Autism, Intelligence and Consciousness

By RODNEY M. J. COTTERILL

Biologiske Skrifter **45**

Det Kongelige Danske Videnskabernes Selskab
The Royal Danish Academy of Sciences and Letters

Kommissionær: Munksgaard · Copenhagen 1994

The Royal Danish Academy of Sciences and Letters

publishes four monograph series, an Annual Report and, occasionally, special publications. The format is governed by the requirements of the illustrations, which should comply with the following measures.

	Authorized Abbreviations
Historisk-filosofiske Meddelelser, 8°	Hist.Fil.Medd.Dan.Vid.Selsk. (printed area 175 x 104 mm, 2700 units)
Historisk-filosofiske Skrifter, 4° (History, Philosophy, Philology, Archaeology, Art History)	Hist.Filos.Skr.Dan.Vid.Selsk. (printed area 2 columns, each 199 x 77 mm, 2100 units)
Matematisk-fysiske Meddelelser, 8° (Mathematics, Physics, Chemistry, Astronomy, Geology)	Mat.Fys.Medd.Dan.Vid.Selsk. (printed area 180 x 126 mm, 3360 units)
Biologiske Skrifter, 4° (Botany, Zoology, Palaeontology, General Biology)	Biol.Skr.Dan.Vid.Selsk. (printed area 2 columns, each 199 x 77 mm, 2100 units)
Oversigt, Annual Report, 8°	Overs.Dan.Vid.Selsk.

The Academy invites original papers that contribute significantly to research carried on in Denmark. Foreign contributions are accepted from temporary residents in Denmark, participants in a joint project involving Danish researchers, or partakers in discussion with Danish contributors.

Instructions to Authors

Manuscripts from contributors who are not members of the Academy will be refereed by two members of the Academy. Authors of accepted papers receive galley proof and page proof which should be returned promptly to the Editor. Minidiscs etc. may be accepted; contact the Editor in advance, giving technical specifications.

Alterations causing more than 15% proof changes will be charged to the author(s). 50 free copies are supplied. Authors are urged to provide addresses for up to 20 journals which may receive review copies.

Manuscripts not returned during the production of the book are not returned after printing. Original photos and art work are returned when requested.

Manuscript

General. – Manuscripts and illustrations must comply with the details given above. The original ms. and illustrations plus one clear copy of both should be sent to the undersigned Editor.

NB: A ms. should not contain less than 32 *printed* pages. This applies also to the Mat.Fys.Medd., where contributions to the history of science are welcome.

Language. – English is the preferred language. Danish, German and French mss. are accepted and in special cases other languages. Where necessary, language revision must be carried out before final acceptance.

Autism, Intelligence and Consciousness

By RODNEY M. J. COTTERILL

Biologiske Skrifter **45**

Det Kongelige Danske Videnskabernes Selskab
The Royal Danish Academy of Sciences and Letters

Kommissionær: Munksgaard · Copenhagen 1994

Abstract

A new theory of autism is presented, and its connection with the concepts of intelligence and consciousness are explored. The theory aims at explaining those aspects of autism which are not encompassed in the currently-popular view that the syndrome is related to the autistic person's lack of a *theory of mind*. These other characteristics include ritualistic and bizarre movements, an insistence on sameness, and impaired motility. The new theory also attaches considerable significance to the fact that about one third of all autistic patients also suffer from epilepsy. This leads to the suggestion that the underlying cause of autism is what experimental psychologists refer to as *breakthrough*, the inappropriate occurrence of this in the autistic patient being caused by a failure of inhibition at the neuronal level. It is argued that such breakthrough will influence the patient's ability to process sensory information, and that it will thus affect certain aspects of the patient's intelligence. It is also argued that a proper analysis of intelligence requires an understanding of consciousness, and tentative theories of both those attributes are put forward. Finally, an attempt is made to link certain aspects of intelligence to possession of a *theory of mind*, and thus explain all observed features of autism in a unified manner.

RODNEY M. J. COTTERILL,
Division of Biophysics,
Physics Department,
Danish Technical University,
DK-2800 Lyngby,
Denmark

© Det Kongelige Danske Videnskabernes Selskab 1994
Printed in Denmark by Special-Trykkeriet Viborg a-s
ISSN 0366-3612 ISBN 87-7304-254-4

Contents

Introduction

It is often the case in scientific investigation that understanding of the general is furthered by contemplation of the particular. When a phenomenon is the product of several contributing agencies, the complex interplay between the latter may be illuminated by observing what happens when one of them malfunctions; as the saying goes, *treasure your exceptions*. The history of brain science includes many examples of advances that came through considering the consequences of accidental or deliberate lesions to one or another part of the brain. This approach is especially useful when interest is focussed on a well-defined attribute closely related to sensory input, such as the visual perception of form, colour or motion. When the faculty is more general, it seems reasonable to assume that a greater number of brain regions will be participating. This would presumably make it more difficult to identify the neural circuits involved. We are beginning to get an idea of which cortical regions are implicated in the various aspects of vision, for example, but where is the seat of personality? What, for that matter, are the real determinants of intelligence and consciousness? Amongst other things in this paper, I am going to be making an attempt to answer those latter questions, and I shall be drawing on observations of a particularly tragic exception to the norm, namely *autism* (Kanner 1943, Asperger 1944, Rimland 1964, Park 1967, Wing 1971, Bartak & Rutter 1976, Rutter 1978, Ornitz 1983, Schopler & Mesibov 1987, Frith 1989, Gillberg 1989, Treffert 1989, Naruse & Ornitz 1992, Gillberg & Coleman 1992).

That affliction will not merely serve to provide data for the discussion of intelligence and consciousness, however. Far from it, because the main object of what follows is to essay a theory of autism itself. The article starts with a fairly concise survey of the various symptoms of autism, the great majority of these being displayed by all autistic people. I then review a most interesting idea that has emerged recently which suggests that autistic people lack what is called a *theory of mind*, which is to say that they are not aware that other people have thoughts. In order to put my own ideas in context, there then follow fairly extensive reviews of recent work on working memory, on attention, with an emphasis on the phenomenon of breakthrough, and on consciousness. There is then a discussion of embedding, which has a direct bearing on the concept of a theory of mind, and which I argue may be a good indication of intelligence itself. Finally, I introduce my own theory of autism, which perceives *breakthrough* as the key to the infirmity, and I link this to a biochemical imbalance in the distribution of excitatory and inhibitory neurotransmitters. The article closes with two predictions, one at the genetic level and the other regarding a possibly observable effect with magnetoencephalography.

Characteristics of Autism

General Remarks

Autism affects about 4 out of every 10,000 people, apparently from birth, the incidence being markedly higher amongst males. (The actual proportion appears to be very roughly 3 males to each female.) Its main characteristic is severely impaired social interaction, and this has led some to liken the typical autistic's existence to living in a invisible shell. The syndrome was first properly described by Kanner (1943), and independently by Asperger (1944). The term comes from the Greek word *autos*, meaning self, a reference to the withdrawal that is the affliction's hallmark. It was coined in 1911 by Bleuler, who is better known for his pioneering identification of schizophrenia, with which autism was initially coupled. Kanner pinpointed two aspects of autism which he took to be its defining features, namely social isolation and resistance to novelty; the autistic is profoundly aloof to the human environment and obsessively desires constancy. But Kanner also noted the frequent presence of isolated abilities, particularly those related to memory. Apserger's list of characteristics included such additional items as a paucity of facial expressions and gestures, lack of eye contact, and meagerness or absence of language.

As we will see shortly, these inventories of deficits have been augmented over the intervening years, and there are now roughly twenty autistic characteristics which respect neither race nor colour. A large fraction of them can be rationalized into a group of three, as a result of a extensive survey by Wing and Gould (1978, 1979). Carried out in the London borough of Camberwell, it led to the highlighting of a *triad of impairments* associated with Autism (Wing 1981). (The Camberwell study screened 35,000 children, up to 14 years of age, who lived in the borough on 31 December 1970. Of the 914 known to be physically or mentally handicapped, 7 had the symptoms described by Kanner, an incidence of 2 in 10,000.) The three traits are: (a) marked truncation or total absence of verbal and non-verbal communication; (b) severe impairment of (reciprocal) social interaction, particularly with peers; and (c) absence of imaginative pursuits, including pretending, coupled with the presence of stereotyped behaviour. We will return to this well-established triad later, in connection with an intriguing study carried out by the techniques of cognitive science. Before moving on to that, however, we will do well to consider the various items of the larger list. This will give us a better feel for what autism really entails.

The First Group of Symptoms

For reasons that will become apparent later, I am going to divide the various autistic characteristics in two broad classes. Included in the first of these are emotional detachment, aversion to affection, aversion to physical contact, lack of eye contact, lack of personal identity, lack of facial expression, disturbed communication and lack of imaginative play. Let us consider them in that order. The most striking thing about autistics is their isolation. It is almost as if they are inhabiting another world, from which normal people are excluded. And if anything, this impression is strengthened by the fact that their appearance is otherwise normal. This detachment is particularly distressing to the parents of the autistic child; they feel unwanted, and starved of the feedback that plays such an important part in the rearing process. Coupled to this is the frustration of not know-

ing whether the child is happy under a given set of circumstances. Even worse is the parents' feeling of being shunned, because the autistic child actively opposes affection. There is the impression that fond attention is either not being understood or that it is unwelcome. And if the frustrated parent attempts to embrace the child, as a consequence, the physical contact is vigorously resisted.

Another disconcerting facet of the syndrome is the lack of eye contact. Normally, the establishing of this bond between mother and child presages a succession of developmental stages. In the case of the autistic, the development is stopped in its tracks, and the child seems to pass into a state of suspended emotional growth. It is as if life has been put on hold. The invisible shell is firmly in place, and the child appears to be perfectly happy with this state of affairs. The impression of suspension will be reinforced, over the years, by another peculiarity of the typical autistic, namely the almost total absence of wrinkles; the autistic face is remarkable for its lack of weathering. This is attributable to the scarcity of facial expressions. The history of a normal person's life is etched in the lines of the face. Every frown and smile leaves its indelible mark. The autistic face is singularly free of such markings; it is like an unwritten page.

A contributing factor to the paucity of facial expression may well be the absence of a sense of identity. Autistic speech, when it is present at all, is devoid of the word *I*. And the mystery is often compounded by the fact that some autistics nevertheless do refer to themselves, by using their first names (Park 1967). Moreover, the use of other people's names appears to present them with no difficulty whatsoever. Indeed, those who possess (a rudimentary form of) language will often go out of their way to enquire as to a person's name, apparently for ease of subsequent reference.

A major aspect of autism is disturbed communication. Although the typical autistic has difficulties with language, as will be discussed later, this is not the only factor. There is, after all, more to communication than the spoken (or written) word. Normal people also communicate with hand gestures, with facial expressions, with all manner of non-verbal vocal sounds, as well as with such subtler things as the raised eyebrow and certain tones of voice. The autistic seems not to appreciate such modes of signalling. The importance of this particular class of deficit in the diagnosis of the affliction could hardly be exaggerated, for it suggests that the invisible shell may be imposed not by choice but by an essential limitation in the ability to handle sensory input. It must be stressed that the faculties underlying both these factors, volition and cognition, have neuronal origins. There are parts of the brain that control drive just as there are parts that mediate the reception of information from the environment. And the two systems are not unrelated, of course.

To round off what I have chosen to call the first class of symptoms, there is the absence of imaginative play. The autistic child is indifferent to dolls, puppets and toy animals. These are treated in just the same way as other inanimate objects, no inclination being shown to attribute living qualities to them. This too is deeply disappointing to the parent of the autistic child. What would normally be regarded as traditional nursery accoutrements are simply ignored, and the feeling of frustration is compounded when it is observed that preference is given to seemingly uninspiring objects like odd bits of wool or string, dog-eared pieces of paper, and even the grain in a wooden floorboard. This autistic trait is in particularly sharp contrast to what is observed in the normal child. Even while speech is still developing, the normal child is according dolls a quasi-human status, and constructing plots for them that can be impressively involved.

The Second Group of Symptoms

In the next section, I am going to be describing a clever investigation which has produced an idea that can explain every one of the symptoms belonging to the first class. On the other hand, it appears to have no bearing on the symptoms of what I am calling the second class. Let us now move on those other traits, therefore, and consider the bare facts. The class comprises that remarkable resistance to change, fixation on certain objects, bizarre ritualistic behaviours, abnormal anxiety patterns, impaired motility, disturbed perception and attention, language deficits, lack of initiative, and retarded intellectual development.

Autistic people are the arch conservatives of this world. Their tolerance of change is virtually non-existent. And their characteristically good memories exacerbate the problem that this can represent for all those who must interact with them. Many a parent of an autistic child has been driven to distraction by the insistence on sameness, both with respect to the physical aspects of the home and to the daily routine. Returning after an absence, the autistic child will invariably breeze through the house, quickly pinpoint any changes that have occurred, and then start demonstrating for a return to the *status quo*. This is the case even for things which would normally be regarded as insignificant, such as the exact placing of the various pieces of furniture. There is an accompanying tendency to prefer simple geometric order. I once had the opportunity of observing the behaviour of an autistic boy who happened to possess a dozen or so identical rubber penguins. These were always lined up on a particular shelf, the spacings and the orientations being standardized to within millimeters and degrees. I succumbed to the temptation of slightly displacing one of them, while he was briefly out of the room. Upon his return, he promptly restored the symmetry, uttering little squeaks of frustration until the job was done. The insistence on sameness of routine can place quite a burden on the autistic child's family. Any departure from the perceived norm can lead to pandemonium.

We have already noted the preference an autistic person can display for trivial objects. In many cases the preference is transformed into a full-blown fixation. Seemingly inconsequential items can be studied for hours on end, as if they held the secret of some deep wisdom or mystery. And any attempt to terminate such a seance, on the part of parent or guardian, has to contend with that overwhelming desire for sameness. Not infrequently, the fixation acquires a ritualistic dimension, elaborate and repetitive behaviours being built up around the object of attention. A common manifestation of this autistic trait is the interminable twirling of a piece of string or rocking of a curved object. In some cases, the ritual involves no object, and it is merely a part of the body that is rocked, usually the head.

Abnormal anxieties are another feature of the typical autistic person's existence. Things which clearly pose no danger can become the focus of extreme and illogical fear. There are autistics, for example, who dread passing through certain doors, even though there has been no history of anything adverse connected with them. Conversely, things or situations which clearly *are* dangerous often seem to pose no threat to the autistic individual. This is obviously cause for considerable concern, because it makes such a person vulnerable to accident. It is indeed the regrettable fact that autistics are prone to serious injury because of this. I believe that both these features of the autistic existence might be connected to something we will be coming to shortly, namely impaired perception.

The word motile merely means capable of movement. The impaired motility that is seen in many autistic people refers to their awkward gait, particularly when they have to negotiate

uneven surfaces. This is particularly noticeable during walks in wooded terrain, when roots and fallen branches are a common hazard. Either of these types of obstacle can present the autistic person with a surprisingly difficult challenge. Conversely, an uncommonly flat surface, such as a dance floor, can elicit an almost rapturous response, and the autistic will often spontaneously launch into a series of runs and hops. Impaired motility sometimes affects other movements, and it usually the case that autistic reactions are slower than normal. This does not seem to apply to withdrawal from painful stimuli, however. In general, the level of coordination is lower than in normal people. It is worth adding that both crawling and walking tend to be achieved at a later stage than in the case of normal children.

Disturbed perception is a common feature of this affliction, and I have already suggested that it could underlie the abnormal anxiety patterns. It primarily takes the form of diminished awareness, but there is also a strong suggestion of warpage, with prominent details of objects and situations ignored at the expense of trivia. An associated deficit concerns the typical autistic person's difficulties with attention. It is not unusual for these people to be initially diagnosed as being deaf, for instance, because of their failure to react to sounds, even though these might be quite loud. It can take an expert to establish that deafness is not the cause of the trouble. We have already noted the tendency to fixate on certain objects or rituals, and this can be looked upon as an impaired ability to shift attention from one thing to another. A similar rationalization may be applied to the obsessions that are another autistic characteristic. Recent research has tended to attach considerable importance to these deficiencies of attention, and I will be describing anatomical and physiological evidence which supports such a view.

We have already noted that there is more to communication than language. The diminished facility with language that is a prominent feature of autism nevertheless deserves consideration, not the least because of the increased understanding of the neuronal basis for language which is currently emerging. Unfortunately, they are at present mostly limited to nouns and verbs, which have been identified with specific cortical regions. The cortical areas serving adjectives, prepositions and other types of word have not yet been pinpointed. It seems that their production in the cortex is a more nebulously affair. When an autistic person possesses some language, it is usually nouns and verbs that are present, indeed, whereas there is scant evidence of adjectives, prepositions, adverbs or pronouns.

Another thing that strikes anyone who has studied autistic people is their lack of initiative. Left to their own devices, they seem perfectly content to continue whatever they are doing, often for hours on end. The continuity is broken only by the calls of basic bodily needs. A concomitant of this feature of autism appears to be a lack of curiosity, but my own feeling about this is that it might not be so pronounced as some have taken it to be. My reason for saying this is that autistics can nevertheless be found to have considerable stores of passive knowledge. It is not that they are furtive about their acquisition of facts. Rather, it is the lack of communication which allows the learning process to pass unnoticed.

Given the woefully disadvantaged state with respect to language, as well as the reticence in taking the initiative, it is hardly surprising that the average autistic person is intellectually backward. The autodidactic processes that is so important an aspect in normal learning obviously labours under difficult conditions. Moreover, it takes great patience on the part of a parent to continue attempts at communication when feedback is sparse or totally absent. Another contributing cause, which I will now in-

troduce into this discussion of autism, appears to be mental retardation. Let me say immediately that this side of the issue is not clear-cut. There are autistic people who have normal IQs, and there are people with low IQs who are not autistic. As we are going to see later, however, there is a clear trend which makes autism more common as the IQ decreases.

I will round out this survey of autistic characteristics with just two more aspects of the story, one that has already been mentioned, namely islands of ability, and one that does not usually appear in lists of the type that we have been contemplating. Let us first consider the remarkable feats of memory which many autistic people manage. A good case in point concerns music, since many autistic individuals show considerable aptitude for it. I have had numerous opportunities to observe this in autistic people and have often been deeply impressed. I once played for an autistic girl the record of a piece of music which she had almost certainly never heard. The piece lasted about four minutes, and when it was done I managed to induce her to hum it back to me. She did this without making a single error, as far as I could judge. It was as if she possessed a sort of internal tape recorder, which she played back upon demand. I must nevertheless draw a distinction between such acts of recall and the sort of thing achieved by what is referred to as an *idiot savant*. *Idiot savants* are much less common than autistics, and their accomplishments are both narrower and more extraordinary than the latter. There of people belonging to this class who can multiply fifteen digit numbers in their heads, while others are able to say on which day of the week a particular date fell in a remote year. This sort of thing is quite beyond the typical autistic person. Moreover, remarkable computational achievements of this type are just as frequently associated with people who are not autistic.

I will close with something that is so much more specialized than anything else we have been considering that it would seem to fit only awkwardly into this section. It is nevertheless of great interest, because it probably reveals something about the way in which the autistic mind works. Indeed, it would seem to lay bare one of the limitations of autistic thought. It concerns the way in which the typical autistic person tackles a jigsaw puzzle. This is again something which I have had numerous opportunities of studying, and the pattern is remarkably uniform amongst autistic people. When a normal child is confronted with a new puzzle, he or she invariably starts by examining the picture on the cover, whereafter a reasonably orderly start is made on assembling the pieces. A favourite strategy is to begin by finding all the edge pieces, since these will collectively constitute the frame of the completed puzzle. Another useful tactic is to select all similarly coloured pieces, and thereby build up the final picture's major areas.

The autistic approach is quite different. There is no consulting of the picture on the cover. Instead, a single piece is selected, apparently at random, and then all the other pieces are closely inspected to see if they match with the first. This is done with such accuracy that the proper fit is virtually inevitable, once the second piece has been found. The procedure is then repeated, with all the remaining pieces being inspected until a third piece has been found which will fit onto *either* of those already in place. The picture thereby grows around an initial nucleus, the autistic person appearing to have an internal photograph of the perimeter, at each and every stage. There is an almost humorous consequence of this radically different process, and again this is something that I have observed many times. The autistic puzzle builder will occasionally stop and request the recovery of a piece that has dropped on the floor. This invariably comes as a surprise, especially when the puzzle is still far from comple-

tion. But a search always reveals that a piece has indeed fallen off the table, and it is immediately placed in its correct position.

My own interpretation of this unusual approach is that it exposes a limitation in the way the autistic mind handles information. I believe that the indifference shown toward the picture on the cover stems from an inability to work with large areas of the field of view. It shows, I feel, that autistics experience difficulty in obtaining a sufficiently broad view of things. The picture on the cover of a jigsaw puzzle is only an example, of course, and one of only minor consequence. More significantly, if this idea is correct, the limitation may often cause things to be missed that would otherwise have been grasped. This lack of overview may well underlie the disturbed perception and attention that we considered earlier, and I believe that it could even have a bearing on what I referred to as impairments belonging to the first of the two classes.

This completes our initial look at the common symptoms of autism. Let us now move on to those tests which appear to explain the items in that first class of symptoms.

Theory of Mind

The interesting idea has recently emerged that autistic people are unaware that other people think. Whether or not this has to imply that they are also unaware of their own thoughts is a more difficult question, to which we will have to return later. The term employed to describe the situation is *theory of mind* (Leslie 1987), and the idea is that the mental equipment of the typical autistic does not include this. Such an inference is properly viewed as a theory, because the state in question is not directly observable, and its existence can be used to make predictions about the behaviour of others. The term is now widely adopted, not only for the affliction being discussed here but also in connection with the mental powers of other animals (Premak & Woodruff 1978). It appears first to have been linked specifically to autism in a suggestion by Barlow (1980). Let us begin by considering the resourceful manner in which absence of a theory of mind was established in a large fraction of autistic children.

