for without it the brain would be enslaved to the particular. The capacity to abstract is probably also imposed on the brain by the limitations of its memory system, since it does away with the need to recall every detail. Memory is a critical step in obtaining knowledge but, as Descartes saw, it could not be trusted in an unqualified way, even within the exalted and certain world of mathematics. A cardinal step in mathematical formulations is the deductive one. But in deductive arguments, the mind has to rely on memory in tracing the earlier steps in a deductive process. And there is no guarantee that the memory process itself may not be at fault (as indeed it often is).

We are of course able to abstract in different fields, from the world of mathematics on the one hand to the equally important world of our daily visual experience on the other. Given the profound functional specialization of the visual brain, with different visual areas involved in the processing of different kinds of visual signals, and each contributing to perception and therefore to knowledge in its specialized sphere, it is worth asking whether each area has its own mechanism for abstraction or whether abstraction is a more specific property conferred on a given area or areas. My strong prejudice in favour of functional specialization, dictated by experimental results, inclines me towards the former supposition while admitting that there is no present evidence in its favour. Whatever the answer may turn out to be, it is obvious that the same eliminative process must be applied in every case of abstraction, and hence a general feature in the unity of knowledge may be supposed to be that of abstraction.

Another definition of abstraction, applicable especially, but not exclusively, to art, is the process by which the product does not represent or symbolize objects (non-iconic abstraction). This kind of abstraction was taken in a different direction by Piet Mondrian, who sought in his art to find and represent the constant elements in nature as regards form, ones which would be constituent of all forms. He wrote: 'Art shows us that there are also constant truths concerning forms' and considered that it was the function of art to 'discover consciously or unconsciously the fundamental laws hidden in reality' (Mondrian 1937). He had been much attracted to Cubism but 'Cubism did not accept the logic of its own discoveries. It was not leading abstraction towards its ultimate goal, the expression of pure reality ... To create pure reality plastically it is necessary to reduce natural forms to the constant elements' (my ellipsis).

This kind of abstraction shares a similarity with the first in that it is also eliminative. In terms of colour vision, we can ask whether the brain can construct colour in the abstract, that is without assigning it to a particular recognizable object. Recent imaging experiments on the human brain have shown that the activity produced by viewing a multicoloured abstract composition with no recognizable objects (a Mondrian) is limited to the V4 complex, the area that is critical for constructing colours, and the areas that feed it. By contrast, when colours are the properties of recognizable objects, additional areas located in the temporal and frontal lobes are recruited (Zeki & Marini 1998). Other experiments show that abstract forms arranged by arbitrarily assembled straight lines activate more restricted areas of the brain than do the same lines when rearranged to form recognizable objects (L. Marini and S. Zeki, unpublished results). One would thus presume that the abstract involves the activity of a more limited set of areas than the non-abstract. Perhaps one conclusion to be drawn from this is that more restricted areas may contribute to the process of abstraction, but no one knows through what neural mechanisms.

(ii) *The construction of ideals*

The process of abstraction leads naturally to the formation of ideals. Plato used the term 'Ideal' to mean a universal by contrast with particulars, derived from the intellect alone. To him, only Ideals had a real existence. Beauty, for example, is an Ideal to be contrasted with particular things that may appear beautiful in one viewing and not so in another or beautiful to one person but not another, a mutability that implies a variability not only in one individual with time, but also between individuals. Hence the intimate link between the problems created by the construction of ideals and by variability. Kant (1781) tells us in the *Transcendental dialectic* of his *Critique* (first edition) that Plato meant by the term 'something that not only is never borrowed from the senses, but that far surpasses even the concept of understanding ... inasmuch as nothing ever congruent with it is found in experience', adding that 'I do not here want to ... establish what meaning the august philosopher linked with his expression. I shall point out only that ... by comparing the thoughts uttered by an author on his topic we understand him even better than he understood himself, because he did not sufficiently determine his concept' (my ellipsis). Perception, which is the source of all knowledge, is necessarily that of the particular at any given time. But this, in the Platonic system, leads to a superficial impression and an opinion at best, because of the Heraclitan doctrine of constant change. Perception is always of things 'which always become and never are', while to acquire knowledge about entities we have to learn about things which 'always are but never become and never pass away' (Schopenhauer 1859*a*). This constitutes the Ideal in the Platonic system and the 'thing-in-itself' in the Kantian one.