I am not sufficiently conversant with the techniques of cognitive science to know whether the experiments I am about to describe are typical examples of work in that field, but they are certainly impressive, given the difficulties of interacting with handicapped individuals. The ingenious Sally-Anne test is sometimes attributed to Wimmer and Perner (1983), but its substance was anticipated in a discussion by Bennett (1978), Dennett (1978) and Harman (1978), all of whom credited Lewis (1969) with the underlying idea. The test proceeds in the following manner. A little scenario involving two dolls, Sally and Anne, is acted out in front of the subject being tested. Sally has a basket, and initially also an object she is fond of, such as a marble, while Anne only has a box. Sally puts her marble into her basket, covers it with a cloth, and retires from the stage, leaving the basket behind. During Sally's absence, Anne transfers the marble from the basket to her box, closes the lid, and then she too departs.

Sally returns to the scene of the crime, and the question arises as to where she will look for her marble: in the basket or in the box. A normal person, not having been present when the marble was transferred, will obviously look in the basket. Baron-Cohen, Leslie and Frith have applied this test to a number of children, all with a mental age above three, some being autistic, some normal, and some mentally retarded (Down's syndrome) but not autistic (Baron-Cohen, Leslie & Frith 1985). The majority of the children in the latter two groups, when questioned, arrived at the correct answer: Sally would look in the basket because this is where she believed the marble to be. The majority of the autistic children, on the other hand, opted for the box. This was where the marble was located, and this is where they felt that Sally would look. They were clearly unable to attribute belief to her, and the logical conclusion was that they thus lacked a theory of mind. This experiment was as simple as it was elegant.

A question naturally arises as to the validity of this test. Given the absence of imaginative play in autistic children, as noted earlier, the choice of dolls for this little play was not particularly fortunate. Leslie and Frith therefore repeated the procedure with a different group of autistic children (who were actually more able than the ones tested earlier), replacing Sally and Anne by themselves, and using a coin rather than a marble (Leslie & Frith, 1988). The outcome was roughly as before, with 15 out of 21 going for the inappropriate option.

To put this result in perspective, *random choice* would have given 10.5 failures (i.e. 50%). The *standard deviation*, which is used in statistical tests of significance, is a measure of the amount of scatter of a series of numbers or measurements about their mean value. It is defined as the square root of the average value of the squares of the deviations from their mean value. If the deviation from the random-choice score is within the standard deviation, the result has dubious significance. The standard deviation for 21 measurements, all having equally probable outcomes, would be the square root of 21x0.5x0.5, namely 2.3, so 15 out of 21 *is* significant.

Leslie and Frith also looked for the appreciation of beliefs in others by subjecting the children to a test in which a (tubular) Smarties box contained a pencil stump rather than the usual sweets. The children were tested individually and in sequence. Each, after discovering the inedible content, was asked what the next person in line would say was in the box. Only 4 out of 20 correctly realized that the answer would be *Smarties*.

There are points to be made about the methodology of these experiments. One concerns their involvement of verbal reports by the autistic subjects. As we saw earlier, autistic people tend to have language deficits, and their performance with prepositions is indifferent at best. Although this would no doubt have made the trials more difficult to carry out, it would have been preferable to work with a protocol that remotely checked for surprise, without making any demands regarding language. This is now routinely achieved by monitoring the size of the pupil; when an individual is surprised, the pupil dilates, and this is of course a totally involuntary process. A more subtle question concerns that test with the Smarties box. We should note that the testing paradigm was in that case changed from the *where* of the Sally-Anne problem to the *what* of the Smarties box contents problem, and these things are handled by different parts of the cerebral cortex. We could turn this to advantage, however, and say that because there was nevertheless roughly the same failure rate, the neuronal cause of the autistic deficit, whatever it is, must affect reasonably widespread areas of the cortex.

Lack of a theory of mind does appear to account for the symptoms that I referred to

above as falling in the first class. The triad of impairments listed earlier (Wing 1981) are all in this class, if one takes lack of verbal communication *not* to embrace language deficits. Frith calls this triad the core symptoms of autism, while she relegates the items in the second class to peripheral roles (Frith, 1989). This may prove to be counter-productive, in the long run. Lack of a theory of mind does not obviously account for resistance to change, and it is difficult to see what relevance it has to the questions of fixation and ritualistic behaviour. The same could be said of the language deficits that are frequently seen in autism. It is also hard to see how not having a theory of mind could lead to delayed crawling and walking in autistics; the theory would seem quite incapable of explaining impaired motility. Thus although one must admire the ingenuity that has gone into what is surely a significant piece of progress, it would be unfortunate if this milepost came to be regarded as the end station. One should try to look even deeper into the problem. In order to do this, I am going to have to review a considerable amount of recent work on memory, attention and consciousness.

Working Memory

The key step in mental arithmetic is retention of the result of previous operations while a new one is being carried out. When adding a string of numbers, for example, one has to remember the current sum while adding to it the next number in the sequence. The key step in the use of verbal language is retention of what has just been spoken while suitable new words can be selected and arranged for one's continuation. These are just two examples of the effortless procedure that is going on during much of our waking time: mental retention of items while new ones are either received through our senses or retrieved from memory. The mechanism that enables us to carry out this obviously important process is referred to as *working memory*. It also involves erasure, in that items need not be held on to beyond their immediate usefulness. In mental addition, the current sum ceases to be of interest when a new number has been added to it, and we have no need to remember every word we use during conversation. It is for this reason that Just and Carpenter made an appropriate choice when they referred to working memory as *the blackboard of the mind*.

Evidence that has been accumulating in recent years suggests that working memory requires the participation of the *prefrontal lobes*. The main source of this information has been observations on people with injury to that part of the cortex. One of the techniques most frequently used in the study of working memory is known as the *delayed-response test*. The subjects have often been monkeys, because they appear to possess working memories that are reasonably similar to our own; they suffer similar deficits when they have sustained damage to their prefrontal cortices. The essence of the delayed-response test is that it gauges the animal's behaviour in a situation requiring reaction to something experienced in the recent past. In a typical experiment of this kind, a monkey is first trained to give priority to a centrally-locat-

ed spot on a fluorescent screen, and to fix its gaze on this irrespective of what other objects are displayed elsewhere. The monkey then fixates on the spot while a single target object, a cross for example, is briefly flashed in one of the corners. After a further time lapse of several seconds, the central spot is extinguished and the investigator checks to see if the monkey then shifts its gaze to where the target had appeared. It will do so only if it has been able to retain an internal representation of the target's location.

The prefrontal cortex clearly plays a part in this ability because monkeys with lesions in that area of the brain cannot remember where the target appeared. As Goldman-Rakic (1992) has put it, *out of sight* for these animals means *out of mind*. The ability is also lacking in infant children and in monkeys below the age of two months, and it is satisfying to note that synaptic development in the monkey's prefrontal cortex is most rapid between the second and fourth months of life. Investigations of activity in the prefrontal area at the neuronal level, during delayed-response testing, have shown that different neurons respond at different phases; some become active as soon as the target is presented, while others are active during the few-second delay period, and those of a third group remain dormant until the monkey initiates its eye-shifting response.

Funahashi, Bruce and Goldman-Rakic (1989), have shown that each neuron in the second of these three groups has what they call a *memory field*: it reacts during the delay period only if the target happened to fall within a certain visual area relative to the center of gaze. Collectively, the neurons in that second group look after the spatial aspect of working memory. These neurons lie in the vicinity of a region known as the principal sulcus. It was subsequently demonstrated that there is another area somewhat below the principal sulcus in which the neurons respond to the attributes of objects rather than to their spatial positions. Examples of such features are shape and colour. For either set of neurons, it is the case that initiation of activity follows automatically when the target has been briefly exposed, the only exception being in the event of a sufficiently strong distraction.

The question naturally arises as to what these neurons are accomplishing with their activity, as they bridge the interval between target stimulus and motor response. The answer appears to lie in the interactions between the principal sulcal and other brain regions. It had been established anatomically that the principal sulcus is reciprocally connected to several other areas, prominent amongst which are the major sensory regions, the premotor cortex, and the hippocampus. Moreover, the 2-deoxyglucose technique has been used to demonstrate simultaneous activity in the principal sulcus and the hippocampus during a delayed-response test, with its attendant demands on the ability to rapidly update information. Just as importantly, it was found that the neuronal activity level in those same regions is much lower when the monkey is merely being required to associate.

One of the sensory regions discovered to be implicated in this shared activity is the *parietal lobe*. This is significant because that area is known to make a decisive contribution to the general sense of awareness. The involvement of the hippocampus is equally intriguing, given the structure and possible function of that region. Goldman-Rakic and her colleagues believe that the hippocampus consolidates new associations, whereas the prefrontal cortex somehow retrieves the products of previous associations from elsewhere in the cortex. What is still lacking, however, is a detailed picture of how the activities in the various areas, and especially the interactions between them, conspire to produce the faculty that working memory provides. The fact that the mechanism involves both the brief retention of things present and

the rapid retrieval of things past is suggestive of something fundamental. It would seem to touch on the fascinating issue of how one *knows*.

It appears to be possible that *everything* required to characterize observable features of the external world is served by one or another region of the prefrontal cortex. There are regions which variously handle position, shape and colour. Knight has located another which looks after auditory stimuli. And he has investigated the electroencephalographic (EEG) responses of patients with prefrontal lobe lesions to what is sometimes referred to as *the oddball paradigm*. The subject is exposed to a regular series of sounds, such as a succession of clicks, and every so often one of these is substituted by something unexpected. The latter might be a click of a different tone, or it might even be the omission of a click altogether. When normal people are subjected to such a disturbed series, their EEG traces show a characteristic surge in amplitude at around 300 milliseconds. The corresponding traces for patients with prefrontal damage have a much smaller amplitude around the 300 millisecond mark, and in some cases this part of the trace is missing altogether. A strongly truncated amplitude around the 300 millisecond mark is also observed in many autistic patients. This may indicate that autistics suffer from a deficit in their working memories.

Structural damage to a cortical area naturally represents a rather gross disruption of the system. More subtle disturbances occur when the trouble lies at the biochemical level. Goldman-Rakic and her colleagues have established that the prefrontal cortices of monkeys are rich in dopamine, which is a prominent member of the catecholamine family of neurotransmitters. It has transpired that upsetting the distribution of dopamine in the prefrontal cortex impairs working memory, in much the same way as is observed with physical lesions. In aging monkeys, such an imbalance occurs naturally, and poor performance in delayed-response tests is one of the consequences of advancing years. If these senior animals are given dopamine injections, their abilities are restored to the level seen in young monkeys.

It is interesting to note that dopamine imbalance had previously been linked to schizophrenia, and it has been established that malfunctioning of the prefrontal cortex is the cause of that affliction. It is especially noteworthy that the typical characteristics of schizophrenia, such as thought disorders, inappropriate or truncated emotional responses, lack of initiative, reduced attention span, and an inability to plan ahead, bear a striking resemblance to the abnormalities seen in patients with damage to the prefrontal region. Schizoid patients can manage routine procedures, and their reflexes are intact, but they perform poorly in situations requiring verbal information or symbol manipulation. They also do rather badly on delayed-response tests.

Schizophrenic patients have been subjected to a particularly revealing type of test, namely a measurement of their eye-tracking ability. It has been discovered that they are quite bad at anticipating the projected trajectories of moving objects. This too is rather suggestive in that the eye-movement neuronal groups involved in such predictive tracking are located in the posterior region of the prefrontal cortex. Indeed, it is found that the same deficit is present in monkeys with damage to that cortical area. In view of the fact that these various pieces of evidence are in such harmony, Goldman-Rakic (1991) has suggested that the time is ripe for a change of view regarding the Schizophrenic syndrome. She prefers to see the impairment as a failure of the link between representational knowledge and behaviour. The neural connections in the prefrontal region are continually updating our inner models of reality, because of the constantly changing circumstanc-

es in which we find ourselves. Those connections serve short-term memory, and thereby guide behaviour. If they fail in this task, perception becomes fragmented and a continuous train of events is seen as a series of snapshots. This produces the familiar symptoms of schizophrenia because behaviour is then dominated by the impulses of the moment, rather than by the evened-out influences of past and current information.

When considering memory, we are confronted with a series of antithetic pairs. First there was short-term memory, as opposed to long-term memory. Then there was declarative (explicit) memory, as distinct from non-declarative (implicit) memory. And we have just been considering working memory, which presumably has its counterpart in what could be called storage memory. Squire (1987) has further discussed semantic memory, which contrasts with episodic memory, and there are also active memory and passive memory. It is still too early to say whether all these types really are different, and we do not know whether there are still other forms waiting to be discovered. It will be exciting when a proper synthesis becomes possible.

Attention

Cortical Mechanics

It is not often the case that a qualitative picture of a phenomenon is sufficient for a full understanding of what that phenomenon entails. Far from it. If study of the phenomenon has not yielded the relevant numbers, and if these have not been accounted for in a quantitative manner, what appears to be understanding may well be misunderstanding. Before the early astronomers started to attach numbers to their observations, for example, interpretations of celestial phenomena tended to be subjective at best. There was insufficient contrast between astronomy and astrology. The situation changed dramatically when Kepler put the planetary trajectories on a firm mathematical footing. By deriving a reliable prescription for determining when a planet would reach a given position, and thereby adding a predictive dimension to the issue, he established the science of celestial mechanics.

During the hundred or so years that have passed since the discovery of the neuron, the phenomena associated with neural function have been studied with increasingly sophisticated techniques. And each advance has furthered our understanding. The invention of the cathode ray oscilloscope, and its use in conjunction with suitable electrodes, paved the way to elucidation of the microscopic events that underlie the nerve impulse. This was possible because experimenters had been given access to time intervals well below the millisecond duration of that impulse. A recent and further advance in that particular technology is currently permitting neuroscientists to simultaneously record from many neurons, and the fascinating things that have come to light will be concerning us later when we consider cortical oscillations. There is also still much to be learned from single electrode experiments. I am going to give an account of what they have

recently revealed about the speed with which the cortex goes about its cognitive function. It is now possible to determine when signals reach a given position in the cortex, and this has given quantitative significance to the term *cortical mechanics*. Before starting on that, however, it would be advisable to reflect on what we already know regarding the characteristic times of the nervous system.

The nerve impulse occupies about a millisecond, so this is a natural choice for the unit of duration. Impulses are generated by a neuron at a rate determined by the amount by which its threshold has been exceeded, the frequency saturating at very roughly a hundred per second. The shortest interval between successive impulses is thus around 10 milliseconds. For a myelinated axon, the conduction speed can be as high as a hundred meters per second, whereas the figure is far more modest for unmyelinated axons, namely a couple of meters per second. If I prick my finger with a pin, the fastest signals can thus make the approximately one-meter journey to my brain in about 10 milliseconds, via what is called the *A-delta route*, whereas the slower signals that convey the sense of pain travel through unmyelinated (*C-fibre route*) axons and they take about 500 milliseconds. These numbers obviously have a bearing on the speed with which one can respond to sensory input, but because reactions involve muscular movement, other processes must be involved as well. And it is here that uncertainty creeps into the picture. My reaction to that pinprick would generally come before those 500 milliseconds have elapsed, as Libet and his colleagues have established (Libet, Wright, Feinstein & Pearl 1979, Libet 1982), and yet I am left with the impression that I have responded to the pain rather than to the piercing of my skin.

The situation is further complicated by the fact that our nervous systems are able to anticipate the course of events through the use of internally-generated models. This is probably one of the main functions of the cerebellum. When we undergo tests of reaction time, therefore, our responses are as much a product of this faculty for internal modelling as they are of signal transmission from our sensory receptors. Moreover, it is vital to distinguish between tests that involve simple reactions and those in which there is also an element of discrimination, or even choice. In 1868, Donders determined a duration of 50 milliseconds for reactions involving discrimination, and 150 milliseconds when selection was added. Experimental psychologists now take these figures with a pinch of salt, because of those anticipatory factors. Indeed, tests of pure discrimination are looked upon as being more fruitful than those in which the subject is also being asked for a quick reaction.

A particularly interesting case of pure discrimination was published by Shibuya and Bundesen (1988). It was an example of what is referred to as a *partial report experiment*, the subjects being asked to discern between randomly-arranged digits and letters, in displays which were presented on computer screens for different periods of time. More specifically, the task was to report the digits (called the *targets*) from a circular array of digits and letters (called the *distractors*), the brief exposure of the arrangement being terminated by what is known as a *pattern mask* (i.e. replacement of the digits and letters by meaningless symbols of the same size). The number of correctly reported targets was analyzed as a function of the total number of presented targets (2, 4 or 6), the total number of presented distractors (0, 2, 4, 6, or 8), and the exposure time.

Shibuya and Bundesen rationalized their observations with a theory known as *the fixed-capacity independent race model*, and the outcome of their work was a set of quite intriguing numbers. The visual processing capacity of normal human adults was found to be 45 items per sec-

ond, (for targets and distractors of that particular complexity), which is equivalent to just over 20 milliseconds per item, while the short-term storage capacity turned out to be a mere 3.5 items. The other product of their work was a value for the longest *in*effective exposure duration; this was 18 milliseconds. It seems reasonable to interpret this latter period as being the shortest time required for something in the visual image to make any impression at all on the cognitive system. This is apparently what Coltheart (1983) refers to as iconic memory. From what we noted above regarding the frequency of emission of nerve impulses, this period would span a mere two or three action potentials. This does not mean that the act of cognition is *completed* in those 18 milliseconds, however. The number simply measures the *temporal span* of the signal which subsequently generates cognition higher up in the system. It is nevertheless fascinating that two or three action potentials should provide the system with enough to work on.

Those two or three impulses subsequently travel from the retina to the lateral geniculate nucleus (LGN) of the thalamus, and thence to the primary visual cortex. They thereafter pass through the hierarchy of visual areas, in a parallel and distributed manner. They ultimately impinge upon the association areas, there to elicit the appropriate response from the motor system, be this a limb movement, a spoken word, a written word, or whatever. The question obviously arises as to *when* these signals reach the various cortical areas; we would like to know something about the underlying mechanics. The remarkable fact is that it is now possible to supply the numbers. The basis for this possibility has arisen from the fact that places have been located in the cortex in which there are neurons that respond to well defined features in the visual field. An example that is rapidly becoming a classic is the presence in the inferior temporal visual cortex of neurons which are selective for faces. (Here, and later, we should bear in mind that the temporal cortex handles information on *what*, while the parietal cortex deals with *where* and *when*.) This has been the remarkable achievement of Gross and Rolls, and their respective colleagues, the studies having been carried out on monkeys (Gross, Bender & Rocha-Miranda 1969, Perrett, Rolls & Caan 1982).

Once such a face-selective neuron has been located, it is a relatively straightforward matter to determine the arrival time for signals at different intermediate points on the route from the retina. A picture is flashed in the visual field of the attending monkey, and the (electronic) stopwatch is started at the same instant. Electrodes are already in place at the positions of interest, it having been established beforehand that they are impaling neurons which are involved in the responses to the picture in question. The arrival time at the inferior temporal visual cortex itself lies around 100 milliseconds. The timetable for stations *en route* is 40 milliseconds for V1 (the primary visual cortex, which is the Brodmann area 17), 55 milliseconds for V2 (the secondary visual area), 70 milliseconds for V4, and 85 milliseconds for the posterior inferior temporal visual cortex (Rolls 1992).

These numbers are quite enlightening. We see that the signals can progress from one cortical area to the next in as little as 15 milliseconds. Examination of the underlying circuitry shows that at least two synaptic junctions would have to be traversed by the signals as they pass through the different layers of a single cortical area. It seems reasonable to add a further synapse for each transit from one area to another. This gives us a total of three neurons and their associated synapses to be crossed for each additional cortical area visited. Given that the travel time down a dendrite, across the soma, out along an axon (this latter part occupying an admittedly negligible amount of time), and

across a synapse would not be less than about 5 milliseconds, we find that the measured transit time between cortical areas can just about be accounted for.

The fascinating thing about these transit times is that their duration roughly corresponds to those same two or three action potentials that emerged from the work of Shibuya and Bundesen. Each cortical step has to respond on the basis of this small number of signals. We may conclude, therefore, that the perceptual process is impressively efficient. Just as interesting is the fact that the looped circuits made possible by the reciprocal projections, discussed by Van Essen, Anderson and Felleman (1992), and Felleman and Van Essen (1991), do not participate in these *initial* stages of the cognitive process; there is simply not enough time available during the 15 millisecond inter-area transit time to fully activate them. This is not the same as saying that they do not *ultimately* play an important part in the cognitive process, however, because even the above-quoted 100 millisecond period is insufficient to establish conscious awareness. In contemplating this disparity, we are getting the first inklings of the distinction between preattention and attention.

A further important point is that the ten millisecond interval between adjacent action potentials applies only in the case of maximal output frequencies from the neurons involved. We can turn this around and say that *only* those neurons which are driven at their maximum rates can be contributing to the perceptual process. This provides a more quantitative criterion for the winner-take-all idea that is now so common in analyses of neuronal processes.

In the human being, the times consumed by two further processes are of interest. The first is that of the onward passage of signals from the inferior temporal visual cortex to the appropriate association areas, which can be looked upon as the deepest recesses of the cortex, the regions most remote from the sensory inputs. This remainder of the afferent (inward) journey would take about 15 milliseconds if, as seems possible, the inferior temporal visual cortex and the relevant association areas are adjacent in the hierarchy. That gives a total of 115 milliseconds for the trip from the receptors to the association areas, in the case of the visual system. Subtracting the 40 milliseconds it takes for the signals to travel from the receptors to V1, we get a time of 75 milliseconds from V1 to the association areas. That number is going to be important in our further contemplation of the cognitive mechanism.

The second time-consuming process is that of the signals' subsequent progress from the association areas to the cortical regions responsible for generating the response, which is an efferent (outward) journey. To see what this involves, let us consider what happens after the signals have reached the association areas. By a mechanism which is still a mystery, and which we must therefore discuss, a visual *percept* will somehow give rise to a *concept*. After suitable associations have been struck up, the latter could ultimately lead to the utterance of a word or a phrase. We know that this reaction involves both Wernicke's area and Broca's area, and it ultimately leads to activation of the relevant muscles underlying vocalization, including those of the diaphragm. The trouble is that we do not have direct measurements of the amounts of time these latter efferent processes require.

One can make a very rough guess at their duration, however, by comparison with a process that maximally involves both afferent and efferent signalling. When the hearing of a word conjures up a visual image, afferent signals travel from the auditory receptors to the association areas, and efferent signals then travel from the association areas to the early visual areas. (The involvement of the latter part of this route is implied by the fact that the ear-

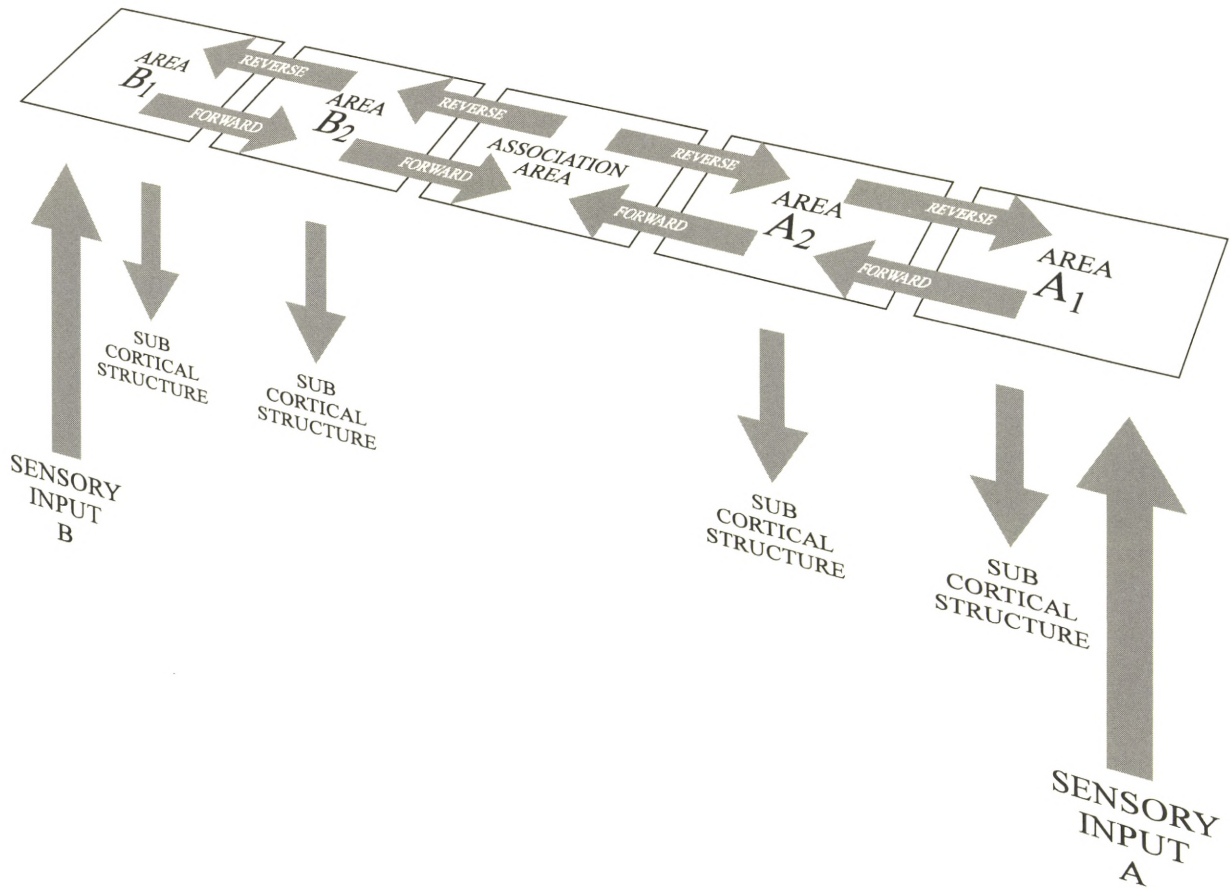


Figure 1 In this highly schematic view of part of the cerebral cortex, each of two different sensory modalities, A and B, is served by two successive areas, arranged hierarchically. (No attempt has been made to show the multiplicity of cortical areas present at some levels of the hierarchy; these *are* shown in figure 4.) All inter-area links in both routes involve forward and reverse projections, and they ultimately feed signals into an association area. Each area also sends projections to a suitable subcortical structure, although in practice this might not be the case for all areas. (These sub-cortical projections, which invariably stem from layer V of the cortex, are believed, *inter alia*, to mediate the feelings generally referred to as qualia.) The key point of this diagram is that the reverse projections of sense A appear to be forward projections for sense B, and vice-versa. (See also Cotterill and Nielsen 1991).

ly visual areas are involved in visual imagination (Zeki 1992).) If we assume that the number of synapses between the receptor cells serving *any* sense and the relevant association areas is pretty much the same, irrespective of the type of sense, it seems logical to conclude that our estimated value of a roughly 115 millisecond interval from receptors to association areas would *always* apply. But the forward projections of one sensory modality are the backward projections of another (see figure 1). If we further assume that the rates of progress of signals do not depend upon whether a forward or a backward projection is being used, the time lapse between sound waves striking the ear and a visual image being created would be roughly

180 milliseconds, this number being composed of the 115 millisecond afferent part and the 75 millisecond efferent part. It must be emphasized that this is little more than a guess, but the value does happen to be in quite good agreement with the 200 milliseconds often quoted for a simple cognitive process. If we make the further gross assumption that all efferent target areas are roughly equidistant, in synaptic terms, from the relevant association areas, that value of 75 milliseconds might not be too bad an estimate for the missing efferent duration in the case of muscle activation.

Even if this exercise in approximation can be looked upon as being satisfactory, it does not mean that our understanding of the cognitive process is adequate. That happy state of affairs is still some way off, for it is quite literally the fact that there is more to this issue than meets the eye. Let us return to the fascinating work of the Gross and Rolls groups. They reported the existence of neurons responsive to faces, amongst other things. They said nothing about neurons specifically responsive to vases, but it seems safe to assume that such neurons would have been discovered if they had been looked for. The question is, why on earth should I bring vases into the story? What I have in mind is the intriguing sort of picture that Edgar Rubin gave prominence to in 1915. One of the best known of these looked like the example shown in figure 2, and a little contemplation reveals its ambiguity; depending on what is regarded as the background, the figure shows either a vase or two counterdisposed faces. The way in which our visual systems distinguish between an object and the background was of great concern to *Gestalt psychology*, which was founded in 1912, the pioneers being Wertheimer, Koffka and Köhler. The term is German, and the usual English translations *shape, pattern* or *form* fail to do adequate justice to what is really a rather subtle concept. The Gestalt is the *wholeness* or *configuration* of a figure, and it is this which enables

Figure 2 This celebrated picture, first published by Rubin in 1915, can be interpreted in two ways. It shows either a black vase against a white background or two white counterdisposed faces against a black background. It is impossible to perceive both alternatives simultaneously, but switching from one interpretation to the other can be achieved in as little as about half a second. (See Rubin 1915).

one to separate it from the background. It is what would now be referred to as an *emergent property*, the idea being that the whole is more than the sum of its parts.

Let us return to those face-detecting and vase-detecting neurons, both presumably located in the inferior temporal visual cortex, and both therefore triggered around the 100 millisecond mark. We now see that it cannot be their mere activation which produces even a percept, let alone a concept, because we are aware of *only one* of the interpretations at any instant. We see the face *or* the vase, but not both simultaneously. There would thus have to be an additional passage of time, during which the cognitive process was completed. The duration of that supplementary period has not been measured, but there again appears to be the possibility of a rough estimate. Using a stopwatch, I have timed myself over ten reversals from face to vase, interspersed with ten reversals the other way. Making an intense effort to do this as rapidly as possible, and repeating the entire thing five times in order to achieve good statistics, I find that I can get the average reversal time down to 550 milliseconds.

During this test, I noticed that the center of

my gaze was systematically shifting from the region immediately between the two noses to a point somewhat higher up, around the center of gravity of the vase, and back again. Such *saccades*, as these eye movements are called, have been measured to occur every 250 milliseconds or so when the eye freely scans over an object, and the actual movement occupies less than 50 milliseconds. This movement is therefore too rapid to have been a major factor in the 550 milliseconds of those cognitive reversals.

What is going on during that approximate half second? We cannot be sure at present, but we can at least make a simple calculation of what would be possible, using the rudimentary cortical mechanics that we have been establishing. In 550 milliseconds, there would be time for seven or eight of those 75-millisecond V1-associator or associator-V1 journeys, or nearly four V1-associator-V1 round trips. (The recall of visual images would also involve early visual areas other than V1, and the return-trip transit times for these would be shorter. One could think of the early visual areas as functioning together in such a process. I will loosely refer to them collectively by the *unofficial* designation EV.) If that is indeed what is happening, it would not merely be a case of the same signals passing backwards and forwards along this route; such ping-pong dynamics would achieve nothing, apart from possibly supplying the cortex with a *reverberation period* that could have relevance to *very short-term memory*. It seems likely, however, that the signals would be progressively modified as they run to and fro between the various cortical areas. Whether or not part of that modification has to *undo* the old concept before it can make a start on establishing the new one is a moot point.

Bandwidth and Breakthrough

Treisman's work provides a good starting point for considering the very important phenomenon of attention (Treisman 1960, Treisman 1964, Treisman & Gelade 1980). She too used patterns that were briefly exposed to subjects, but in this case the patterns were somewhat more complicated. In one of them, the target was a red or blue H, say, amidst a number of red O and blue X distractors. Interest centered not only on the subject's ability to identify the target, after brief exposure to the display, but also to properly locate it once it had vanished from the screen. It transpired that errors in remembering the target's location were more frequent than mistakes in simply identifying it. In a second set of displays, the target was a blue O or a red X amongst a number of red O and blue X distractors. This task was fundamentally different from the previous one, because the target had something in common with each of the two types of distractor: form with one and colour with the other. In such a display, one speaks of features being *conjoined*. The subjects' scores for this variant of the test were quite different from those for the previous challenge: correct identification was now strongly correlated with correct localization. This demonstrated that attention has to be focussed on a specific location if the features of a target are to be mentally combined, and thus properly mimic the actual conjunction in the display.

These studies were obviously related to what was discussed in the preceding section. The various elements of the displays were naturally being handled preattentively, and they were then being combined further up in the visual hierarchy. One refers to such consolidation by the term *bottom up*, and because conjunctions of letters with colours is an abstract exercise, bottom-up consolidation was probably the only process operative in these tests. As Treisman emphasized, many conjunctions have the additional dimension of *prior knowledge*. Blue bananas and red eggs are not things that one often encounters, and if such items occur in test dis-

plays, one could imagine that they will meet with a certain amount of resistance, from the cerebral powers that be. One speaks of *top-down constraints* in such cases. Biederman has shown that identification of objects occurs more rapidly if top-down strictures are not being defied.

Treisman and her colleagues have explored top-down influences further. They used displays in which three target figures, equally spaced along a line, were flanked by two distractors. The distractors were digits, while the targets were simple geometrical shapes which could nevertheless be linked with familiar objects, the idea being to thereby introduce the possibility of top-down influence. An elongated isosceles triangle will be reminiscent of a carrot if it is coloured orange, for example, while an ellipse will resemble a lake if it has the appropriate shade of blue; a thickly-drawn circle could be an automobile tire, if it is black. The entire five-member display was obliterated by a pattern mask after exposure to the subject for 200 milliseconds, a pointer simultaneously indicating the position previously occupied by one of the three targets. The subject was required first to report which two digits had been used as the distractors, and then to identify the object that had lain in the position indicated by the pointer.

The subjects in these experiments were divided into three groups, which were required to perform under different test protocols. Members of the first group were told that they would be seeing carrots, lakes and tires, but that a quarter of the objects would appear in the wrong colour, to preclude short-cut identifications through colour alone. The second group was told that the objects would be orange triangles, blue ellipses and black circles, again with a certain fraction of deviant figures such as orange ellipses. Members of the final group were made to expect only the natural colours; their operational paradigm was the same as for the first group, except that they were not informed beforehand that false colourings might occur. Not surprisingly, the first group scored better than the second; their reports included fewer erroneous identifications. The real interest centered on the third group. Treisman and her colleagues discovered that the subjects did not generate illusory conjunctions to fit their expectations. In cases where the pointer indicated the position of a triangle that was unexpectedly blue or black, the likelihood that they would erroneously report it as having been orange was quite independent of whether that colour had appeared elsewhere in the display.

Treisman and her colleagues drew two fundamental conclusions from this test: prior knowledge and expectations aid attention in properly conjoining features of a target figure, but they do not promote illusory exchanges of features in order to make unexpected combinations revert to what was anticipated. Illusory conjunction thus appears to arise at an earlier stage of visual processing than that which has semantic access to knowledge of familiar objects. Conjunction of features apparently occurs in the bottom-up manner characteristic of preattention; it seems not to be subject to top-down interference. But how does the bottom-up mechanism feed into the upper echelons of the hierarchy? The term cognitive compression is used when describing what happens to the inward-travelling information. Let us consider an example cited by Treisman herself. A bird perched on the limb of a tree is perceived to have a certain size and shape. As it flies away, these attributes change, and yet we still see it as the same object. We can even cope with the radical change in colour that can occur as the bird unfolds its wings. How does it maintain its perceptual integrity in the face of such variability? What mechanism insures that the cognitive compression will nevertheless funnel these changing sensory inputs into the same conceptual pigeon-hole?

Treisman and her colleagues have suggested that the intermediate consolidation process involves what they called a *spotlight of attention*. (Some authors prefer the term *searchlight of attention*. This conveys the impression of a slightly more active role, which may or may not be appropriate.) The crux of their model is the construction of a temporary representation that is specific to the object's current appearance, this being continually updated as the object changes. The temporary representation's contents are then passed up the hierarchy, for comparison with what is already stored in memory. Treisman compared this latter part of the process to the scrutinizing of a file. Continuity of the perception of an object will be guaranteed if the contents of the spotlight continue to fit one particular file better than they fit any other. In order to test this idea, Treisman and her colleagues devised a test in which identity had to be robust against change of position. Two different letters were briefly flashed on a screen, each surrounded by a square that remained visible when the letters disappeared. The empty squares then moved to new positions, and a letter was briefly flashed in one of them. This second singleton letter that fleetingly appeared in the randomly selected square could either be the same letter that had been there the first time or it could be different.

When there has been prior exposure to the same letter, that letter's first appearance is referred to by the term *priming*. It is well known that priming speeds up recognition; if the second letter is not identical with the primer, recognition takes slightly more time. The key point in the experiment was whether priming is effective only if it has occurred in the same square. This is just what was observed; subjects recognized the match an average of 30 milliseconds faster if it was taking place in that location. This confirmed that the conjunction of the letter's identity and its link to a particular square was being maintained despite the movement of the square. Indeed, it is as if the visual system is attributing movement to the letter as well, by association.

As a result of these experiments, Treisman proposed a scheme for visual processing which accorded the spotlight of attention a decisive role. She saw the presence of individual features of a stimulus as being signalled without this necessarily including an indication of their whereabouts. She also conceived of a *master map of locations*, one area of which was instantaneously highlighted by the spotlight. According to her theory, only that area currently being highlighted can be the goal of instantaneous links from the various features. It is the momentary establishment of those links which gives the spotlight its content, and it is that content which is of relevance for the abovementioned file. Finally, Treisman suggested, the contents of that file are compared with descriptions stored in a recognition network, that is to say in an appropriate part of memory. This network specifies the defining characteristics of the things involved, and it supplies their names, their likely behaviour, and their current significance. According to Treisman's view, *it is the interactions between the file and the memory bank which are the basis of conscious awareness.*

The various stages of the hierarchy invoked in Treisman's model might appear to be identifiable with the various stages of the visual system (if it is vision that is implicated, of course). In her scheme, the feature modules come before the map of locations, that map comes before the object file, and the object file comes before the recognition network. Could it be that the feature modules are simply in the primary visual cortex, that the map of locations and the object file are in the subsequent visual areas, and that the recognition network constitutes the appropriate association cortices, functioning collectively? I believe that this would be

too facile an interpretation of her ideas. In my view, there are a couple of things which militate particularly strongly against such direct equating of her levels and the chain of areas that undoubtedly exists in the cortex. For a start, such a scheme leaves no role for the reverse projections. Two-way traffic of neural signals occurs in the Treisman strategy only between the uppermost two levels, that is to say between the object file and the recognition network. Two-way traffic in the cortex, on the other hand, is seen between many of the areas, even in those which receive the sensory input. Even more seriously, Treisman's conclusion that the feature modules might not be place-specific is difficult to reconcile with the actual situation in the cortex; patches earmarked for colour, orientation and direction coexist side by side in the quilt arrangement of areas now known to be present in the cortex (Felleman & Van Essen 1991).

There is a third, rather more subtle, objection that can be levelled against too direct an interpretation of the attractive model that Treisman proposed. As a matter of fact, its roots can be traced back to the work of Perky (1910), and plenty of corroborating evidence has emerged since that time, most notably the recent extensive work by Zeki (1992) on the various visual areas. When we imagine something, we often speak of seeing something *in our mind's eye*. The point of Perky's investigation, and of the experiments that followed in its footsteps, is that it showed that the mind's eye and vision's eye have a lot in common; we apparently use the same areas when imagining as we use when seeing. The cortex is apparently parsimonious with its areas. It is this economy which necessitates a more subtle interpretation of what Treisman was postulating. It means, I feel, that her feature modules and her map of locations might share the same part of the cortex. If this is so, the unidirectional conduits which she conceived as passing signals from the former to the latter might in fact be two-way connections. They might be a combination of the forward and reverse projections that we know are a very common attribute of the cortex's hard wiring. Such a scheme would permit us to use the fact that things like colour, orientation and direction *are* spatially mapped in the primary visual cortex. Before continuing with this alternative picture, however, we ought to consider other evidence that has a bearing on this quite central issue.

We could hardly choose a more appropriate starting point than the pioneering contributions of Broadbent (1958, 1982). He investigated a number of situations which strongly pointed to essential limitations in the nervous system's ability to handle signals (see also Duncan 1980). We instinctively allow for these limitations, hardly realizing that we are doing so. We sometimes close our eyes in order to concentrate on a piece of music, or on something being said. Conversely, we turn off the radio when we wish to give all our attention to something we are reading. The distract*ors* used by the experimental psychologist and the distract*ions* of everyday life are, of course, close cousins. Broadbent's early work in this area was motivated by the type of practical problem that accompanied the development of communication networks and the growth of mass transport. A typical example is seen in the case of the flight controller, who must make rapid decisions while being bombarded with numerous streams of information.

Experimental work in this area benefitted from the invention of the tape recorder and the development of stereophonic techniques. These advances gave Broadbent the means of varying the rate at which competing flows of auditory information are fed to the two ears. He used what is known as a *split span protocol*, three (prerecorded) spoken digits impinging on the left ear while a different set of three is simultaneously fed to the right. The task of the

subject was to recall as many of the six digits as possible, and it was found that the optimal rate of delivery was two digits per second. Despite the fact that the subject was being presented with three synchronous pairs of digits, Broadbent found that they were being recalled in two groups of three: the entire set for one ear, followed by all three digits for the other ear. His conclusion was that the subject attended to one ear first, and that the information simultaneously reaching the other ear could be retained in an unattended form of short-term memory for a few seconds. Attention was then switched to that other information, while retention of the first three digits became the responsibility of short-term memory. These observations formed the basis of Broadbent's *filter theory*, the gist of which is that the brain constitutes a single-channel information processing system with limited capacity. This channel, according to his findings, could select only one sensory input at a time, the switching of selections occurring at a maximum rate of about two per second. This is about the same as what I was able to achieve with that faces-vase picture that we encountered earlier.