For Kant, the mind itself determines the form in which reality appears to us and the thing-in-itself (loosely, the Platonic Ideal) must remain forever unknown to us. In fact, as Schopenhauer (1859*a*) pointed out, the Platonic Ideal and the Kantian 'thing-in-itself' share substantial similarities; both concepts arose naturally from the ephemeral nature of the sensible world, the Heraclitan concept of flux. The formation of ideals is necessarily linked to that of abstraction, and may indeed be an extension of it since it is difficult to believe that the capacity to form ideals—with its indifference to the particulars on which it ultimately depends—could be achieved without the capacity for abstraction. These Ideals, Plato and Kant believed, belong in the supra-sensible world, are arrived at by the process of pure thought and are alone capable of giving knowledge about the world.

There is in fact much neurobiological merit in the notion of a supra-sensible world of knowledge, though perhaps not as intended by Plato and Kant. That we do form ideals in different areas of experience cannot be doubted, nor can the fact that these ideals often do not correspond to particulars either in our past or present experience. In fact, ideals and their formation are fundamental properties of the brain, which neurobiologists will have to study. Through what neural process are ideals formed and where? Philosophers and neurobiologists alike would agree that experience and exposure, and therefore also memory, are critical. One might add that the process of pure thought somehow combines the experiences. Indeed, how can we define the Platonic Ideal or the Kantian 'thing-in-itself' in neural terms? One answer I have given (Zeki 1999) is unsatisfactory, its sole merit lying in focusing attention on the fact that the formation of ideals is a neurological problem. I defined the Platonic Ideal of an object as the brain's record of all the objects that it has ever seen, dependent on a multitude of particular views, and from which it synthesizes an 'ideal'. Though inadequate in terms of knowing what neural mechanisms are involved, there is an interesting experiment in this regard. Logothetis and his colleagues (Logothetis *et al.* 1994) have studied this process of abstraction in single-cell recordings from the monkey brain. They have shown that when monkeys are exposed to different views of new forms generated on the computer, forms which they have never seen before, most cells in the temporal lobe discharge to one view only and their response declines as the object is rotated in such a way as to present increasingly less familiar views. But a small proportion of cells, amounting to about 1%, respond in a view-invariant manner, amounting to a process of abstraction. Moreover, for all, the actual size or precise position of presentation in the field of view is irrelevant, once again amounting to a process of abstraction. We can therefore say that exposure to many different views allows a minority of cells to become indifferent to the particularities of view. One supposes that by such neural procedures ideals are created.

A question that one can ask is whether there is a unique process in the brain that is applied to the formation of all ideals, or whether the formation of ideals is also a distributed system. In the light of present evidence, one would argue for the latter, though without much certainty. It is hard to believe that a patient with a lesion in V4 can form ideals of colour or that a patient with a lesion in the face-encoding area (in the fusiform gyrus) can form ideals of faces, and so on. Yet it is difficult to deny that a similar process, perhaps duplicated in many cortical areas must be involved. For in essence, abstraction and idealism involve, through a neural process that we know nothing about, the selection of some features (the most important one to the subject), the discarding of others and a synthesis of selected features, again by a neural process of which we are ignorant, into a new entity related to the individual experiences but indifferent to any one particular experience.