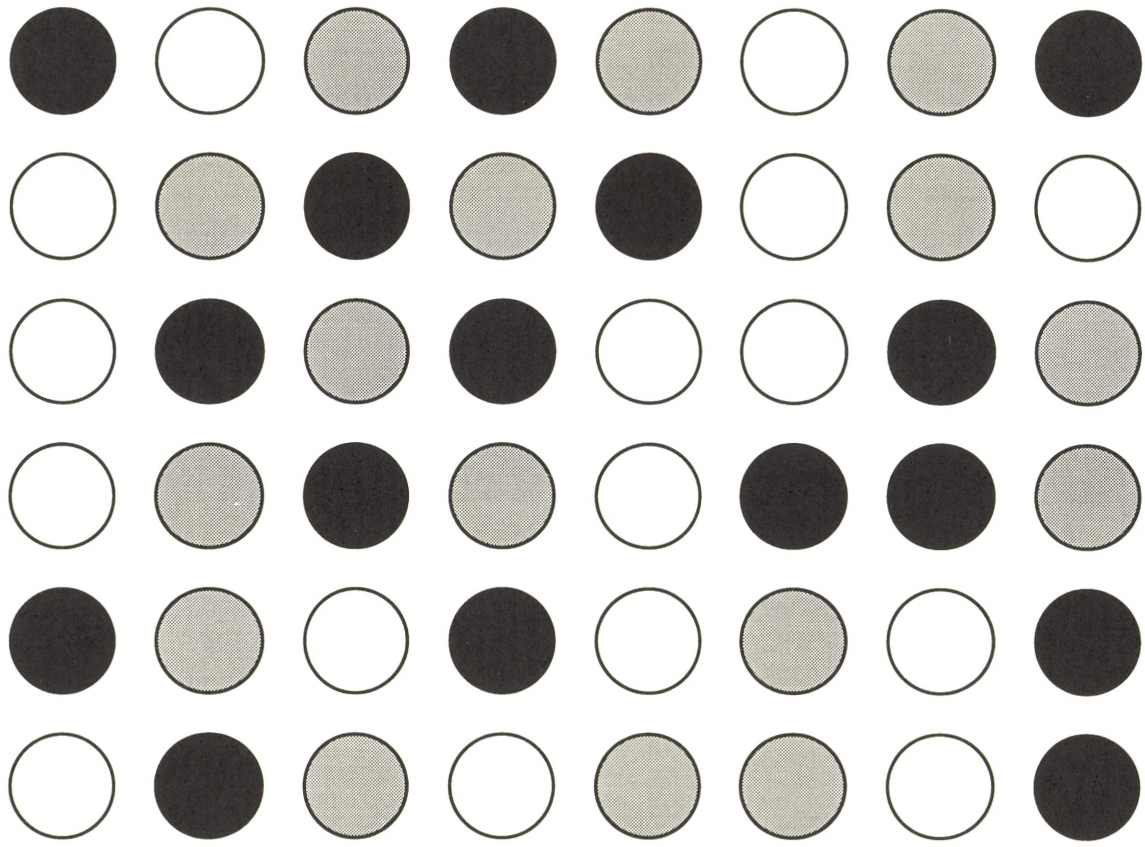
Limitations of information-carrying capability are often encountered in signal transmission technology, the measure of capacity usually being what is referred to as the *bandwidth*. Broadbent's conjectures are particularly reminiscent of a technique known as *time-division multiplexing*, in which several information streams can be handled simultaneously, if it is permissible to pay full attention to any one of the streams only periodically (Sakrison 1968). Brief attention is rapidly and cyclicly shifted from one signal to the next, this mechanism being known as *gating*. The strategy works only if the persistence of any stream is longer than the time lapse between attentional shifts. If that is not the case, information will be lost. In time-division multiplexing, the bandwidth can be roughly defined as the ratio of the period over which the signals persist to the time lapse between shifts; this determines how many signalling streams can be coped with by this approach. In the Broadbent scheme just described, the persistence of a signalling stream was simply that of short-term memory. We will have to suspend until later discussion of what it is in the brain that determines the rate of gating, but from Broadbent's observations we can see that the rate is about two per second; the minimum time that must be given to any one thing is the period that we are already familiar with, namely *500 milliseconds*.

Broadbent's early work was followed up by studies of what is known as *speech shadowing*, the pioneer of this technique being Cherry (1953). Broadbent himself was active in this area, as were Treisman (1960) and Moray (1960). The subject is asked to repeat aloud a piece of text that is presented simultaneously with one or more distractor passages. A number of different factors were found to aid concentration on the correct signal, including tone of voice, the direction and distance from which the various signals were arriving, and the cadence of the delivery. Treisman (1960) discovered attributes of a distractor which enable the unattended passage to impose itself, at the expense of the correct text. One of these factors is emotional content, an example being the sound of one's own name. Another potent distractor is text that is closely related to what one is supposed to be concentrating on. This is rather revealing, because it touches on that issue of bandwidth. When two signals bear a close mutual resemblance, it might no longer be correct to regard them as being independent; the routes that they take through the nervous system might have a lot in common. This particular distractor phenomenon is quite logically known as *breakthrough* and we have good reason to be grateful for its existence. There are occasions when breakthrough can be vital to life and limb, indeed. Let us consider an example.

I am often shocked, when travelling in my car, to discover that my mind has not been on my driving. There are occasions when I am lost in thought, even though the car is moving at considerable speed. It is almost as if the vehicle were driving itself, with me as a passive passenger. There are other occasions when an impatient driver immediately behind me has to sound his horn, to shake me out of my reverie, and draw my attention to the traffic light that long since turned green. One would think that such absent mindedness would put me at constant risk. Yet the fact is that in thirty-five years of driving, I have never had an accident. What guardian angel has been watching over me all this time? The benevolent agency has probably been a combination of Broadbent's bandwidth and Treisman's breakthrough. The bandwidth is probably greater than he initially envisaged, however, and the threshold for breakthrough is probably higher in the driving situation than it is when two streams of speech are vying with each other.

Day dreaming when driving is especially interesting because the latter enterprise makes considerable demands on one's sight. It is thus mysterious that one can nevertheless see things with *the mind's eye*, at the same time. I say this because it has been emphasized already that the visual component of imagination appears to use some of the same cortical areas as those employed in sensory vision. How can it be that this double usage does not lead to confusion? What prevents scenes perceived with the real eyes and scenes imagined with the mind's eye from getting all mixed up? One factor must be the difference between what is being seen and what is being imagined, because that experiment performed by Perky (1910) showed that a mix-up *does* occur when the difference is inadequate. One could say that her experiment explored the limiting case, in which breakthrough was mandatory. When there *is* a difference, the possibility of keeping them separate must stem from the fact that the items are being handled by different sets of neurons, in the appropriate area of the cortex. We saw earlier that there is a cortical region in which groups of neurons are sensitive to faces, and it seems safe to conclude that there are other groups of neurons which handle roads, fields and trees. It is this separation that provides the safety factor when one imagines someone's face while driving through the countryside. Breakthrough is impressively illustrated in the type of test first described by Stroop, an example of which is shown in figure 3.

Distribution of perceived items amongst different cortical regions can only be part of the story, however. There remains the important question of when priority is being allocated during the processing of information; is selection made late, after the stimuli reaching the senses have been fully operated upon, or much earlier? If we could answer that question with confidence, we might be closer to an understanding of how the brain really handles sensory input. A strong hint is provided by situations in which the system is not provided with opportunities for selection. In such cases, interference is found to depend on individual features of the stimulus and not on the responses they provoke. We saw examples earlier, when considering Treisman's work. If the subject knows what part of a stimulus field to select, there is no interference between the various features of the input. The interference must thus occur before the choice of response has been made; selection is made early. We might have anticipated this by contemplating the cortical map. In vision, for example, the number of alternative paths confronting signals as they proceed up the hierarchy steadily increases. Early in the system, therefore, there must be more sharing of neuronal resources than there is later on. If we accept that there will be more interference between two tasks in regions where they are competing most vigorously, we see that the case for early selection is quite persuasive.



BLACK WHITE GREY BLACK GREY WHITE GREY BLACK

WHITE BLACK WHITE GREY WHITE BLACK WHITE GREY

BLACK GREY GREY WHITE GREY WHITE GREY BLACK

BLACK WHITE GREY BLACK WHITE GREY BLACK GREY

WHITE GREY BLACK WHITE BLACK GREY WHITE GREY

WHITE BLACK GREY BLACK WHITE GREY BLACK BLACK

Crick and Koch (1992) have recently described a delightful illustration of interference. For reasons which will become apparent shortly, it is known as the *Cheshire Cat experiment*, and it stems from what is known as *binocular rivalry*. The phenomenon can be observed with the aid of a convenient object, a mirror, and a featureless area such as a white wall. The mirror is positioned with one of its edges touching the tip of the nose and the middle of the forehead, as if it were about to slice vertically through the head. An object is now positioned so that it can be seen only by the eye lying on the non-reflecting side of the mirror. In the relaxed state, the other eye will also point in the direction of the object, but it will actually be receiving the uniform illumination from the blank wall; the subject will not be aware of the fact that the object is being seen only with one eye. If the hand on the reflecting side is now moved in front of the wall, so that it apparently passes through the part of the object actually being attended to by the other eye, that part of the object will briefly vanish. If the object is the face of a cat, and the moving hand passes across the appropriate portions of it, only the smile will not be briefly obscured. Hence the allusion to Lewis Carroll's famous animal. In this fascinating experiment, the hand's movement momentarily captures the brain's attention, apparently because motion is accorded the higher priority. The missing parts of the image rapidly reappear, once the distraction is gone. I have timed this restoration to occur within as little as a couple of seconds.

◀ Figure 3 Breakthrough, the interference of one percept by another, can be demonstrated with what is known as a Stroop test (Stroop 1935). In the example shown here, the perceptions are of colour and shape. When the shape is a circle, there is no conflict between the two percepts, and breakthrough does not occur. If shape is used to convey meaning, however, this may well clash with the perception of the colour, and breakthrough will be possible. That this happens in the case shown here can readily be demonstrated by comparing the times taken to read aloud the *actual colours* in the sequences above and below.

The question naturally arises as to the nature of the neuronal interactions that underlie binocular rivalry. The phenomenon has yet to be given a satisfactory explanation, but it seems reasonable to assume that inhibition must play a role. Perhaps the inhibition is of what could be called the *en passant* (or perhaps *spoiler*) type, with collaterals being used to activate conveniently-located inhibitory neurons, while the main axons project to the region of (their) primary interest. The neuronal pathway that carries the signals for motion detection might dispatch axon collaterals toward the pathway that handles the signals for stationary objects, and it might inhibit the latter through the mediating agency of inhibitory interneurons. There is some anatomical evidence, admittedly tentative, that could endorse this idea. The point is that investigations of the cat's visual system by Gilbert and Wiesel (1985), have revealed the presence of two separate pathways which have been designated the X and Y subsystems. Both X and Y pathways project to the primary visual area, which seems to be mostly concerned with the stationary attributes of an observed object, whereas only the Y pathway projects to the secondary visual area, which apparently lies on the main motion-detecting route. Now the spatial response of an X-route neuron displays high resolution, whereas it is temporally sluggish; it is well suited to stationary objects. Conversely, Y-route neurons show good temporal resolution but poor spatial discrimination; they are ideal for the detection of motion. The anatomical scheme just described therefore makes good sense, apart from that fact that the primary visual area also has Y-type neurons. They would seem to be unnecessary and misplaced. What are they doing in the primary visual area? Could it be that their mission is *en passant* in-

hibition? Are they there as potential spoilers, waiting to give motion priority if it occurs?

Much painstaking work will be required to settle that particular issue, but if such *en passant* inhibition does play a role, it would have important repercussions for theories of attention. In that binaural set-up employed by Broadbent, for example, it could explain how first one ear and then the other is attended to. To explain breakthrough, however, we would have to postulate that the inhibition can fail if the competing signals are too similar. Just how this would happen is not clear, but it might involve the reverse projections alluded to earlier. If this is the case, it would be an example of a higher cognitive process influencing the early stages of signal processing. As we have seen earlier, there is independent evidence for this sort of thing. Perhaps it is the occasional intervention of such a top-down mechanism which has made it difficult for Broadbent and his peers to settle the early selection vis-a-vis late selection issue.

The introduction of *en passant* inhibition into the story raises another interesting possibility. It could even force us to adopt a different attitude toward the sort of hierarchical cortical diagram that is now becoming familiar in the literature. The various parts of figure 4 are typical examples, for the macaque monkey (although the equivalent diagrams for the human cortex are not yet available). One is impressed by the welter of inter-area routes, many but not all of which are reciprocal. Could it be that a sizable fraction of these routes arise to serve the needs of *en passant* inhibition? If that is so, it could mean that the hierarchy actual accomplishes for *groups* of neurons what McCulloch and Pitts (1943) tried to attribute to single neurons. Perhaps there really is a sort of logic operating in the cortical hierarchy, but one which permits different cortical *areas* to variously compete or join forces. Being based on large groups of neurons, such a scheme would not be subject to the catastrophic vulnerability that so blighted the McCulloch-Pitts single-neuron model.

I mentioned in connection with the Cheshire Cat experiment that restoration of the disrupted image takes a couple of seconds. This puts that process on the outer reaches of our main interest here; we have been considering what happens in the first half second. However, the work of Broadbent suggests that other *mutually* inhibiting processes (as opposed to *en passant* inhibitory processes) may well operate around the half-second time scale. We saw that his experiment involved first listening to a train of sounds impinging upon one ear, and then switching attention to a different stream arriving at the other ear. We must now enquire as to what mechanism permits identification of the sounds being attended to by a given processing route. We were earlier identifying visual events that apparently occupy as little as 18 milliseconds, but nothing was said about the process which conspires to produce cognition from such brevity. When discussing the work of Bundesen and Shibuya, I concluded that a winner-take-all mechanism could govern things in the visual domain, because the time scale seems to fit this possibility. Could a winner-take-all strategy also apply to hearing? Before attempting to answer this question, let us check whether that brief time scale is also relevant to hearing.

Spoken words are composed of phonemes. At what rate can the ear intelligibly receive those atoms of articulation? The highest speed recorded in a public speech was achieved by the late President John F Kennedy, in December 1961 (McWhirter and McWhirter 1963). In one particularly rapid burst, he reached a rate of 327 words per minute. This corresponds to just under 200 milliseconds per word. Allowing an average of two syllables per word, this gives just under 100 milliseconds per syllable. If we further allow for an average of two phonemes

per syllable, we arrive at the final figure of just under 50 milliseconds per phoneme. Although this is slightly longer than the elementary time that emerged from the visual studies by Bundesen and Shibuya, it is nevertheless similar to the value we have encountered earlier for those switches in visual attention that do not involve eye movements. It does seem possible, therefore, that the underlying mechanisms rely on similar neuronal processes. Just what these are, and whether they do indeed involve the winner-take-all strategy, is not yet known. But

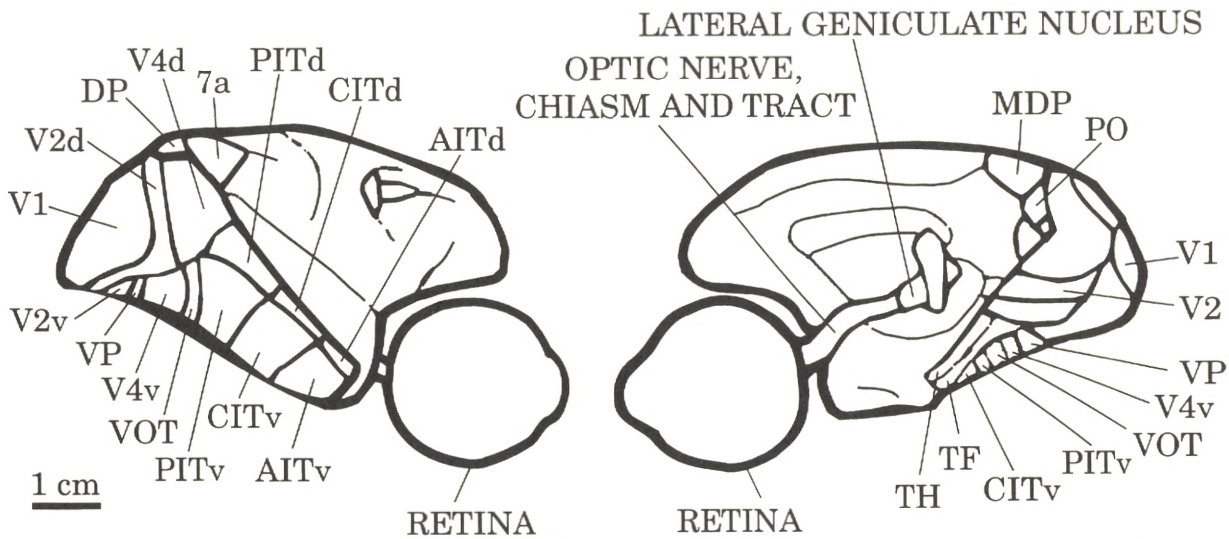
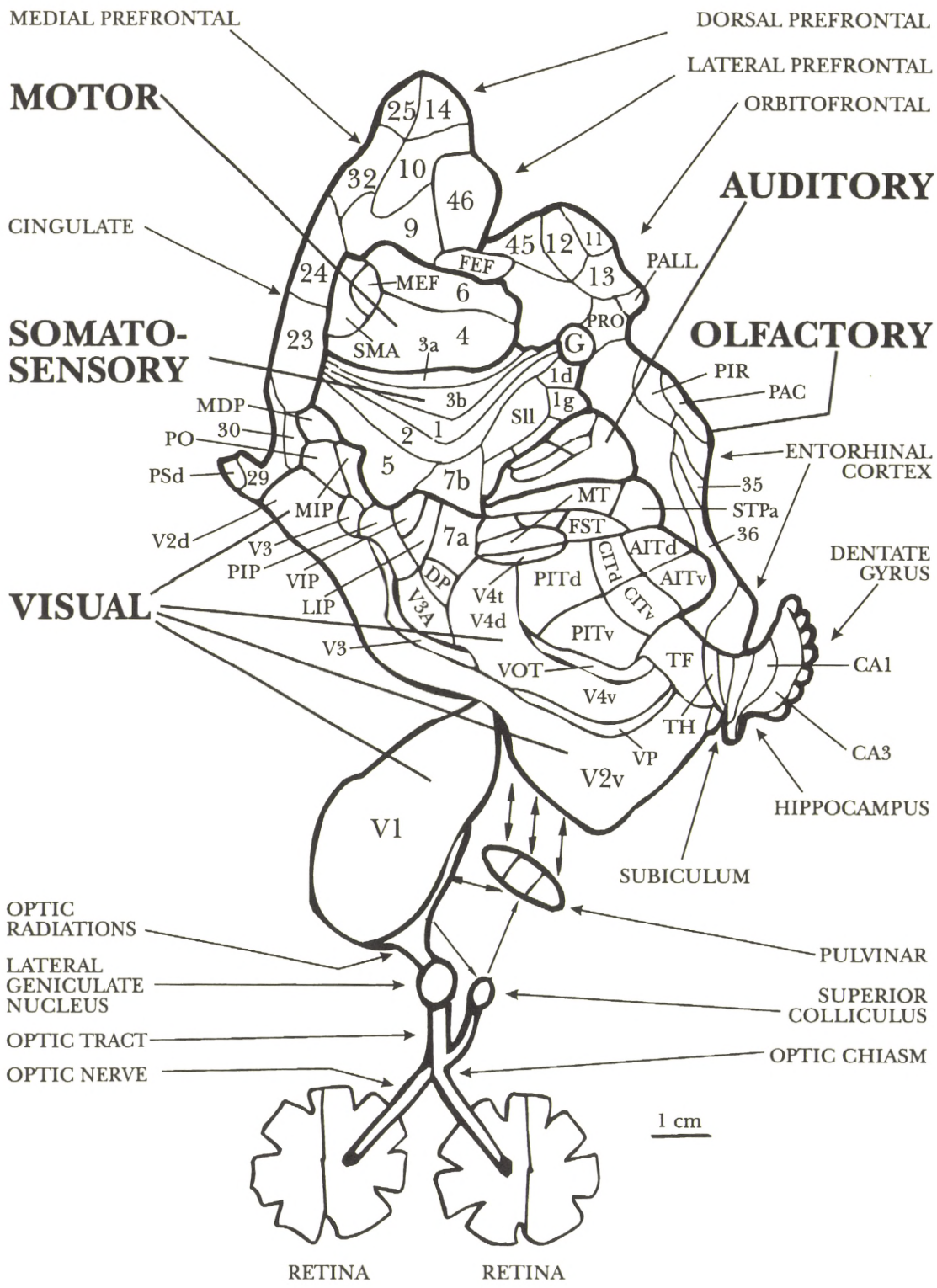
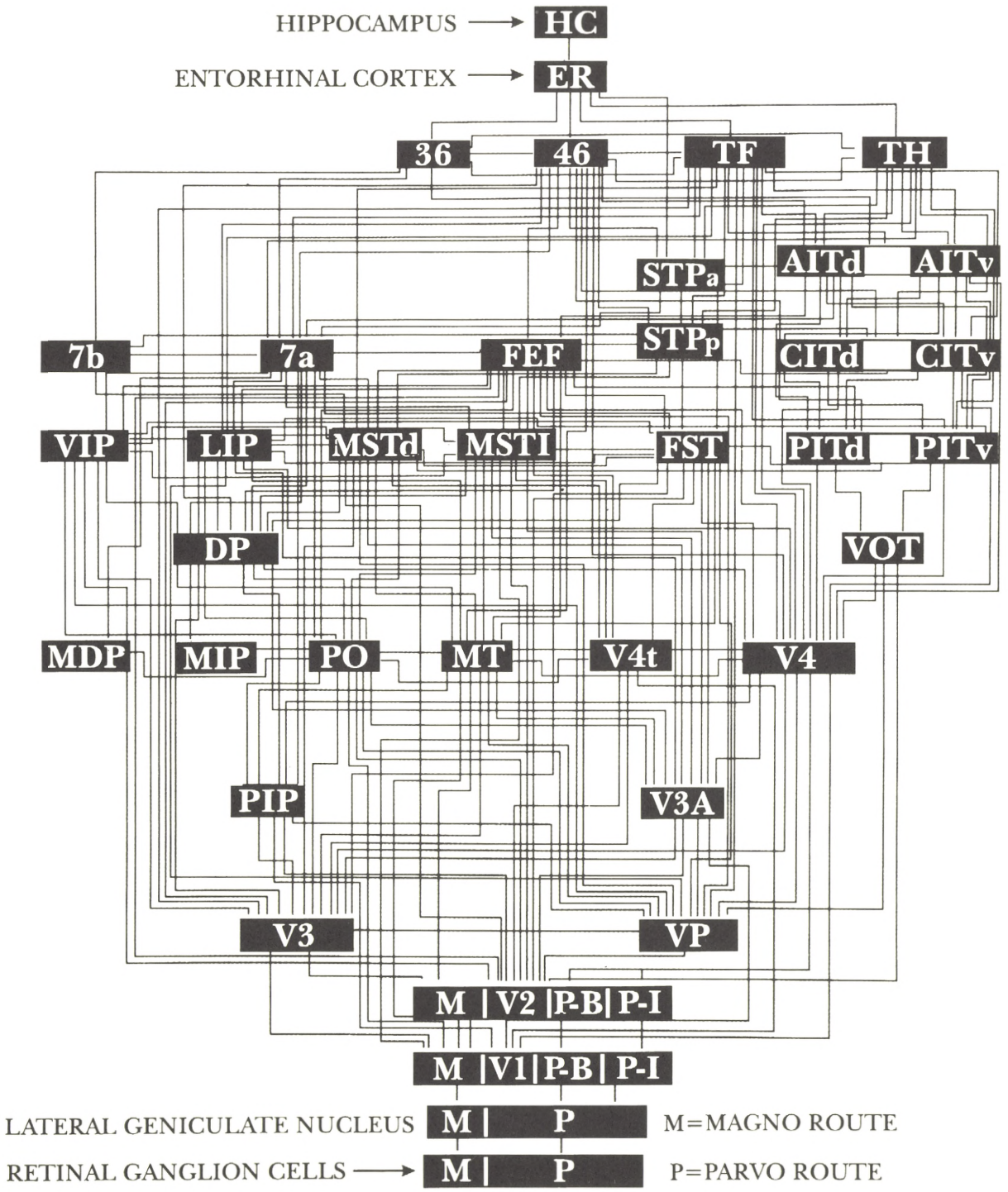


Figure 4 The diagrams reproduced on this and the following two pages indicate the layout of the various cortical areas of the macaque monkey. (Equivalent diagrams for the human are not yet available; it is expected that they would be at least as complex, and that the human cortex might comprise even more areas.) They have been adapted from the work of Felleman and Van Essen (1992), whose papers on the subject have drawn on data published by a large number of other authors. The diagrams on this page show a lateral view of the right hemisphere seen from the right (left-hand diagram) and a mid-sagittal section of the right hemisphere seen from a vertically oriented median plane, looking toward the right (right-hand diagram). The early visual areas such as V1 and V2 are seen to be located in the occipital lobe, the letters d and v designating dorsal and ventral respectively. Prominent amongst the features shown in the temporal lobe are the anterior inferotemporal area (AIT), the central inferotemporal area (CIT), and the posterior inferotemporal area (PIT). The ventral occipitotemporal area (VOT) is located at the border between those two lobes. The medial dorsal parietal area (MDP) is a major feature of the parietal lobe, while the parieto-occipital area (PO) is another borderline region. The diagram reproduced overleaf shows how one entire cerebral hemisphere appears when it is spread out on a table, with all the gyri flattened out. The retinae have been included, to show how they project to the visual cortex via the (appropriate) lateral geniculate nucleus, superior colliculus, and pulvinar. The features seen in the lateral and sagittal views are easily identifiable, as are many other cortical areas. This spread-out view gives a good impression of the relative positions of the main entry regions for the various senses (compare with figure 1), as well as the position of the motor cortex. The frontal eye field (FEF) is of particular interest to the issue of working memory, because it appears to cooperate with several distantly-positioned areas in providing the system with a sort of erasable sketch-pad or blackboard. The manner in which some of these areas are inter-connected is illustrated on the facing page overleaf. The entry point for visual signals is shown at the bottom, while the hippocampus appears at the top. The central role played by the frontal eye field is reflected in its central location in this circuit diagram. Each line represents the millions of individual axons of a known inter-area (white-matter) route, no effort having been made to indicate vergence. Many of these routes comprise both forward and reverse projections. As can be seen from this final diagram of the series, the various cortical areas can be arranged in a hierarchy, running from the early visual areas to the hippocampus.


MEDIAL PREFRONTAL
DORSAL PREFRONTAL
LATERAL PREFRONTAL
ORBITOFRONTAL
MOTOR
AUDITORY
CINGULATE
SOMATO-SENSORY
OLFACTORY
VISUAL
25
14
32
10
46
9
45
12
11
13
24
FEF
MEF
6
PALL
PRO
23
SMA
3a
4
G
1d
1g
3b
Sll
2
1
5
7b
MDP
30
PO
PSd
29
MIP
V2d
V3
PIP
VIP
LIP
7a
DP
V3A
V4t
V4d
V3
MT
FST
PITd
PITv
CITd
CITv
AITd
AITv
PIR
PAC
ENTORHINAL CORTEX
35
STPa
36
DENTATE GYRUS
CA1
CA3
HIPPOCAMPUS
SUBICULUM
TF
TH
VOT
V4v
VP
V2v
V1
OPTIC RADIATIONS
LATERAL GENICULATE NUCLEUS
OPTIC TRACT
OPTIC NERVE
PULVINAR
SUPERIOR COLLICULUS
OPTIC CHIASM
1 cm
RETINA
RETINA


HIPPOCAMPUS
HC
ENTORHINAL CORTEX
ER
36
46
TF
TH
STPa
AITd
AITv
7b
7a
FEF
STPp
CITd
CITv
VIP
LIP
MSTd
MSTI
FST
PITd
PITv
DP
VOT
MDP
MIP
PO
MT
V4t
V4
PIP
V3A
V3
VP
M | V2 | P-B | P-I
M | V1 | P-B | P-I
LATERAL GENICULATE NUCLEUS
M | P
M=MAGNO ROUTE
RETINAL GANGLION CELLS
M | P
P=PARVO ROUTE

there is increasing evidence that an important role is played by the *pulvinar*. In order to appreciate how this sub-cortical structure comes into the story, we should first consider some important studies of eye movements carried out during the past few years.

Eye movements are an important exception to the fact that attentive processes are usually hidden from us. They are necessitated by that lack of uniformity in the retina which gives priority to the centrally-located fovea, with its high acuity. The oculomotor system swivels the eyes in their sockets, so as to point the fovea in the direction of an item of potential interest. The swivelling is not smooth and continuous. It proceeds in the jumps known as saccades, which occupy a mere 50 milliseconds. It is between such saccades that the system inspects the details that can be encompassed within the foveal region. Yarbus (1967) established that these momentary fixations focus predominantly on the most salient features of an object. It is worth emphasizing, however, that attention can also be paid to an object not lying in the direction of gaze. We humans can master such surreptitious attention, concentrating on things seen out of the corner of the eye. Monkeys can be trained to do the same thing, the incentive to remain fixated on a particular spot being a suitably edible reward.

Wurtz, Goldberg and Robinson (1980, 1982) have studied the responses of neurons in various parts of the cortex, and also in the superior colliculus, in monkeys that had been trained to fixate. It is worth recalling that the *superior colliculus* controls the direction of gaze, and that it too receives input from the retina, even though the bulk of the retinal signals are passed to the cortex via the thalamus. (The situation is just the opposite in more primitive species, such as frogs and fish, the main target of retinal output being the *optic tectum*, which is the counterpart of the superior colliculus.) In addition, we should recall that the receptive field of a neuron anywhere in the visual system is simply that portion of the total visual field to which that neuron can respond. Receptive fields (of neurons in the primary visual cortex) for the foveal region are typically four degrees, whereas extra-foveal regions may give rise to receptive fields four or five times larger than this.

Wurtz and his colleagues made a serendipitous discovery. They were studying the response to a spot of light lying at an angle to the direction of gaze, by probing the activity of a particular extra-foveal neuron. Suddenly, their monkey inadvertently released its fixation and glanced at the spot, about 200 milliseconds after it had appeared on the screen. The investigators observed a surge of activity in the neuron they were monitoring, starting a mere 50 milliseconds after the spot had appeared. Subsequently establishing that there was no such response if the eye movement was made in darkness, Wurtz and his colleagues concluded that this activity guided the swivelling of the eyes, the location of the probed neuron relative to the fovea defining the direction of movement. The surge in activity was clearly the precursor to a saccade. Conversely, as the investigators later established, a response is not seen in the appropriate collicular neurons when an extra-foveal spot is merely being attended to by the monkey, without subsequent shift in gaze.

When Wurtz and his colleagues made a similar investigation of the primary visual cortex, they obtained results which differed from those seen with the superior colliculus, in virtually every respect indeed. A slight rise in activity of an extra-gaze neuron could be detected, admittedly, but this did not appear to define the direction of a saccade. It seemed that the increased activity was related not to attention but rather to the general state of awareness. The conclusion was that the primary visual cortex and the superior colliculus have widely different significances for the visual system. The response of that region of the cortex is highly

discriminative with respect to the input, but its resultant activity pattern is not directly related to any immediate reaction by the oculomotor system; the primary visual cortex appears to be oblivious of significance. The response of the superior colliculus is crude, on the other hand, but it is eminently purposeful; for this brain structure, significance is of the essence.

Wurtz and his colleagues went on to show that there are two other cortical regions which behave more like the superior colliculus than the primary visual area. These regions are known as the *frontal eye field*, which lies in the frontal cortex, and the *posterior parietal cortex*. The former had been shown by Ferrier to provoke eye movements when it is directly stimulated, while damage of the latter, in one cerebral hemisphere, is known to lead to neglect of parts of the other side of the body. A person who has sustained a lesion in the posterior parietal cortex may leave one side of the hair uncombed, and one side of the face unshaven. Mountcastle showed that cells in this cortical area display a rise in activity immediately prior to an eye movement, just as Wurtz and his colleagues reported for the superior colliculus. Bushnell, Golberg and Robinson (1981) were able to show that there is nevertheless a difference from the collicular response. The appropriate neurons in the posterior parietal cortex begin to react even when an object in their receptive field is merely being attended to. The corresponding neurons in the frontal eye field and the superior colliculus come into play only some tens of milliseconds later when a saccade is immanent.

We should pause here and briefly contemplate the broader significance of what this work was establishing. The animal can presumably exploit the differential routing of the various characteristics of sensory input. One can liken this apportioning of attributes to what happens because of the parallel distributed processing (PDP) in a perceptron (Rumelhart, Hinton and Williams 1987), with different aspects of the input being handled by different pseudo-neurons in the hidden layer, even though the analogy with that artificial neural network must perforce be a loose one (Crick 1989). And it must also be emphasized that such PDP, in the case of the cortex, does not mean that the whole sensory message can be regarded as a conglomerate of independent parts. Far from it, indeed, for the brain is so constructed that this wholeness is carefully protected by the binding mechanism, to which we will return later. Differential routing certainly does occur in the cortex, however, and one has a good example in the visual system, with the temporal region handling the *what* of visual input while the parietal region looks after the *where* and *when*. Now let me come to the point of this little digression. If the designer of the brain wished to tap off from the cortex information that would be useful for attention, which region would be the favoured choice? Bearing in mind that the items being attended to might not be stationary, the parietal region is seen to be a particularly strong candidate. Viewed in this light, the results described in the preceding paragraphs have a satisfying logic about them.

We have not yet done with the parietal region, and its significance for attention, and there are two other parts of the brain that we will also have to take a close look at before the end of this section, namely the *thalamic reticular nucleus* (*TRN*), also just as commonly known by its Latin name *nucleus reticularis thalami*, and the *brainstem reticular formation* (*BRF*). But before returning to our main thread we must clear up another piece of unfinished business. It concerns those 50 milliseconds that are consumed by the eye movement of a saccade. Had it not been for what was described earlier, we might have been inclined to look upon this as a period too brief to be worthy of further consideration. But that visually-significant interval of

18 milliseconds, as measured by Bundesen and Shibuya, suggests that we might overlook the period of movement at our peril. Indeed we would! As we are about to discover, what happens in those 50 milliseconds is a piece of biological orchestration that is of the utmost importance. It is also possessed of a compelling beauty. And the conductor of the piece is the *pulvinar*. Let us start by taking a closer look at this centrally-located component of the brain.

The pulvinar is located in the thalamus, partially sandwiched between the lateral geniculate nucleus (which serves vision, of course) and the medial geniculate nucleus (which plays an analogous role for audition). Its size has been found to be an excellent indicator of the evolutionary rank of a species; the more advanced the animal, the larger is the pulvinar with respect to the rest of the brain. Although we do not need to go into all the details, it should be noted that it is a composite structure, several parts of which have been found to comprise *retinotopic maps*. The latter term refers to the spatial arrangement of the receptive fields, which are congruent with the arrangement present in the retina. Several of the pulvinar components receive strong projections from the superior colliculus and various early visual areas of the cortex, and they dispatch reverse projections to those same regions. It is particularly noteworthy, in the present context, that the cortical areas involved in this exchange of signals include the posterior parietal area and the parieto-occipital area (both of which are on that *where-when route*), and the middle temporal area (on the *what route*). Given the posterior parietal area's involvement in spatial attention, which we have just been discussing, its inclusion in this list is intriguing.

Earlier in this long section (and we are far from finished yet), the term *filter* cropped up in connection with the pioneering work of Broadbent. It transpires that vision-related pulvinar neurons carry out a filtering which discriminates between movement of the retinal image caused by movement in the external visual field and movement of the retinal image caused by movement of the eyes themselves. (The reader may be familiar with a simple demonstration of this difference. Stationary objects observed as the head is rotated are perceived as being motionless. But if the corner of the eye is gently poked, the entire visual field appears to move.) A sizeable fraction of these neurons respond to movements in the visual field during fixation, but their activity falls sharply during a saccade. As Robinson and Petersen (1992) have noted, it is as if these neurons do not see, during eye movement, an object which they see readily during periods of fixation. Moreover, these pulvinar neurons receive projections from the superior colliculus, and Robinson and Petersen have postulated that the latter modulates the activity of the former. The *visual characteristics* of these pulvinar neurons are nevertheless dictated by what they receive from the early visual areas rather than by what comes from the colliculus.

We can now begin to glimpse the reason why the wiring of the brain has to be so complicated. The mechanism we have just been discussing involves the superior colliculus, the early visual areas and the pulvinar, with each of them having responsibility for a different aspect of the overall process. But what a complex process it is. In fact, there is evidence that the pulvinar can accomplish things that are still more subtle. When pulvinar neurons are monitored immediately after an eye movement, it is found that their activity level takes a brief amount of time to establish itself again. There is even evidence of certain pulvinar neurons being able to filter out extraneous information when there is movement in the background, during a period of fixation. This would heighten the salience of the object being attended to. Finally, experimental results are accumulating which suggest that other parts of the pulvinar

can achieve analogous filtering for auditory and tactile signals. The pulvinar appears to be the brain's conductor *par excellence*.

I am going to add my own piece of conjecture to this story, and attempt to go beyond what has already been established experimentally. We have seen that the pulvinar essentially switches vision off during the 50 milliseconds of eyeball movement, and we also know that the eyes stay stationary for about 250 milliseconds thereafter, before the next saccade. The question arises as to what happens to the visual snapshot gained during the preceding 250 milliseconds, when the eyes shift to a new target. It cannot already have disappeared from the system, because very short memory lasts at least a couple of seconds or so. I wish to suggest that the system not only retains that snapshot but that it can accurately fit it to the new snapshot that is being acquired. I believe that this could happen because the superior colliculus is able to keep track of the direction and magnitude of each eye movement. If this guess is correct, it would mean that the visual system is able to build up a composite picture, jig-saw fashion. (The mechanism bears a certain resemblance to the technique of *aperture synthesis* in radioastronomy, in which a number of well-spaced and small antennae are electronically linked up in such a way that they become equivalent to a single much larger radio telescope.) Ultimately, of course, the individual snapshots of this composite array will fade, and they will do so in the order in which they have been acquired. The visual system will thus be in a continuous race against time; it will be as if the motifs on the individual pieces of the jig-saw puzzle gradually fade away. (The use of the term *jig-saw* should not be taken to imply any significance for the concept of interlocking, in this context; one should rather think of the simplest type of jigsaw puzzle, composed of a few simple squares, which are used for very young children. In the balance of this article, I will prefer the more technical term *aperture synthesis*, sometimes more generally applying it to situations that are not limited to the visual domain.) This is no doubt the reason why the eye returns again and again to the same areas in the visual field. Wurtz and his colleagues were easily able to establish this, and they noted that features of particular interested were visited most often by the center of gaze.

An analogous mechanism could operate for the other senses. Let us briefly consider hearing, for example. Although that sense has not been treated to the same degree of detail here, we can quickly lay an adequate foundation. Just as there are no lights flashing around inside the head, so is that inner sanctum devoid of all sound. In both cases, the information has been converted to trains of action potentials. In vision, it is the position and shape of what is being perceived (not to mention colour and movement) that are of the essence, and these are analyzed through the receptive fields of the relevant cortical neurons. In audition, the important attribute is frequency content. But the system converts frequency to position, in what is known as the *tonotopic map*, and this too can be analyzed through the receptive fields of the appropriate cortical neurons. This means that the requisite machinery for *aperture synthesis*, as I am calling it, also exists in hearing. Given that the pulvinar is also implicated in the tactile sense, I will guess that the same sort of thing can even apply to touch.

It must be emphasized that this is mere speculation on my part, but the various components required for such a consolidation process have been shown to exist, so the idea is not particularly farfetched. If this picture is a reliable one, it could provide a mechanism for the bandwidth limitation in Broadbent's filter process. We recall that the three pairs of digits fed simultaneously to the two ears, in his classic experiment, were attended to as two sequential triplets; the three digits impinging upon one

ear were stored first, and the three reaching the other ear were stored thereafter. If the mechanism just described is correct, it would provide the means for storing each of the triplets. And because the span of very short memory is naturally limited, we can readily see that the number of pairs of digits fed simultaneously to the two ears could not be made arbitrarily long. The two trains of three digits could not be replaced by two trains of fourteen (say) digits – not if the subject was to have any chance of remembering them all.

We could go on to further fascinating speculation along this road. Does aperture synthesis also operate at the higher levels of cortical function, for example, and are there conditions which limit the number of pieces that can fleetingly be held in very short memory? I will be returning to such issues later, when we consider the depth of reasoning that any given brain is capable of. For the time being, however, let us return to the parietal lobe, and consider what happens when it has been injured. Posner, Walker, Friedrich and Rafal have emphasized that the cognitive act of shifting attention from one place in the visual field to another can be accomplished covertly, that is to say without movement of the eyes (Posner, Walker, Friedrich & Rafal 1984, Posner & Driver 1992). They have rationalized such shifts as comprising three distinct stages: first there is *disengagement of attention from its current focus*, then there is *transfer to the new target*, and finally there is *engagement of the new target*. From their studies of human patients, they were able to conclude that damage to the parietal region causes a defect in the disengagement stage, if the target happens to lie to the side of the head opposite to that which has the lesion. The researchers also studied patients with damage to other parts of the brain, including the frontal and temporal lobes, and also the midbrain. None of these other damage sites were found to cause difficulties with disengagement, even though it is known that they all receive projections from the parietal region.

In this respect, other projections from the posterior parietal cortex may be rather more significant. As Mountcastle (1978) has stressed, there are particularly strong sets of connections to the superior colliculus and the BRF (brainstem reticular formation). He has described evidence that these components somehow manage to gauge errors in the momentary alignment of the eyes, so that this can be reduced to zero. It was not possible to construct details of the mechanism from the observations, but there are a number of other pieces of work which, if taken together, might provide a clue. I will mention them in the order which best reveals the anatomical loop made up of the structures they invoke. Let us start by tracing the loop itself. The first link is the abovementioned path from the posterior parietal cortex to the BRF. Secondly, there is the influence that the BRF exercises on the TRN (thalamic reticular nucleus). Thirdly, there is the fact that the TRN appears to act as a gate between the thalamus and the cerebral cortex (Singer 1977). The projections from the thalamus to the cortex do not include direct lines to the posterior parietal cortex, which would close the loop. Instead, the line goes from the thalamus (more specifically, the lateral geniculate nucleus, of course) to the early visual areas. These, in turn, feed signals up the hierarchy, and it is in this way that the loop is ultimately closed, when the line reaches the posterior parietal cortex (see figure 5).