Herein lies one cause of unhappiness. The brain selects features which it deems to be of importance, 'throwing off ... useless luggage for the purpose of handling more easily the knowledge to be compared and manoeuvred in all directions' (Schopenhauer 1859*b*). Because of variability between individuals, what one brain deems important and therefore selects is not necessarily the same as another brain. Because of variability over time, what one brain deems important at one time, and therefore selects in the formation of ideals, is not necessarily what it will select at a later time. It is from those selected features that the brain builds an ideal, again by a process about which we know nothing. Whatever the neurological process that leads to that formation, it commonly departs substantially and disappointingly from the particulars from which it is derived in past experience and to which it is applied in present and future experience. In practice, this is not always a real problem for an individual. Most of us, including even mathematicians, probably do not think much about whether the straight line that we have drawn with a ruler departs significantly from the ideal, the thing-in-itself, of the straight line, arrived at by pure thought. Neurobiologically, the latter can only mean the synthesis of all the straight lines that we have experienced, unless we are to accept quite literally Plato's belief of the existence of such entities (straight lines) independent of the perceiving mind. But there are many other instances in which the ideal, as constructed by the brain from its experience, departs significantly from the particular and can therefore be a source of great disappointment. This is true of human relations, of achievements, of ambition and much else that is cardinal in human conduct. Moreover, ideals created by the brain change with time and with experience, making it doubly difficult to tailor a particular to the more general, and abstract, creation of the brain. To this clash between ideals and particulars, between knowledge of subjects as ideals and knowledge of them as particulars, one can trace the cause of much misery and disappointment. But this is the by-product of an efficient knowledge-acquiring system. If knowledge is based on perception, to which the brain makes its formal contribution, and to pure thought, to which the brain contributes exclusively, then abstraction and idealism, with all that they entail in terms of richness and misery, are the necessary by-products. I have sometimes wondered whether I should not substitute the Kantian *a priori* intuitions of space and time with the *a priori* intuitions of abstraction and idealism, innate capacities of the brain developed through evolution in the service of acquiring knowledge, the two great organizing principles into which all experience is read.

(h) *Ideals in art*

The frustration of achieving the ideal in daily life is common knowledge. It is through art that humanity has often tried to represent the ideals formed by the brain, which it often cannot encounter in particulars. I have argued elsewhere (Zeki 1999) that the function of visual art is an extension of the function of the visual brain, namely the acquisition of knowledge. Just like ordinary or exalted knowledge, art has to deal with the problem of constant change, of the Heraclitan flux. Hence Henri Matisse's (1978) statement that, 'Underlying this succession of moments which constitutes the superficial existence of things and beings, and which is continually modifying them, one can search for a truer, more essential character, which the artist will seize so that he may give to reality a more lasting interpretation' is one that both Kant and Schopenhauer would have approved of, as indeed should neurobiologists, though Plato would have disagreed. In Book X of *The Republic* Plato considered art to be lowly since it could only capture one view of a particular object, which is but one example of the more general 'Ideal' object. But, as John Constable (1771) once wrote, 'the whole beauty and grandeur of Art consists ... in being able to get above all singular forms, local customs, particularities of every kind ... [The painter] makes out an abstract idea of their forms more perfect than any one original' (my ellipses). This process, just as in the example of monkeys given above, depends on exposure to many different views of the scene to be depicted, as artists have repeatedly attested, and the synthesis from these views of what Hegel has called a 'Concept', through an elusive thought process. In an almost neurobiological statement, Picasso (1935) once said, 'It would be very interesting to preserve photographically . . . the metamorphoses of a picture. Possibly one might then discover the path followed by the brain in materialising a dream' (my ellipsis). For Hegel (1832), the Concept, which I again interpret to be the brain's record of all the scenes that it has become acquainted with, becomes the idea when transformed on to canvas. Through this process of externalizing and concretizing what is in the brain, art 'furnishes us with the things themselves, but out of the inner life of the mind'.

Perhaps one definition of great art would be the art that comes nearest to being a particular example of the synthesized concepts (in the Hegelian sense) in as many brains as possible. This is an enormous feat to achieve, given the great variability between brains in this regard and few have been equal to the task. One way of achieving this is through ambiguity, a characteristic that is highly prized. Vermeer, for example, instilled a mystery and an ambiguity in many of his paintings, but I use the term ambiguity here, not in the sense defined in dictionaries, but in a neurological sense. It is the opposite of uncertainty. Rather, it is the certainty of many different situations or conditions, each of which has equal validity to the others (Zeki 1999). Why should ambiguity be such a prized quality in art? It is, I believe, a reflection of the fact that it can fit many different ideals or concepts formed in different brains, with their variability in both neurological terms and in experience. In their book on *Cubism*, Albert Gleizes and Jean Metzinger (1913) wrote 'Certain forms should remain implicit, so that the mind of the spectator is the concrete place of their birth'. There can be no better description of the work of Vermeer or the unfinished sculptures of Michelangelo, where nearly all is implicit. It can thus be tailored to different brains, at different times. Indeed where particulars never correspond with the ideal formed by the brain, art (at least to an individual) may be the only outlet. Richard Wagner, after all, composed *Tristan und Isolde* as 'a monument to the greatest of all illusions, romantic love'!