The case for a closed loop thus appears to be a good one, but what would be its significance? We can gauge this by considering each of the links in turn. Let us start with the BRF. Singer (1977) has reviewed the anatomical evidence, and he suggests that the BRF controls the activity level of the TRN, through a process of disinhibition. This mechanism appears to be capable of finer control than that of direct excita-

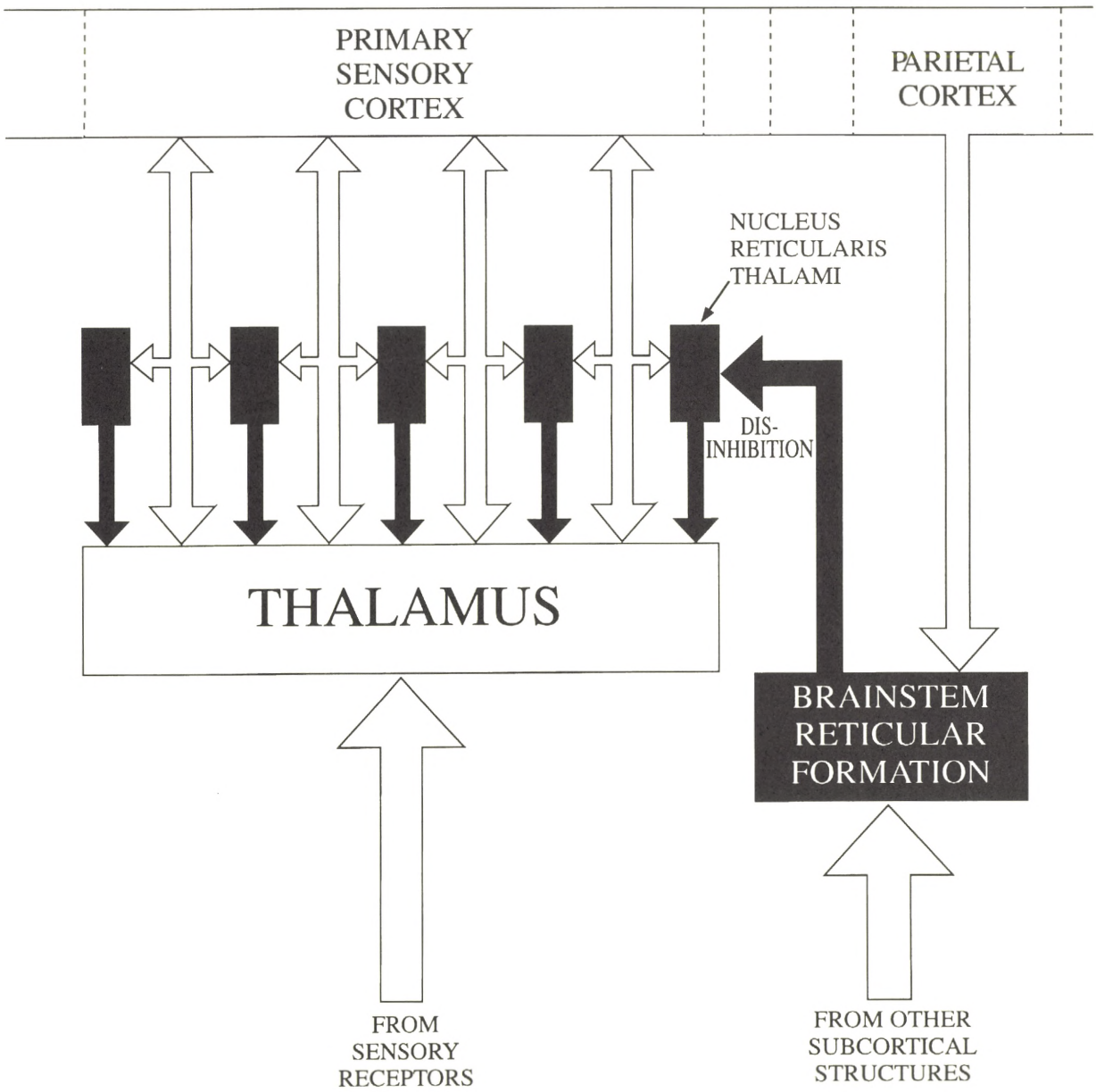


Figure 5 The excitatory pathway from a particular part of the thalamus (the lateral geniculate nucleus, in the case of vision) to the relevant part of the cerebral cortex (areas 17 and 18, in the case of vision) is modulated by the nucleus reticularis thalami (TRN). The inhibitory neurons of that structure form a layer that surrounds the thalamus, and the thalamocortical projections have to penetrate through this sheet. These projections activate the TRN neurons en route, through their axon collaterals, and the activated TRN neurons then return inhibition to the thalamus. The degree of this negative feedback is further controlled by the brainstem reticular formation (BRF), the neurons of which are also inhibitory. The BRF therefore exercises what could be called a gain control, and it does this by disinhibition. Finally, the BRF itself is controlled by, amongst other things, the parietal region of the cortex. It determines the state of arousal of the animal. In this highly schematic diagram, excitatory neurons are shown in white, while their inhibitory counterparts are coloured black.

tion. We must then ask what the TRN achieves for the system. Crick (1984) has suggested that it functions in conjunction with the thalamus to intensify a particularly active thalamic input to the cortex, thereby providing a spotlight of attention. We recall that the existence of such a spotlight was postulated by Treisman. Let us consider how its structure would make it suitable for such a task. The TRN surrounds the thalamus in much the same way as the fingers of one hand can be made to surround the clenched fist of the other. Nerve fibers, that is to say axons, stemming from the thalamus must therefore penetrate past the neurons of the surrounding nucleus, on their way to the cortex.

Now it turns out that these TRN neurons are all inhibitory (the neurotransmitter being GABA), so they oppose the activity of the thalamic neurons via feedback inhibition. Although Crick invoked an observed peculiarity of the thalamic neurons which has not stood the test of time, a spotlight mechanism can still be salvaged from the observed anatomy. The point is that the inhibitory neurons of the TRN will subject the oncoming signals to a winner-take-all process, and this will automatically provide a salience filter; it will function like a selective gate. The beauty of the system, if this scenario is correct, lies in the circular nature of the control. The BRF controls the TRN, which controls, as would a gate, the information passing from the thalamus to the cortex, some of which, no doubt in a modified form, helps to control the BRF, via the posterior parietal cortex. The key phrase here is *in a modified form*, because if it were not for that possibility of modification, the closed loop would seem to have little purpose. The modification in question will be the result of cortical processes, which is to say, ultimately, of thought. Some time will pass, surely, before all the details of this picture are fully worked out. When they have been, we will be much the wiser regarding the mechanism that facilitates a search when we know what we are looking for, and we will probably be closer to an understanding of what enables us to pick out a familiar voice through the din of a cocktail party.

Recent discoveries regarding the anatomical structure of the TRN have removed what was earlier seen as a difficulty for Crick's ideas. His theory required that the TRN be composed of well-defined topographical maps, one for each sensory modality, but evidence of such maps was not available at the time his original suggestion was published. It has now been shown that these maps do exist, however (Mitrofanis & Guillery 1993). The pathways connecting both the thalamus and the cortex to the TRN establish distinct reticular sectors that relate to functionally distinct parts of the thalamocortical pathway. (We recall that the word *reticular* refers to a netlike structure. These new results indicate that the structure stems from the distinct anatomical divisions.) The Crick view has thus been vindicated.

Relevant Limitations in Colour Vision

The idea of a limitation in the signal-handling capacity of the brain was a central theme in the treatment of attention given in the preceding section. But although a good case was made for such a restriction, direct evidence was lacking. In the case of colour vision, the evidence for a restriction is quite compelling, and it is thus appropriate to make a small digression and consider the facts. I must emphasize, however, that I am not going to go into the philosophical question of what it means to perceive colour. That issue belongs to the broader realm of consciousness, which we will be turning to later.

There are three types of wavelength-sensitive (cone) receptor in the normal human retina, and they display peak responses to what we know as the colours red (R), green (G) and

blue (B). It is not the case that these three different types of cone give rise to three separate sets of nerve fibers running from the retina to the thalamus, and thence on to the cortex. Far from it, because there is already interaction in the plane of the retina, between the various types of cone. This gives rise to signals that derive from additions and subtractions of the individual cone responses. Thus neural pathways have been found which signal R-G, G-R, (R+G)-B, and B-(R+G). The sum R+G in the latter two combinations is equivalent to the colour we know as yellow. This is probably the reason why that colour is the fourth psychological primary, even though we do not have receptors for it.

Now how does the visual system exploit the possibilities inherent in these combinations to derive the maximum efficiency from the available resources? It turns out that it uses an approach analogous to that devised around 1950 by television engineers, when they were attempting to produce the first colour receivers. These designers naturally wanted to restrict the amount of information that had to be transmitted, and there was also the requirement that owners of black-and-white sets should nevertheless be able to receive colour programmes. The simplest approach would have been to record scenes through three different filters, namely R, G and B, and arrange for them to be recombined in the colour receiver. But this was found to give a black-and-white picture that was too biased toward what was being detected by the G channel. Another drawback was the trebling of the amount of information that had to be transmitted.

It was then realized that a better alternative would be to let one carrier convey *luminance* information, since this is all the black-and-white sets had use for (Troscianko 1987). They could then simply ignore whatever else was being transmitted in the composite signals. The question then arose, however, as to how colour and luminance were to be separated in the colour receivers. The solution lay in determining the relative sizes of the contributions to luminance made by the three primaries, and the upshot was an equation that read L = 0.3R + 0.59G + 0.11B, where the three letters now stand for the actual *magnitudes* of the respective signals. The three numbers (referred to as coefficients, or weightings), reflect the contributions to luminance made by the corresponding colours. The engineers then derived the two difference equations R-L = 0.7R–0.59G–0.11B and B-L = –0.3R–0.59G+0.89B, with L representing luminance. These latter two equations are said to convey *chrominance* information, and it turns out that less bandwidth has to be devoted to its transmission, because human visual acuity is less for colour than it is for luminance. It can be demonstrated that the employment of one luminance signal and two chrominance signals leads to optimal use of transmission resources.

The fascinating thing is that the visual system apparently avails itself of the same possibilities inherent in the luminance-chrominance strategy, thereby making the most of its limited capacity for information transmission (Fukurotani 1982). This has been a protracted digression, but it has served to demonstrate that bandwidth limitation is both real and important in the visual system. It does not seem like a dangerous generalization to assume that such a limitation would apply to *all* parts of the nervous system. Nor does it seem excessive to assume that the entire system has evolved so as to make the most out of the available resources.

Bandwidth and Working Memory

Despite the clear evidence of limitations on signal-handling capacity, as described in the preceding two sections, we have not yet dealt with the neuronal nature of the bottle-neck. The impression could easily be gained that this is essentially dictated by the span of short-term memory, the determinants of that attribute be-

ing uniform throughout the system. Results that have been accumulating during recent years suggest that this view is too simple.

As has so often been the case, evidence supporting this change of paradigm has come from observations on people with brain injuries. There are, for example, patients suffering from the classic amnesic syndrome, in which short-term memory is unimpaired despite an inability to lay down new long-term memories. The situation became more complicated when Shallice and Warrington (1970) identified a second class of patients whose trouble appeared to be just the opposite; they possessed normal long-term memory but had a short-term memory span limited to one or two items. This observation was difficult to reconcile with the accepted picture of a short-term store acting as working memory, and thereby mediating long-term accumulation. Patients with this deficit are relatively rare, so Baddeley and Hitch (1974) tested normal subjects with a dual-task technique in which they were required to remember digits while performing other tasks. Investigation of a subject's ability to retain a string of digits is referred to as a *digit-span test*, and the investigators first determined the span of each subject. Subsequent observations revealed that the subjects could still reason and learn even when their full digit-span was being required of them (Baddeley & Hitch 1974). This result, and related findings, led to abandonment of the idea that a single short-term memory serves as the working memory.

It was replaced by a tripartite scheme comprising a *central executive* and two subsidiary *slave systems*. The executive was seen as including an *attentional controller*, while the two slave components were envisaged as an *articulatory or phonological loop* and a *visuospatial sketch pad*. The former of these subsidiary components was assumed to maintain speech-based information, while the latter was looked upon as being responsible for setting up and manipulating visuospatial imagery. Figure 6 was inspired by Baddeley's schematic representation of this system (Baddeley 1992). The various parts of this composite have been investigated by different, and complementary, techniques. The central executive is probed by what is known as the *psychometric approach*, which investigates the extent to which performance on working memory tasks can predict individual differences in the relevant cognitive skills. The slave systems, on the other hand, are best examined by the *neuropsychological approach* and the *dual-task approach*, which collectively demonstrate the separability of the memory systems which serve vision and audition.

A typical example of the psychometric approach is the test in which subjects must retain the last word from each of a series of sentences. For example, a person being presented with *The sailor sold the parrot*; *The vicar opened the book*; *The chicken crossed the road*, should respond with *parrot*, *book*, *road*. The *working memory span* is deemed to have been reached when the number of sentences is so large that the subject can no longer recall all the terminal words. If subjects are subsequently tested on standard reasoning tasks, like those used to assess intelligence, it is found that working memory capacity shows a strong correlation with reasoning skill (Daneman & Carpenter 1980, Kyllonen & Christal 1990). The two things are not actually equivalent, however, because reasoning shows the greater reliance on prior knowledge, whereas working memory span is more dependent on sheer speed of processing.

It has been found that the key function of working memory is coordination of resources (Baddeley 1986, Baddeley 1990). Strong endorsement of this proposition comes from observations on patients with Alzheimer's disease. When required to simultaneously perform two tasks, one visual and one verbal, these people manage very badly, despite the fact that they can handle either task individually just as

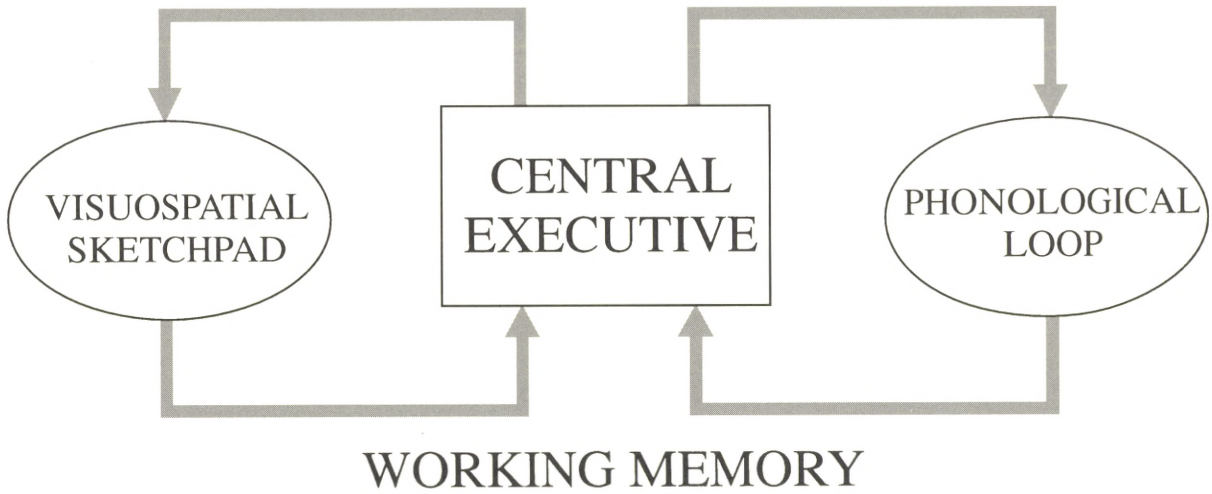


Figure 6 According to the model of working memory advocated by Baddeley and Hitch (1974), there are three components, as indicated in this highly schematic diagram. The central executive, which is an attention-controlling system, might be located in the prefrontal cortex. It marshals reverberatory processes in the two slave systems, namely the visuospatial sketchpad and the phonological loop. These might each be a series of successive cortical areas, their interactions being mediated by the forward and reverse projections that are known to be a common feature of the cortex (see figures 1 and 4).

well as their healthy counterparts. Moreover, although the patients show little deterioration in their ability to handle those individual tasks, as the disease progresses, their facility for combining tasks falls away markedly; it is clearly the ability to coordinate their mental resources which is suffering (Becker 1987, Baddeley, Logie, Bressi, Della Sala & Spinnler 1986).

Our understanding of the peripheral slave systems has progressed even more rapidly, and the dual-task approach has confirmed the separate nature of visuospatial imagery and verbal repetition. Imagery is interfered with if the subject is concurrently required to perform another visuospatial task, whereas no such clash occurs if the secondary task is a verbal one. A similar lack of interference is observed if the primary task is verbal and the intended distractor is visuospatial. Moreover, there is further separation between positional and pattern processing in the visual domain, and the corresponding neural mechanisms have been causally linked to the occipital and parietal lobes, respectively. Such separation is endorsed by neuropsychological data. There are patients who experience difficulties in recalling the shape of a spaniel's ears or the colour of a pumpkin, and yet can still easily navigate their way through a labyrinth. Conversely, there is another type of patient whose deficits are just the opposite.

The phonological loop is the easiest component of working memory to investigate, and it is consequently that part which has been studied most extensively. It appears to comprise two parts: a *phonological store*, capable of retaining acoustic information for one or two seconds, and an *articulatory control mechanism*. The latter plays a mediatory role, and it is this part of the system that permits us to register visual information in the phonological store, by subvocalization. This model of the loop explains observations such as the *acoustic similarity effect*, the *irrelevant speech effect*, the *word-length effect*, and *articulatory suppression*. Let us briefly consider these effects, in that order. We find it easi-

er to recall a list of dissimilar words, such as *pit, day, cow, pen, rig,* than a sequence of sound-alikes such as *man, cap, can, map, mad.* Baddeley attributes poorer retention of the latter group to the fact that similar items have fewer distinguishing cues, and are therefore more easily forgotten. I will suggest an alternative explanation shortly, but let us note here that similarity of *meaning* does not produce this effect; this subsystem appears not to reflect *semantic* coding.

The irrelevant speech effect refers to impairment of recall through interference of irrelevant spoken material. As we noted earlier, the pioneer in this type of study was Treisman. Here too, the semantic character of the material is not important. The word-length effect was discovered when it was found that memory span for words is inversely related to the spoken duration of the words. Finally, articulatory suppression occurs when subjects are required to include a vocalized gratuitous sound, such as the word *the,* in a word sequence that is being subvocally rehearsed. This effect is also found to disrupt the registering of visually-presented material in the phonological store.

The consensus view of the phonological store is that it serves as a backup system for comprehension of speech. The definitive tests leading to this conclusion have involved repetition of nonwords, the subjects usually being children, so that their subsequent development could be gauged. It is found that children who encounter difficulty in repeating nonwords increase their vocabularies at a much slower rate than those with a facility for recalling gibberish. Short-term phonological memory is apparently crucial to the acquisition of new words.

Returning to the system in its entirety, working memory is best regarded as a mechanism that permits performance of complex cognitive tasks through its ability to temporarily store information related to the various sensory modalities, particularly those of vision and audition. In Baddeley's view, working memory stands at the crossroads between memory, attention and perception. As such, its role in cognition could hardly be exaggerated, and any damage to this system would obviously imply serious disadvantage for the sufferer. We will later be contemplating the plight of people who appear to have such a deficit.

What bearing do these findings on working memory have on some of the things discussed earlier? For a start, they are obviously related to the bandwidth concept embodied in Broadbent's filter theory. They have simply served to fill in more of the detail regarding the various components of the system; they have shown that we must differentiate between the active executive function and the passive reverberatory mechanisms, that is to say the visuospatial sketch pad and the phonological loop. The question remains, however, as to where these various components are actually located in the brain. It seems likely that the reverberatory loops are simply the relevant regions of the cortex that are interlinked by those forward and reverse projections, which we contemplated when discussing cortical mechanics earlier. And as for the central executive, which must clearly be accessible to signals emanating from different sensory modalities, we should recall that work of Goldman-Rakic and her colleagues which was cited above.

They made a very strong case for what they were calling working memory being located in the prefrontal cortex, and we learned that this region has connections with many other parts of the brain; it is ideally connected to play a marshalling role, organizing and coordinating information from various other brain centers. (Although the analogy is probably too simple, it might be useful to think in terms of marshalling by traffic lights.) Armed with what the work of Baddeley, his colleagues, and his contemporaries has revealed, it is tempting to reappraise

that other evidence and conclude that the prefrontal cortex specifically provides only the central executive, rather than the entire working memory.

How does it go about its job of steering the traffic, at the neuronal level? We do not know. The fact is that we know very little about such traffic steering in *any* part of the cortex. But I believe there may be a hint at this mechanism in two different aspects of the work of Treisman and her colleagues: breakthrough, and the fact that it is difficult to detect an O amongst Qs, but easy to find a Q amongst Os. Breakthrough, I feel, is suggestive of inhibition that is momentarily being overcome. And we have already noted that it probably underlies the acoustic similarity effect. Moreover, we saw earlier that the results of the O-Q study indicate that the cortex reacts to the *presence* of things, rather than to their absence; the cortex works through neuronal activations, not through a lack of them.

Let us assume that the circle and the short straight line which constitute a Q, in the visual domain, can be compared with the phonemes that constitute words in the auditory domain. When the cortical region corresponding to a given phoneme is activated, it has presumably become the momentary winner in a winner-take-all process. One inhibitory factor in that type of mechanism may be supplied by *feed-forward inhibition*, and one could liken the process to one in which a pattern of lights is trying to penetrate a fog; only the brightest parts of the pattern succeed in getting through. Having been thus excited, the neurons associated with a particular phoneme will maintain their activity only for a limited period of time. Trouble for the system would surely ensue if this were not the case. What ultimately curtails the activity, however? It could simply be the leakage of charge that occurs in any electrical system. But a crisper termination of activity will be achieved if that activity is quenched by inhibition. And the natural candidate for this *second* opposing factor is *feedback inhibition*, because it will arise just where it is most needed, namely near the site which has just been activated; that is to say, at the neural correlate of the phoneme in question.

Unlike the feed-forward variety, however, feedback inhibition involves a time factor. Activation of the relevant inhibitory interneurons, and the consequent return of inhibition to the activators, takes time. That out-and-back trip traverses two sets of dendrites, as well as the related synapses. The net transit time could be at least 20 milliseconds, and complete quenching might take several times that period. We saw earlier that phonemes can follow on the heels of each other at about 50 millisecond intervals during rapid speech. This means that the terminal phoneme of the previous word might still be activated when the leading phoneme of the following word is becoming activated, and the upshot could be a false conjunction further up in the hierarchy. Perhaps this type of process underlies breakthrough in general.

This is a rather qualitative argument, admittedly, but I will be attempting to build upon it later, when we contemplate the question of how we know something. Meanwhile, let us return to the trail of those first 500 milliseconds.

Consciousness

A Consciousness Agenda

We are beginning to home in on conscious awareness. It is not that the discussion has produced a crisp definition, but the concept is becoming delineated by its beneficial consequences. We will consider more of these later. It would be a good idea to halt at this point, however, and ask whether it is known just what one would have to explain in order to produce a successful theory of consciousness. We are going to have a hard time trying to construct a hypothesis if we do not know what we are aiming at. So is there an inventory of such things, is there what could be called an official agenda? Insofar as there is not universal agreement on the issues, the answer is a regrettable *no*. But it is possible to compile a reasonably full list of items to be elucidated, even though it is difficult to say whether it would be exhaustive. One of the most active in this respect, in recent years, has been Searle (1980, 1984, 1992), whose writings on the subject are extensive. Let us contemplate the characteristics that he has identified, commenting upon them as we go.

The most important attribute of consciousness, in Searle's opinion, is *subjectivity*. We must explain how processes at the neuronal level collectively give us our sense of awareness, a sense which we privately experience. A subsidiary issue concerns the *qualia* that often accompany this awareness; we would also like a neuronal explanation of these feelings. Velmans (1991) has addressed the apparent difficulty of explaining subjective experience in what are necessarily objective terms. In what he calls a *reflexive model*, he proposes that the external phenomenal world be viewed as *part of* consciousness rather than *apart from* it. Observed events are then seen as being public only in the sense that they are shared private experiences, and scientific observations are objective only in that they are intersubjective. Velmans believes that the gap between the physical and the psychological domains can be closed by noting that observed phenomena are repeatable only in that they are sufficiently similar to be taken as tokens of the same class of event.

Another item on Searle's shopping-list is the *unity* of conscious experience, what Kant referred to as 'the transcendental unity of apperception'. In keeping with the currently favoured terminology, we have been calling this the *binding problem*. It refers to the fact that our experiences occupy a single conscious field, irrespective of whether we are sensing external events or are occupied by our thoughts. A related issue concerns the *continuity* of the stream of consciousness. Our experiences and thoughts are retained for a few seconds, as if they were being recorded single-file on tape, and this enables us to build upon them.

Searle also stresses the importance of *intentionality*, by which he means that mental states are usually related to, or directed toward, external situations and circumstances. He cedes the point that this is not always the case, however, notable exceptions being the states of anxiety, depression and elation that are merely nebulous moods. One aspect of intentionality concerns *choice*, irrespective of whether this implies the exercise of free will. Even if choice were not really free, the fact that we are able to handle it would still warrant contemplation. Searle's point is well taken. In the science of thermodynamics, there is no place for the word *purpose*, systems merely being treated on a statistical basis. In the eyes of the thermodynamicist, no arrangement of a given set of atoms is any more significant than any other arrangement. By highlighting intentionality,

Searle has identified one of the defining characteristics of the higher organism.

Then again, as Searle points out, there is *the distinction between the centre and the periphery of our conscious awareness*. He cites the example of not being conscious of the tightness of one's shoes when something else is being attended to. But we can easily switch our attention to our shoes if we so desire, at the expense of ceasing to concentrate on what occupied us earlier. There is again a corollary issue; our perceptions form natural hierarchies, some things occupying centre stage while others are relegated to the background.

Searle also identifies *familiarity* as something that requires explanation, his point referring to the fact that we will usually not have seen recognized items from precisely the angles they are currently being viewed from. Likewise, we understand words spoken in our native tongue even when they are intonated in an unfamiliar accent, and we recognize melodies even when they are played by unfamiliar combinations of instruments. As Searle puts it, we seem to have prior possession of categories for all the things we comprehend.

Finally, there are what Searle calls *boundary conditions*. He notes that conscious states are embedded in what one might call their '*situatedness*'. Although we do not constantly have such things in mind, we are nevertheless always aware of where we are, of roughly what time it is, and of the season and the (at least approximate) date. Conscious states are thus experienced in the context of a situation.

Searle's compilation is valuable in that it serves as a guide to what a theory of consciousness must account for. But it does not provide us with a definition. Indeed, Crick and Koch (1990, 1992) suggest that it might be better not to seek a definition of consciousness at this time. They make their point by noting how difficult it would be, for example, to give a neat definition of a gene. That analogy is not an ideal one, however, because a gene is not a process whereas consciousness appears to be one. Perhaps the difficulty in finding a suitable definition is related to the fact that one must distinguish between the underlying mechanism of consciousness and its mere consequences. I will later be making a guess as to the nature of that mechanism. Meanwhile, let us turn to something that has been the subject of recent controversy, namely filling in.

Filling In

When an optician inspects the back of a person's eye, through the pupil, he sees a fine network of blood vessels which feed oxygen and nutrients to the retina's neurons. The widths of these vessels are considerably larger than the diameter of a retinal receptor cell. Yet even when we look at a featureless expanse, such as a cloudless sky or a whitewashed wall, we are quite unaware of these capillaries. (A newly-ruptured blood vessel does in fact give a visible irregular patch, but this rapidly disappears again, as the receptive fields of the surrounding neurons adjust to the changed circumstances. This mechanism lends support to the argument given in this section.) The visual system somehow compensates for those parts of the field of view that they would otherwise obscure. The blind spot represents an even larger potential obstruction, but we are normally oblivious of this too. (This region is properly referred to as the *optic disc*, and it is the place at which the axons of the retinal ganglion cells leave the eye to form the optic nerve. The term *blind spot* is more appropriately applied to the part of the visual field that corresponds to this region of the retina.) Again, the system appears to make the necessary adjustments, automatically and without our intervention. This is not to say that objects whose retinal images fall upon that area will nevertheless be seen, when the other eye is closed and thus

not 'covering' for the missing receptors. (The offsets of the optic discs from the foveas, in the two eyes, are in opposite directions, so one retina has neurons in the zone occupied by the optic disc in the other retina, and vice versa.) On the contrary, such occlusion is readily demonstrated, and I imagine that the majority of readers will have observed it for themselves.

The question arises as to why special steps have to be taken to reveal the presence of the blind spot, even for monocular viewing, and why we are normally not aware of it. The answer appears to lie in the compensation mentioned above, the phenomenon being known as *filling in*. If the area around the blind spot is uniformly and regularly occupied by a series of lines, for example, the brain will fill in the missing region with similar lines, which, moreover, will be in perfect registry with the originals. Gilbert and Wiesel (1992) exposed an important component of the underlying mechanism through an experiment in which a small region of a cat's retina was destroyed with a laser beam, so as to artificially produce a second blind spot. They located the neurons in the primary visual cortex immediately adjacent to those corresponding to the damaged retinal area, and discovered that their receptive fields grew within a few hours, following production of the lesion. This expansion permitted these cortical neurons to take over the role of those that had been rendered inactive. Cortical neurons corresponding to the retinal regions adjacent to those capillary blood vessels must have similarly enlarged receptive fields.

Normally, when both eyes are open, there will be a contribution from the other retina, of course. The region of the primary visual cortex corresponding to one eye's blind spot nevertheless has neurons which receive signals from the other retina. The activity of these monocular neurons has recently been investigated. When two vertical (and collinear) lines are presented on either side of the left eye's blind spot, for example, the monocular neurons in the relevant cortical area respond vigorously, whereas their reaction is much weaker if either line is presented by itself (Gattass, Fiorani, Rosa, Pirion, Sousa, & Soares 1992). So the neural machinery that mediates filling is now amenable to direct investigation.

How does filling in occur, however, and why does it not always dominate? The answer to the first of these questions appears to lie in the feedback processes mediated by the cortex's reverse projections. These pass signals from the higher visual areas (in the case of vision, that is) back toward the primary areas. But the information on which these backwardly-travelling signals are based must have originated from the observed scene itself; it must have been derived from inwardly-travelling signals, routed via the cortical forward projections. In the vicinity of the blind spot, that information will have been gathered by the enlarged receptive fields. If the information is *sufficiently definitive*, it will have the decisive influence on what the brain interprets as being the missing object. Failing this, nothing will be seen at that position. The lines in the pattern referred to above apparently produce such sufficiency, for the result is indeed filling in.

Further exploration of these effects can be quite enlightening. A series of horizontal and closely-spaced parallel lines, each drawn with a single and similarly-positioned break, appears to be traversed by a vertical blank strip. The visual system creates this impression by producing *illusory contours* that span the gaps between the broken ends of the lines. If one arranges for a region containing the breaks to fall on the blind spot, the blank strip will appear to be completed across the occluded area. If the parallel lines are drawn sufficiently far apart, however, the brain opts for the other alternative and it is one of the broken lines that appears to be continuous. Which type of illusion one experiences is apparently determined by those

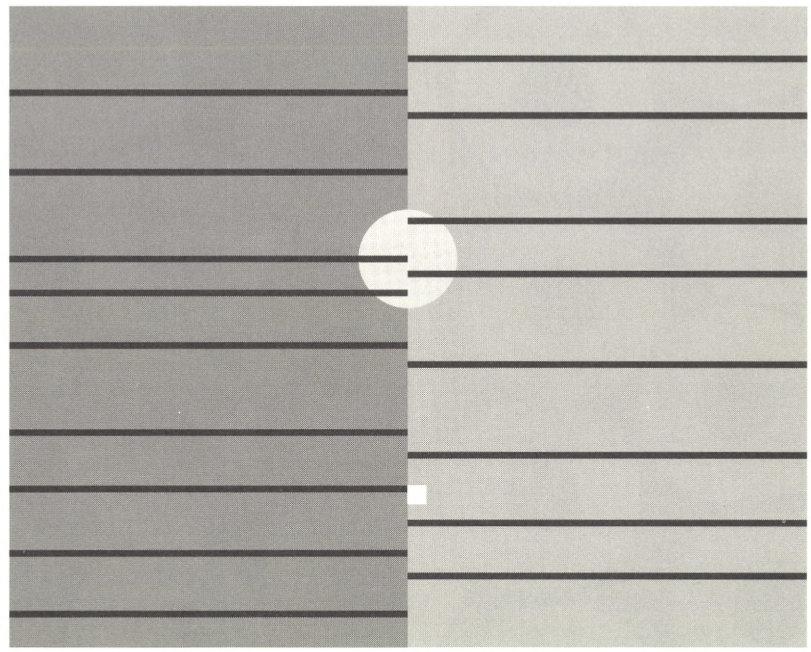


Figure 7 This diagram, which is a monochrome version of one due to Ramachandran (1992), illustrates a form of the process known as filling in. It is of particular interest here because it shows that filling in does not require participation of the blind spot. If one stares intently at the small white square for several seconds, deliberately attempting to confine one's attention to that rectangle, the two different shades of darker grey either side of the small light grey circle will be observed to fill in that feature, gradually obliterating it.

cognitive processes occurring further along in the visual system, processes which then return their verdict along the reverse projections.

The most intriguing effects of this type do not even require participation of the blind spot. Filling in can happen in a wide variety of situations provided that the viewer's attention is not focussed on the region where the compensation is to occur. As Ramachandran and Gregory (1991) (see also Ramachandran, 1992a, 1992b) have demonstrated, a strong effect of this type can be obtained with two adjacent areas that have different pastel colours or slightly different degrees of grey shading. The dividing line can be readily discerned if one looks directly at it, but it appears to evapourate when one stares intensely at a spot located somewhat away from the division. Rather unusual hues can be produced in this manner, including a quite fascinating pinkish green. Figure 7 is a monochrome version of this type of picture; despite its lack of colour, it illustrates filling in quite well.

An important aspect of this latter type of illusion is its dependence on attention. With a little practice, one can make the illusion appear and disappear at will, the trick being to control the *degree* of attention on the offset spot. Our ability to impose such control is impressive, and I will later be making a suggestion as to its origin.

Dennett (1991) has argued that filling in does not exist. His dismissal of it, although questionable, would have been less important had he not tied it so strongly to his views on the nature of consciousness. And because consciousness is one of our central concerns, we should consider Dennett's line of argument. This occupies many pages of his (entertaining) book, and it is naturally not practicable to repeat it *in extenso* here. Suffice it to say that he accounts for our normally being oblivious of our blind spots by assuming that the brain notes what lies in their vicinities and concludes that the missing parts of the images are simply 'more of the same'; that he believes the visual

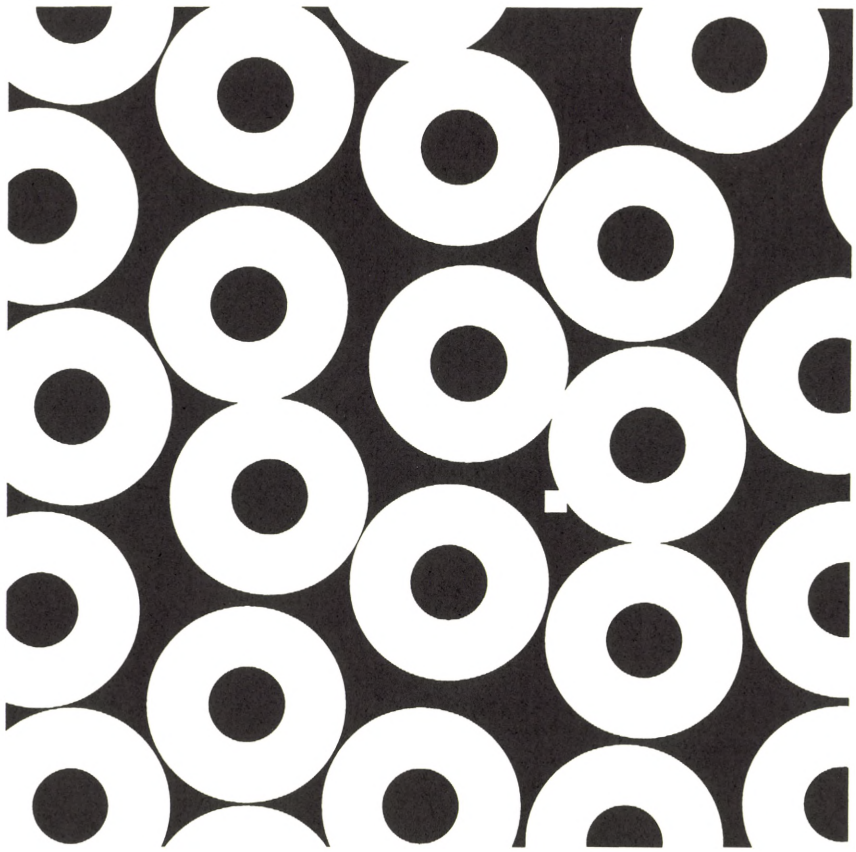


Figure 8 The more familiar example of filling in involves the blind spot. The figure on this page is due to Ramachandran (1992). If one closes one's right eye, and stares intently at the small white square, and suitably adjusts the distance between the eye and the page, the black centre of one of the white annuli will suddenly fill in, causing the resultant full circle to pop out of the background of unmodified annuli. The figure on the opposite page displays the same effect, despite its regularity. This demonstrates that the filling in phenomenon has nothing to do with the arrangement of such a figure's elements.

system simply *ignores* what actually falls on the blind spot; that he postulates there is an affinity between what happens in connection with the blind spot and blindsight; and finally that he uses these arguments to justify the view that qualia do not exist.

Dennett illustrated his first point through reference to the artist Andy Warhol's celebrated picture which consists of a repeating pattern, wall-paper fashion, of Marilyn Monroe motifs. When one inspects this diptych monocularly, one's blind spot does not cause awareness of any break in the pattern. Dennett believes that the brain simply assumes that the missing part would be 'more of the same', and that it can thus safely be ignored. This is related to his next point, to which we move immediately.

The second of Dennett's four beliefs can easily be shown to be fallacious by inspection of the two parts of figure 8, which was inspired by, and builds upon, the work of Ramachandran (1992b). The right half consists of a regularly drawn series of annuli (i.e. doughnut – or bagel – shapes), each of which would cover an area *somewhat larger than the blind spot.* If one of them is made to fall on that region, the compensatory mechanism completes the element as if it were a full circle rather than an annulus. Because the completed element is now *different* from the others, it pops out. According to Dennett's 'more Marilyns' argument, the completed element should have resembled its brethren, so it should *not* have popped out. Our conclusion must be that *the brain does not*

ignore what falls on the blind spot. The left half of the figure is similar, but this time the arrangement of the annuli is *irregular.* The fact that the same pop-out effect is nevertheless observed shows that regularity, such as that employed in Warhol's picture, has nothing to do with the issue.

We ought to pause here, however, and ask how these observations jibe with that experiment of Gilbert and Wiesel cited above. They found that cortical cells corresponding to the retinal cells surrounding the optic disc have enlarged receptive fields. Looking at the elements in figure 8, we find that each of those cortical cells should be receiving from a patch of the retinal image that comprises both black and white portions. Although this would no doubt be a crude approximation, let us assume that this patch is the sector of a circle. Because a sector tapers in toward the center of the circle, and because the radius of the inner black circle is not even half the radius of the complete element, the patch will be more white than black. We can explain the all-white appearance of the element that falls on the blind spot by concluding that each of those cortical cells is registering white as a consequence of a majority-rule process.

Before leaving the subject of visual filling in, and its dependence on attention, we ought to take a brief look at some relevant experiments with rhesus monkeys reported by Moran and Desimone (1985). These studies provide a strong hint as to the underlying mechanism. As

a preliminary, we should remind ourselves that the occurrence of vergence between interacting and hierarchically-arranged cortical areas, leads to a systematic increase in the sizes of receptive fields as one progresses deeper into the hierarchy. That is to say, neurons respond to stimuli within an increasingly large portion of the visual field. Thus whereas a typical receptive field of a neuron in area V1 measures about a degree in any direction, the average neuron in area V4 responds to stimuli within a field that is typically four times larger along each of its edges. In the inferior temporal area, which is even farther up in the hierarchy, the receptive field of the typical neuron is actually not much less than the entire visual field. Moran and Desimone's most interesting results were obtained by investigating responses in V4.

They trained a monkey to fixate on a certain spot, much as we fixate when observing figure 7. But instead of having it attend to a patch surrounding the fixation spot, they made it attend to a patch somewhat offset from that point. In a preliminary investigation, they had identified a conveniently-located neuron whose receptive field was also offset from the fixation spot, and they charted that neuron's responses to various visual stimuli in or near this field. They then measured its response under various conditions, all of which involved one favourable and one unfavourable stimulus, in or near the receptive field. Let us first consider the tests in which the favourable and unfavourable stimuli were *both within* the field. If the monkey was made to attend to a patch centered on the favourable stimulus, the response was strong. The two stimuli remained in their respective positions while the monkey was made to transfer its attention to the unfavourable stimulus. The response was markedly attenuated. So when both stimuli were within the receptive field, the response of the neuron was determined only by the characteristics of the attended stimulus. As Moran and Desimone remarked, it is *almost as if the receptive field has contracted around the attended stimulus* (my italics – see also Crick and Koch (1991)).

They then investigated the situation when the unfavourable stimulus was outside the receptive field of the neuron, the favourable stimulus still being inside. It now transpired that the response was always strong, regardless of which stimulus was attended. Moran and Desimone concluded that when attention is directed outside a receptive field, the latter appears to be unaffected. And because they observed similar responses for this arrangement, irrespective of whether attention was directed inside or outside the receptive field, the indications were that attention does not serve to enhance responses to attended stimuli.

It is perhaps a pity that Moran and Desimone did not investigate the remaining combination, namely that in which attention is directed outside the receptive field, when favourable and unfavourable stimuli are both present inside the receptive field. But based upon what had already been observed, the indications are that the responses would have been exclusively determined by the types and arrangement of the stimuli within the receptive field. This has a bearing on that filling-in experiment illustrated in figure 7. When we are not attending to an area in the visual field, the response is automatic, and determined entirely by the stimuli which happen to be present within that area. When we subject that area to attention, on the other hand, the receptive field effectively shrinks around the attended stimulus, and the features of the latter have a better chance of making a strong impression.

There remains the very important question of what actually happens when attention is directed. It is possible that the work of Georgopoulos, Schwartz and Kettner (1986) has relevance to this issue. They trained a rhesus monkey to reach out and push red buttons if these had been lit. The buttons were located directly

in front of the animal, at shoulder height. There was a central button, and it was symmetrically surrounded by eight others equally spaced on a circle of 12.5 centimeters radius. A particular trial always started with the central button, whereafter one of the others was chosen at random by the experimenters. The goal of the investigation was to discover the neuronal correlates of the resultant arm motion, and an important preliminary was the locating of 224 neurons in the arm region of the motor cortex whose activity was found to be dependent on the direction of movement. The tuning of the individual neurons with respect to direction was rather broad. Subsequent monitoring of the activity of these neurons, as a function of various movement directions, revealed a remarkably logical relationship. Their activity levels always corresponded to what would be expected if they were contributing as *vector components* to the resultant movement. (Readers not conversant with vector summation may be familiar with one of its examples: the parallelogram of forces encountered in mechanics.) If the resultant motion was aimed well away from the preferred direction of a particular neuron, its activity was very low. If the direction eliciting its maximal response almost coincided with the direction indicated by the lit button, on the other hand, it emitted a vigorous stream of action potentials. The satisfaction generated by this rational mechanism received an added boost, a couple of years later, when Lee, Rohrer and Sparks (1988) found clear evidence of similar *vectorial addition* in connection with the activity of neurons in the superior colliculus. As we have noted earlier, the role of this particular structure in the brainstem is to dictate the *saccadic eye movements* with which the visual system rapidly scans an object in the visual field.

To appreciate the significance of what Georgopoulos and his colleagues achieved in a follow-up investigation, we will first have to make a small digression and consider the pioneering work of Shepard and Metzler (1971) (see also Cooper and Shepard 1973). They asked a number of subjects to compare two mutually disoriented pictures of three-dimensional objects and judge whether they were the same. The answers came progressively slower as the offset angle was gradually increased, and from their remarkably linear plot of time delay against angle of rotation Shepard and Metzler concluded that we can mentally rotate pictures at a rate of about 60 degrees per second. The linearity of their plot was a major revelation, whereas there is probably nothing fundamental about the numerical value of its slope; we would expect to be able to mentally rotate simpler figures are a greater speed, and Shepard later showed that this is the case (for a simple *two*-dimensional figure, the speed was 400 degrees per second). This is also what was observed by Georgopoulos, Lurito, Petrides, Schwartz and Massey (1989). This time, they trained a monkey to move its arm to a location lying *perpendicular* to the illuminated button, and they again monitored the activity levels of the implicated neurons in the motor cortex. They were able to show that the various vectorial elements alter their activities in such a way as to rotate the resultant (appropriately summed) vector in the desired direction. Once again, therefore, they had revealed the existence of a remarkably logical mechanism. And given the fact that the monkey's mind was being required to rotate only a *one*-dimensional figure, namely a simple direction, we might not be surprised by their measured value of 732 degrees per second. I say *might*, here, because there is the question of whether monkey and human performances could be comparable; they indeed seem to be. Returning to our original issue, could it be that desire to attend to a certain patch of the visual field sets up a neuronal activity vector, in the superior colliculus as well as in relevant cortical areas.

Libet (1989) has suggested that filling in is a general phenomenon, and that versions of it occur in senses other than that of vision. He calls it *subjective referral of sensory representations*, and noted that the bizarre distortions of sequence observed in the tactile domain are just one of its manifestations. There is not room here to go into such matters, so suffice it to say that interesting observations along these lines have been observed in audition, by Deutsch (1992), and by Shepard(1964), and in touch, by Geldard and Sherrick (1972, 1983, 1986).