3. THE PROBLEM OF VARIABILITY

Abstraction and idealism are thus the privileges of an organ which, through evolution, has developed an exquisite capacity to acquire knowledge through thought processes. Both carry within them a clash between experience of the particular and what the brain has developed from experience of many particulars and both therefore can lead to much disappointment in our daily lives. This disappointment is heightened by the fact that both are subject to variability in time within an individual and between individuals. If these two characteristics, in addition to providing huge selective advantages, also constitute one source of human misery, another critical feature inextricably linked with them in the process of evolution—indeed the bread and butter of evolution—is perhaps an even greater, and possibly the greatest, source of unhappiness. That factor is variability.

The study of Man has shown us that the general somatic variation between people of different cultures and races is trivial compared to the much greater variation in brains, even within an apparently racially and culturally homogeneous population. In *The origin of species*, Darwin (1859) emphasized without referring to the brain that 'When we see any part or organ developed in a remarkable degree or manner in a species, the fair presumption is that it is of high importance to that species; nevertheless it is in this case eminently liable to variation'. This variation is not readily detectable in the general macroscopic anatomical organization of the brain, its broad physiology or its blood supply. Differences between individuals are difficult to detect at this level, which is one reason why physiology has been so successful in charting so much about the way the brain functions. Modern brain imaging techniques have, in addition to charting the areas of the cerebral cortex specialized for different functions, given us a remarkable picture of how similar different brains are in their basic organization. Even the study of colour vision has not shown a great variability in the disposition of the colour centre in the brain, at least at the somewhat gross level. It would thus seem that there is, after all, a great basic similarity between different brains, an important lesson which is worth bearing in mind in discussing the relationship of brain activity to subjective experiences, and to art, beauty and aesthetics generally. Yet in spite of this basic similarity between one brain and another, the mental constitutions of individuals vary greatly, and a glance at human achievements in art and literature, no less than in science, is sufficient to convince one that variability gives our culture—the outward product of our brain—a huge richness and advantage.

The variation that is at the basis of Man's psychological misery lies elsewhere; it is observable behaviourally if not anatomically, and through the observation we infer a difference in brain organization at a level which has so far eluded even the most detailed microscopic examination. It lies in differences between individuals—in their sensitivity, in their intelligence, in their aptitudes, in their desires, in their aspirations and in much else besides, in brief in their mental constitutions. In Darwin's reasoning, this is a clear sign that evolutionary selection is continuing, that we are, as a species, continuing to change. He wrote in *The origin of species*, without reference to the brain, 'It is only in those cases in which the modification has been comparatively recent and extraordinarily great that we ought to find the *generative variability*, as it may be called, still present in high degree. For in this case the variability will seldom as yet have been fixed by the continued selection of individuals varying in the required manner and degree, and by the continued rejection of those tending to revert to a former and less modified condition' (original emphasis). This creates an acute problem for Man. Variability is a cherished quality, both in the evolutionary and in the cultural sense. Yet it is these very differences that society prohibits individuals from expressing, or prohibits them from expressing fully, for the greater good of all. Freud, ever ready to find the source of much misfortune in sexuality, wrote: 'The standard which declares itself in these prohibitions is that of a sexual life identical for all; it pays no heed to the disparities in the inborn and acquired sexual constitutions of individuals and cuts off a considerable number of them from sexual enjoyment, thus becoming a cause of grievous injustice' (Freud 1930). But sexuality, though perhaps the most pervasive, is not the only area in which variability, often prohibited, expresses itself. The capacity and aspiration of some outstanding individuals has commonly to be tempered by the needs of others in society, the extraordinary sensitivity of others is frustrated by the lack of comparable sensitivity in their surrounds. This leads to a paradox, which is at the heart of Man's misery—evolution that has provided the single most variable, and therefore most successful, organ in the history of our planet has, through that, led to a society which, to protect itself, must proscribe that very variability and force all, to variable extents, to adopt a common behaviour, one to which not all are suited by their genetic inheritance to accept.

It is a remarkable fact that the variation in behaviour that society does not tolerate is often expressed to acclamation in art, and the enduring success of this art is itself telling. There are numerous examples of this, and it will suffice here to refer to three relatively well-known ones. Racine's Phèdre, Mozart's (and Lorenzo da Ponte's) Don Giovanni and Balzac's Vautrin are all characters whose behaviour varies significantly from the norm; they are characters whom many would find repulsive if encountered in real life. Yet the root cause and the source of their characteristic deviation from what society, in its self-protective role, has found acceptable is to be found in variability, which also happens to be the rich and fertile ground for the evolution of brains and their potentials. Phèdre falls passionately and helplessly in love with her stepson, a doomed relationship of which she is forever aware, but cannot follow the better course dictated by reason because of her mental constitution. In *Don Giovanni*, Mozart (and Da Ponte) created a character who many would today call a serial rapist or lecher. Yet he, too, is prey to his biological constitution; that compulsive behaviour makes it difficult for him to accept the special relationship in which women want to hold him. He remains throughout indifferent to his successes and, significantly, it is not he but his servant Leporello who keeps the list of these successes in the captivating aria, *Madamina*. In the end, he accepts his biological destiny with courage and even dignity. Balzac, in novel after novel, portrayed characters that pursue their uncontrollable passions, be they of greed, love, hate or ambition, to their inevitable self-destruction. His most prodigious creation, and one of the most extraordinary creations of romantic literature, is Vautrin, a cynical criminal and murderer, with an inflexible will to dominate the society which he despises and whose moral weakness he has identified. But he is also of romantic disposition, attached to those he loves and wanting to dominate society through them, to 'love his creature, shape it, manipulate it for his use' (*Illusions perdues*), to become the man 'who marries events and circumstances in order to lead them' (*Père Goriot*).

There are many other examples of what we may call pathological behaviour that one could give from art, but it is significant that artists have chosen to portray these characters and even more significant that we as a society attach such importance to these creations, judging by the enduring success of these works. It is art, in a sense, that has presented the acceptable face of human variability, has made that variability safe in its pages and scores, while at the same time allowing us all to empathize with the destroyed characters, because the seeds of their destruction are in us too, to variable and differing extents. It is indeed a huge tribute to the capacities of the brain to abstract and generalize that the story of Phèdre, for example, can be easily applied to so many other contexts which are quite remote from the story as developed by Racine. At the turn of the Millennium, science has had little success in tracing the neurological substratum of that variability and has a very imperfect answer as to whether differences in mental constitution are to be looked for in differences in connectivity, in biochemical or pharmacological properties or in some other factor yet to be identified. If found, the differences in brain architecture that lead to such variants will not alleviate the unhappiness of individuals. But it will create a huge problem for society: this has already happened to a certain extent, where criminals and others whose behaviour is judged undesirable, have pleaded in court, with varying degrees of success, that their constitution left them no alternative but to behave in the manner that they have.

4. CONCLUSION

I have tried in this essay to show the direction that neurobiology will take in the coming century, towards problems that have been at the heart of philosophical thinking for millennia. I have tried to show that the problems will be more manageable neurobiologically if dissected into components, without losing sight of the grander questions. I have tried to argue that the variability that is the breeding ground for evolutionary development of the brain and its success is also a cause of its misery; that the evolutionary development of an efficient knowledge-acquiring system depends not only on knowledge of particulars but also on its capacity to abstract and to form ideals, which carry with them the seeds of misery and disappointment. Whether a better understanding of the capacities of the brain to abstract and formulate ideals, or of the determinants of its variability expressed in these capacities, will bring more happiness to Man is not at all clear. I myself rather doubt it. But when, and if, neurobiology faces these challenges, it must also address a wider question with important implications. Supposing that we are ever successful in eradicating that variability, which is at the basis of our developing brains, and supposing that through this we were to have brains that formed identical ideals, would we not also be eradicating the sources of the richness of our culture, as expressed in the achievements of Man? If this hypothetical future were to make men happier, would it be worth pursuing, knowing the price that will be extracted in return?

I do not claim any originality for the thoughts expressed here. Indeed, given my profound ignorance, of which I become more aware with every new reading, I would not be surprised if many of these thoughts have been expressed by others. But I hope that I have also managed to celebrate the glories and lament the tragedies of the human race, both of which have common sources.

END NOTES

[1] That the rewarding richness of our brains should carry with it the makings of human misery is implied in the title of this article, which I have adapted from the title of the novel by Honoré de Balzac, *Splendeurs et misères des courtisanes*.

[2] I use the German title here because the word *Vorstellung* is not easily rendered into English, the closest being that of 'placing before'. An early translation of the work into English by B.

Haldane and J. Kemp gave the title as *The world as will and idea*, whereas the translation that I have used, that of E. F. J. Payne, renders it into the *The world as will and representation*, even though Payne tells us that Schopenhauer himself translated *Vorstellung* as 'idea' in his criticism of Kant's philosophy. Schopenhauer used the term *Vorstellung* to describe what he called 'an exceedingly complicated physiological process in the brain of an animal, the result of which is the consciousness of a *picture* there'. In his translator's note to volume 1 of Schopenhauer's work, Payne tells us of the translation by Haldane and Kemp that, 'the interests of truth and the importance of this work in the history of philosophy require that attention be drawn to the many errors and omissions in their translation, over a thousand of which came to light when it was compared with the German text, and which seriously detract from its merit as a work of scholarship'.

REFERENCES

Bartels, A. & Zeki, S. 1999 The architecture of the human colour centre. *Eur. J. Neurosci.* (In the press.)

Constable, J. 1771 *Discourses on art* (ed. R. Wark, 1975), no. 4.

Critchley, M. 1965 Acquired anomalies of colour perception of central origin. Brain **88**, 711–724.

Darwin, C. 1859 *The origin of species*. London: J. M. Dent (Reprinted in 1928).

Einstein, A. 1916*a* Physics and reality. *J. Franklin Inst.* **221**, 313–347.

Einstein, A. 1916*b* Ernst Mach. *Phys. Z.* **17**, 101–104. Quoted by Miller, A. I. 1984 *Imagery in scientific thought: creating 20th-century physics*. Boston, MA: Birkhäuser.

Freud, S. 1930 *Das Unbehagen in der Kultur*. Translated by J. Riviere (1957) as *Civilization and its discontents*. London: Hogarth.

Gleizes, A. & Metzinger, J. 1913 *Cubism*. London: Fisher Unwin.

Harlow, H. F. & Harlow, M. K. 1962 Social deprivation in monkeys. *Sci. Am.* **207**, 136–146.

Hegel, G. W. F. 1832–1845 *Aesthetics*, vol. 1. Oxford, UK: Clarendon Press. (Translated by T. M. Knox.)

Hubel, D. H. & Wiesel, T. N. 1970 The period of susceptibility to the physiological effects of unilateral eye closure in kittens. *J. Physiol.* (*Lond.*) **206**, 419–436.

Hubel, D. H. & Wiesel, T. N. 1977 The Ferrier Lecture: functional architecture of macaque monkey visual cortex. *Proc. R. Soc. Lond.* B **198**, 1–59.

Kant, I. 1781 *Kritik der reinen Vernunft*, 1st edn. Translated by W. S. Pluhar (1996) as *Critique of pure reason*. Indianapolis, IN: Hackett.

Kant, I. 1783 *Prolegomena*. Translated by P. G. Lucas (1953). Manchester University Press.

Kant, I. 1787 *Kritik der reinen Vernunft*, 2nd edn. Translated by W. S. Pluhar (1996) as *Critique of pure reason*. Indianapolis, IN: Hackett.

Land, E. H. 1974 The retinex theory of colour vision. *Proc. R. Instn Gt Br.* **47**, 23–58.

Leibniz, G. W. 1714 *Monadologie*. Translated by R. Latta (1898) as *Monadology*. Oxford, UK: Clarendon Press.

Lewandowsky, M. 1908 Ueber Abspaltung des Farbensinnes. *Monatsschrift für Psychiatrie und Neurologie* **23**, 488–510.

Logothetis, N. K., Pauls, J., Bulthoff, H. H. & Poggio, T. 1994 View-dependent object recognition by monkeys. *Curr. Biol.* **4**, 401–414.

Logothetis, N. K., Guggenberger, H., Peled, S. & Pauls, J. 1999 Functional imaging of the monkey brain. *Nature Neurosci.* **2**, 555–562.

Matisse, H. 1978 *Notes d'un peintre, La Grande Revue, LII.* Quoted in J. D. Flam (1978), *Matisse on Art.* Oxford, UK: Phaidon.

Maxwell, J. C. 1872 On colour vision. *Proc. R. Instn Gt Br.* **6**, 260–271.

Miller, A. I. 1984 *Imagery in scientific thought: creating 20th-century physics*. Boston, MA: Birkhäuser.

Mollon, J. D., Newcombe, F., Polden, P. G. & Ratcliff, G. 1980 On the presence of three cone mechanisms in a case of total achromatopsia. In *Colour vision deficiencies*, vol. 5 (ed. G. Verriest), pp. 130–135. Bristol, UK: Adam Hilger.

Mondrian, P. 1937 Toward the true vision of reality. In *The new art—the new life, the collected writings of Piet Mondrian* (ed. and trans. H. Holtzman & M. S. James). Boston, MA: G. K. Hall (1986).

Moutoussis, K. & Zeki, S. 1997 A direct demonstration of perceptual asynchrony in vision. *Proc. R. Soc. Lond.* B **264**, 393–399.

Newton, I. 1704 *Opticks: or, a treatise of the reflexions, refractions, inflexions and colours of light*. New York: S. Smith and B. Walford.

Penrose, R. 1994 *Shadows of the mind*. Oxford University Press.

Picasso, P. 1935 In an interview with Christian Zervos, in *Cahiers d'art* **X**, 173–178. Quoted in H. B. Chipp 1968 *Theories of modern art*. Berkeley, CA: University of California Press.

Poincaré, H. 1894 Sur la nature du raisonnement mathèmatique. *Rev. Mèt. Mor.* **2**, 371–384. Quoted by Miller, A. I. 1984 *Imagery in scientific thought: creating 20th-century physics*. Boston, MA: Birkhäuser.

Poincaré, H. 1898 On the foundations of geometry. Translated by T. J. McCormack *Monist* **9**, 1–43. Quoted by Miller, A. I. 1984 *Imagery in scientific thought: creating 20th-century physics*. Boston, MA: Birkhäuser.

Rakic, P. 1977 Prenatal development of the visual system in the rhesus monkey. *Phil. Trans. R. Soc. Lond.* B **278**, 245–260.

Russell, B. 1914 *Our knowledge of the external world*. London: Open Court.

Schopenhauer, A. 1854 *Uber das Sehn und die Farben: Ein Abhandlung*. Translated by E. F. J. Payne (1994) as *On vision and colors: an essay*, p. 84. Oxford, UK: Berg.

Schopenhauer, A. 1859*a* *Die Welt als Wille und Vorstellung*, vol. 1, 3rd edn. Translated by E. F. J. Payne (1969) as *The world as will and representation*. New York: Dover Publications.

Schopenhauer, A. 1859*b* *Die Welt als Wille und Vorstellung*, vol. 2, 3rd edn. Translated by E. F. J. Payne (1969) as *The world as will and representation*. New York: Dover Publications.

Senden, M. von 1932 *Raum- und Gestaltauffassung bei Operierten Blindgebornen*. Translated by P. Heath (1960) as *Space and light*. London: Methuen.

Von Helmholtz, H. 1911 *Handbuch der Physiologischen Optik*. Hamburg, Germany: Voss.

Wittgenstein, L. 1977 Bemerkungen uber die Farben. Translated as *Remarks on colour* (ed. G. E. M. Anscombe). Oxford, UK: Basil Blackwell.

Zeki, S. 1990 A century of cerebral achromatopsia. *Brain* **113**, 1721–1777.

Zeki, S. 1993 *A vision of the brain*. Oxford, UK: Blackwell Science.

Zeki, S. 1999 *Inner vision: an exploration of art and the brain*. Oxford University Press.

Zeki, S. & Marini, L. 1998 Three critical stages of colour processing in the human brain. *Brain* **121**, 1669–1685.

Zeki, S., Watson, J. D. G., Lueck, C. J., Friston, K. J., Kennard, C. & Frackowiak, R. S. J. 1991 A direct demonstration of functional specialization in human visual cortex. *J. Neurosci.* **11**, 641–649.