Deferment Of Automation

Filling in might have broader implications. Let us consider reactions in simple creatures, whose systems are apparently *always* fully switched on. That is to say, such creatures do not display the sleep-wake cycle of higher species. What their responses might not include, is the *ability to anticipate*. The possessor of this apparatus would not necessarily be able to project into the future, and use its predictions to figure out before hand its best tactic. This is an obviously desirable faculty for any animal, and it is interesting to contemplate its equivalent in the provinces of mathematics and engineering. In the former, it goes by the name *extrapolation*. If the trend of a variable can be sufficiently well described over a given interval, it is possible to predict its behaviour in regions outside that interval. The better the characterization, the more reliable is the prediction and the greater is the extended interval over which it can be applied. This procedure is very common; one plots a graph using a given set of points, draws a smooth line through them, and then uses the extension of that line to make one's prognosis. The forecasting of weather, although far more complicated because it involves many variables, is merely another example of this same strategy. Those who play the stock market dream of being able to make reliable extrapolations.

Likewise, in engineering practice, control often has to be of the feed-forward type, and one could say that the designer attempts to give his machine the power of extrapolation. This is embodied in the strategy known as *predictor-corrector*, and its biggest difficulty arises from non-linearity, which essentially means lack of simple proportions between stimulus and response. The brain is a machine that can be dauntingly non-linear if it is owned by a sufficiently advanced creature. This complexity notwithstanding, the numerous processes that are carried out within the device, and indeed *by* it, have to be controlled. The nervous system in its fully switched-on state appears to be forced to respond to every stimulus; it has to come up with a reaction (I am including in this repertoire of possible reactions the decision to remain immobilized), no matter what is thrown at it by way of sensory input. (The one proviso that ought to be added here is that the brain seems to be able to spare itself the bother if the input is merely random noise.) It also has one considerable advantage over the engineer: it does not have to work via mathematical equations. Its underlying structure is genetically dictated, and it can build upon this endowment by learning from experience. These twin factors determine the system's modelling capabilities.

An internal model can have considerable predictive utility. There is one which gives the frog its ability to shoot out its tongue at just the right moment to catch a passing gnat. And there is another which gives the bat its even more remarkable proficiency at intercepting a moth in flight, despite the fact that its victim is itself no mean aerial performer. Yet these impressive feats of anticipation may nevertheless be nothing more than sophisticated examples of automation. Is it possible that some modelling systems are much more impressive than the above examples? In particular, are possessors of conscious awareness capable of more complicated anticipatory feats? I believe that

this is the case. Let us recall the conceptual linking of obligatory responses to the phenomenon of filling in. Let us also remember that filling in is *not* a feature of something that is being attended to. Putting these items together, we arrive at the conclusion that the dual benefits of attention may be deferment (or possibly even suspension) of automation and exploitation of the advantages inherent in modularity.

There are still several other major issues to be contemplated before we can consider drawing conclusions as to the nature of consciousness awareness. It is tempting, however, to take stock at this point, and put what has just been discussed into a broader perspective. In so doing, we will be using our power of projection, the very faculty that we have just been deliberating. We have seen that things outside our immediate attention appear to be handed over to automation. This suggests an automatic nature for the unconscious. Then again, conscious awareness may provide the system with its only means of learning things that have novel timings and sequences. This is a more formal way of putting a point stressed by James (1910). He noted that a major task (and possibly *the* major task) of consciousness awareness is the handling of surprise. (Surprise provides the basis for a large part of humour, of course, and it stems from the conflict between the actual and the anticipated. Anticipation is probably an inevitable product of unconscious mental processes, and is thus a close relative of filling in. It seems possible that it cannot even be avoided through conscious effort.) It is the unexpected event which is invariably of the greatest interest to an organism. If that organism is primitive, unusual events may be able to impress themselves only by their alteration of the statistics of the conditioning stimuli. And because these statistics will change only gradually, the creature will not be able to respond immediately to the novel event. In more advanced species, attention, and the associated conscious awareness, may accelerate the rate at which the unforeseen occurrence can be assimilated into the repertoire of situations that must be handled. If the surprise is great, and of sufficient significance to the animal, its assimilation may even be immediate. Indeed, its very survival may be predicated on such immediate assimilation. As is well known, it is the emotions which serve to adjudicate significance.

The performance of a frog in catching a gnat was cited earlier as a good example of the type of prediction that can be mediated by a nervous system. It is an accomplishment which we ourselves would not be able to match, and we thus find it impressive. It is not difficult to expose this creature's limitations, however. If a bead is swung before it, on a thread, the frog will go for this too, subsequently spitting it out only when it has proved not to be the assumed tasty morsel. Although we are not as quick as a frog, relatively speaking, we seldom make this sort of blunder. Our superiority stems from our more advanced cognitive abilities, and although the implementation of these consumes a certain amount of time, they nevertheless give us a decisive edge.

What is the nature of this advantage? Broadly speaking, it is the ability to distinguish between stimuli that are more complex than those that can be handled by simpler animals. Such distinction requires sophisticated machinery, however, and we have already had hints that its specialty is modularity. There is in fact plenty of evidence for this, on a gross scale at least. Our command of spoken language stems from our ability to string together successions of phonemes, the variety of which is relatively limited. Similarly, the thousands of written words that comprise even the more modest personal vocabulary are constructed from a much smaller selection of individual letters. And these letters themselves have a modular construction that is based on a series of line elements. The principle is reminiscent of the

(metallic) Meccano sets that were very popular during my schooldays; given a few pieces of each basic element, and a copious supply of nuts and bolts, one could build all manner of interesting structures. Today's (plastic) equivalent is called Lego; putting the pieces together is easier in this system, but the underlying idea is the same.

A prerequisite for filling in is lack of focussed attention. The nervous system has characteristics which are acquired by the dual agencies of genetics and experience. Filling in thus appears to be automatic, and steered by what the system already comprises. This raises the question of whether the only things that can be remembered, and thus added to the system's repertoire of automatic responses, are those that have been the subject of attention. The answer must be *no*. If an animal does not possess an attention mechanism, conditioning would be the only way of altering the system. But when attention is possible, it would at least be expected to take preference over unconscious learning. Unconscious learning will still be possible, however, as subliminal television commercials demonstrate only too well. But it seems probable that things learned unconsciously will also influence our reactions in an unconscious (or subconscious) manner. The surreptitious commercial might steer us to a particular brand of goods in the supermarket without us ever being aware of the fact.

All this has a familiar ring; it is evocative of the memory mechanisms studied by Squire (1987). Let us recall his main observations. Removal of the medial temporal lobe and the hippocampus in humans leaves both short-term memory and long-term memory intact, but it severely impairs the transfer of items from the former to the latter. Although it is still not known how many different forms of memory there are, it does appear that this removal (surgically or by injury) seriously handicaps those forms that require a conscious record. These latter are called *declarative* or *explicit memory*. Forms of memory that do not require the participation of consciousness are said to be *nondeclarative* or *implicit*, and such classes are not affected by temporal lobe lesions. Finally, explicit learning is fast, while implicit learning is slow.

Adding these facts to what we have just been discussing, it seems that a synthesis is possible. Implicit memory is characterized by a slow learning speed, as well as the independence of attention (if we loosely link this to conscious awareness, that is). Explicit memory, likewise, bears a striking resemblance to what we have suggested is the simultaneous handler of surprise, exploiter of modularity, and relier on attention. Those animals which possess only the machinery for implicit memory thus begin to be seen as the victims of circumstance. Their future is determined exclusively by their past conditioning and by happenstance; if something unexpected crops up, they will be unable to find a new strategy on the spur of the moment. They are rather disadvantaged.

Trains Of Thought

James (1910) was in no doubt as to the primacy of two of our faculties, namely *discrimination* and *association*, and his analysis of brain function hinged mainly on these attributes. Indeed, he saw discrimination and association as serving all cogitation through a sort of relay race, in which the concept currently at the centre of one's conscious awareness is continually passed back and forth between these two endowments. (A slightly better analogy would be the less common type of race known as a *parlauf*, in which just two runners alternate in carrying the baton around a closed circuit.) And James noticed two quite profound things about association: items in memory tend to be recalled in sequences, but each element is not merely cued by the thing that immediately precedes it.

In other words, we tend to experience *trains of thought*, and these trends are dictated by forces that are not purely local. James concluded that things stored in memory are multiply connected, and that the context for recollection stretches over many other items in the sequence.

Regarding the actual mechanism of associative recall, James was intrigued by the fact that this sometimes takes a considerably length of time, even though one has the distinct feeling that the missing item is *on the tip of one's tongue*. His conjectures about the underlying dynamics came close to invoking the sort of mechanisms that are now so common in writings on neural networks. But the interacting units in his analysis were clearly larger than individual neurons, because they were able to store entire concepts. In the James theory, the way in which the focal point of nervous activation shifts from one item to another, as one strains to dredge up the missing information, bears more than a superficial resemblance to modern ideas about associative recall. And James saw habit amongst the neural elements as being the chief agency that dictates the course that a train of thought takes. *When two brain processes are active together or in immediate succession* he wrote, *one of them, on reoccurring tends to propagate its excitement into the other*. If one replaces the term *brain processes* by *neurons*, one obtains a rule for learning that is strikingly reminiscent of the Hebbian mechanism that is so often invoked in writings on the subject.

Although the writings of James were, perforce, couched in terms that are primitive by today's standards, he was in no doubt as to the lack of anything mystical in the mechanics of the mind. *The order of presentation of the mind's materials*, he wrote, *is due to cerebral physiology alone*. He used the term *irradiations* to describe what would now be called sets of nerve impulses, and he believed that it would be a long time before physiologists would be able to follow the route taken by these packets of activity, under a given set of mental conditions. Indeed, he expressed the opinion that this might never come to pass. But *never* is a dangerous word in science, and James would no doubt have withdrawn this pessimistic prognosis had he been able to witness first hand the sophisticated and precise types of monitoring that are routinely carried out nowadays. Just as importantly, access to the sort of detail that we now have of cortical anatomy would probably have put James on the trail of the mechanics of conscious awareness. Let us see how far one can now get with that quest.

It seems safe to say that all objects of conscious awareness can be divided into two broad categories. There is awareness of things in our environment, and there is awareness of things that are exclusively in our thoughts. Awareness of the external world would appear to implicate a greater share of the brain's resources, because sensory input is involved. A fuller picture will emerge, therefore, if we consider that particular manifestation of the phenomenon. We saw earlier that the neural activity pattern corresponding to a particular percept is generated within 50-100 milliseconds of the sensory stimulus. We also learned that as the percept gives rise to a concept, there is a backward influence on perception itself, mediated by the cortex's reverse projections. After the initial brief period, therefore, the neural activity pattern is gradually modified, increasing degrees of feedback exerting an ever stronger influence.

We should bear in mind the interesting finding by Zeki and Shipp (1988), that the reverse projections tend to be more spread out than their forward counterparts. This would promote the capturing of correlations between various features in the sensory input. A good example of this process at work is seen in Selfridge's famous two words shown in the following figure.

There is no physical difference between the sloppily drawn H and the equally careless A, but the contexts are sufficient to remove the ambiguity in each case. And those contexts are themselves the product of concept-directed perception.

The influence of concept on percept causes us to miss even the most glaring of typographical errors when reading; once the gist of a text is grasped, one starts to form mental images that can actually influence the way in which the printed words are scrutinized. This effect must be general; there is no reason why it should be limited to the written word. The mental images are provided by the associations that are already resident in memory, and their conjuring up has all the signs of being automatic. They can be regarded as a dynamic form of filling in. A proof reader wishing to avoid missing typographical errors will often resort to a tactic that defies automation, and aids awareness, namely reading the text backward. When that method is used, the incorrect spelling of a word can more readily be detected. But what is the actual mechanism of that detection? Why does it have conscious awareness as a prerequisite? I believe that it all has to do with timing, and with the duration of very short memory, and with something that I tend to think of as the device of *unfinished business*. Let us see how this works.

A Comparator Mechanism

Approaches to understanding the brain have at their disposal two broad routes of attack: the top-down and bottom-up. The problems we are addressing here have no monopoly on such division. On the contrary, diversity in the possible lines of attack crops up quite frequently in science. A current goal of physics, for example, is the complete description of nature's four fundamental forces, and a full account of the relationships between them. If recent indications are reliable, there appear to be grounds for cautious optimism that this will be accomplished. Much of the recent work in this area has focussed on the microscopic realm of subatomic particles, which is an example of the bottom-up approach *par excellence*. But the history of fundamental forces has also enjoyed its top-down successes. Newton's falling apple and Oersted's deflected compass needle heralded the greatest of these. The top-down and bottom-up approaches will have met when a subatomic explanation is found for the force of gravity, and for its relationship to the other three fundamental forces, which have already been brought under one roof.

Where will the top-down and bottom-up approaches to the brain find their meeting place? I believe that it will be in *the mechanism which enables us to know*. One could call this mechanism *the comparator*, and we obviously would like to know how it works at the neuronal level. There is one thing about the comparator that we can be pretty sure of: it cannot function in an automatic fashion, because *avoidance of automation* is being styled here as the litmus test of conscious awareness. This means that mere anatomy and physiology are not by themselves likely to produce a neural mechanism which carries out the job of comparison for us. An simple example will serve to illustrate the point. Those schooled in elementary geometry know the theorem of Pythagoras, which states that the square of the hypotenuse of a right-angled triangle equals the sum of the squares of the other two sides. Now although there are certainly neurons in the primary visual cortex which respond to lines, and thus also to the sides of triangles, it is highly unlikely that there are also neurons which carry out the various processes of squaring and summing to enable us to run a

Pythagoras check on a candidate triangle. The brain is not capable of that degree of automation; the task is too complex. Moreover, and here is a vital point, that theorem of Pythagoras is just convention; it is no more objective than is the colour red, which is just a product of our subjective minds. This might sound like an outrage against one of the most venerable products of Greek mathematics, but the point is easily seen when we realize that the theorem only holds in what is known as Euclidean space. And although the brain's overall geometry might be describable in Euclidean terms, there is no reason why the *connectivity* of its neural networks should also be Euclidean.

There is one considerable advantage associated with this lack of an automatic mechanism for squaring and summing. It relieves the cortex of the need to find places for triangle components with all manner of shapes and sizes, not to mention their squares and the sums of these. A neural depository for even this single task of simple geometry would have to be very large. Moreover, and this is a more subtle point, it relieves the brain of having to have the components of a Pythagorean computation suitably arranged in space at just the right moment for the squaring and summation processes to be carried out. To put this in other words, it spares the system the need of having a large blackboard on which to scribble all the individual contributions that go into a Pythagorean calculation. This is just as well, because who is going to read these numbers in any case? And yet, I *have* written about the brain's blackboard, in connection with the work of Goldman-Rakic, and also that of Baddeley and Hitch. We saw that it is even possible to locate one component of that blackboard (the central executive) in the frontal lobe. In what way is that blackboard more acceptable than the one we have just dismissed?

We have reached a most auspicious point in this story, for if it is not a question of arrangements in (cortical) *space*, what then? I believe that the alternative is different kinds of arrangements, and in a different dimension: *time*. Furthermore, I believe that this distinction carries over to that other great schism that we have been identifying, namely the one that separates unconscious wakefulness from conscious wakefulness. Unconscious processes are parallel and automatic, conscious processes are sequential and negotiable.

This is all very well, but we still have to ask how the comparator actually works, and why it needs the time dimension. I believe that the various bits and pieces of the mechanism are already to hand in what has been described earlier. Let us take a whirlwind survey of them, and attempt to synthesize an overall picture: the brain in a nutshell, as it could be called.

Information emanating from the environment makes impressions on the various types of receptor cell, and thereby causes signals to be dispatched from the different senses toward the deeper recesses of the nervous system. Some of these signals are not passed to the brain, travelling instead to neurons which are able, unaided, to elicit responses. These responses occur unconsciously; they are automatic. The large majority of the signals are transmitted to the cerebral cortex, however, many of them passing through the thalamus *en route*. In the wakeful state, the signals from the various sensory modalities enter the cortex at different places, and they do so in versions that are modified by the vergence (di- and con-) that always occurs when signals are projected from one group of neurons to another. Each modality is served by a number of cortical areas, vision claiming the majority share. The signals, continually modified as they travel, progress via the forward projections through the various areas as through a hierarchy, and they do so in a parallel distributed manner, ultimately to arrive at the appropriate association areas of the cortex. Further modified by the associations

that lie collectively stored in those areas, they provoke signals that continue onward toward the early areas of the other sensory modalities, this later part of their journey being mediated by the reverse projections in those target senses. Signals in one sensory modality are thereby able to provoke signals in other modalities, as when vision conjures up auditory associations, and vice-versa.

At appropriate positions in the array of cortical areas, signals are tapped off for projection either to subcortical regions of the brain or to the motor cortex. The subcortical targets are also the sources of signals sent *to* the brain via separate projections, and this arrangement establishes loops. Some of these loops are special in that they permit control over the sensory input, a particularly noteworthy example being the thalamus, which is therefore seen to be much more than a mere relay station between sensory input and the cortex. Indeed, the thalamus is under even further control in that there is a more circuitous route that runs from the cortex via the brainstem reticular formation (also known as the reticular activating system) to the thalamus, and then back to the cortex, through the gating control of the thalamic reticular nucleus. This route is special in that it can even control the degree of arousal of the animal.

The importance of precise timing to the overall system is apparent in the structure and functioning of several of these loops. The one that involves the cerebellum controls the timing of movements. Finer control is served by the basal ganglia which, unlike the cerebellum, do not send signals directly to the spinal cord. Timing is also of the essence in the hippocampus, which, through its connections with many cortical areas, imbues the animal with a sense of its own personal space, but which also mediates remembering that is contingent on conscious awareness. Because all events relevant to the animal are related to its personal space, these two things are intimately connected, and the role of timing may be akin to that familiar in radar. (This analogy does not refer to the emission and reception of radiation, of course. It is the mechanism of precise timing which is important.)

The various signalling processes consume a certain amount of time in reaching their destinations, primarily because of the slowness of signals in dendrites, and modification of the signals during early sensory processing thus occurs unconsciously. The signals also have a certain duration or lifetime in the system, this *reverberation* appearing to last several seconds; it can be regarded as a very brief form of memory, though it is not clear whether it involves synaptic plasticity. (If it does involve synaptic plasticity, the important receptors may well be the NMDA-type glutamate variety.) Conscious awareness of a sufficiently strong stimulus develops after about half a second, which is the time required for neuronal adequacy, as measured by Libet and his colleagues (Libet, Wright, Feinstein & Pearl 1979).

Although we still have to decide what conscious awareness actually is, let us consider a consequence of these temporal delays and reverberations in the system. Action potentials arriving at a given cortical area will produce two main responses. There is the excitatory response, by which activation is passed on to subsequent areas in the cortical hierarchy, and there is an inhibitory response which causes the activity in the given area to diminish, and ultimately to die out altogether. This inhibition will in general arise through two familiar mechanisms, namely the feedback and the feed-forward varieties. We have already seen that the latter type (activated sufficiently early by a prior processing stage) can set up a cloud of inhibition which will then exercise a winner-take-all constraint on the arriving signals. Those signals that succeed in getting through this cloud, by dint of their strength and coherence, will

proceed up the hierarchy and ultimately give rise to return signals, which travel via the reverse projections. They will also leave behind them a residual pattern of inhibition which arises from the feedback mechanism.

The information carried in the reverse direction will be the product of automation. Indeed, it will be a generalized form of filling in. If the original input is fully recognized by the system, the activity pattern set up in the original area by these return signals will not differ from that set up by the original input. If we assume that the pattern of feedback inhibition faithfully reflects the originally injected pattern of activity, and that seems like a reasonable assumption, this means that the return pattern will precisely match the inhibitory cloud, and the result will be instant extinction. One could say that there will be no *unfinished business* to attend to. If, on the other hand, the returning signal does not match the residual inhibitory pattern, there will be some remnant activity. This will be concentrated around the mismatch between the input and return patterns, and it would be available to provoke the appropriate responses by the system. It seems likely that one aspect of this response would involve continued attention, so as to give the system a better chance to examine the origin of the discord. Conversely, the lack of mismatch would presumably be registered by the brain's reward system. I believe that this is the way in which the comparator functions, and its ability to detect novelty is clear.

The winner-take-all and unfinished business mechanisms will clearly require different components of neuronal circuitry to be present. The former will need to avail itself of inhibitory neurons which receive input from relatively distant regions of the system. It is encouraging to note that there is evidence of GABA (inhibitory) neurons in the cortex which receive direct synaptic contacts from thalamocortical axon terminals (Freund, Martin, Soltesz, Somogyi & Whitteride 1989, Hendry & Jones 1989). The unfinished business mechanism, on the other hand, requires a strictly localized source of feedback inhibition. Failing this, the areas excited by the input will be insufficiently defined; the pattern will be too fuzzy to be useful. It is thus encouraging to note that there is also evidence that the axons of most cortical GABA cells termiinate relatively close to the parent cell bodies (Jones 1993).

It would be easy to underestimate what possession of this comparator mechanism would do for an animal. As an example of its usefulness, let us go back to that theorem of Pythagoras. As I emphasized earlier, the procedures required to implement its logic are not wired into the brain's neural networks, waiting merely for the lengths of the triangle's sides to be fed into the system. What *is* in the system? Let me reminisce for a moment on the time when I learned the theorem in school. I sat there in the class-room, surrounded by my fellow pupils, and the teacher informed us that we were going to be shown a piece of mathematical elegance. She started by *reciting* the proposition, pretty much as I stated it earlier. Then she took the 3-4-5 triangle as a specific example, and demonstrated the correctness of the proposition by adding 3 times 3 to 4 times 4, and noting that the sum equalled 5 times 5. My *whole body* registered the equality of the two 25s, arrived at by their separate routes. Numerous loops within my nervous system contributed to the feeling of well-being at the successful outcome of the teacher's little calculation. My muscles relaxed; I felt at one with the entire atmosphere of the occasion; no doubt, my various glands were registering the harmony of the situation. The teacher then used the 5-12-13 triangle as an encore, and my body became even more relaxed. *I was being told that there was no unfinished business, and my synapses were adjusted accordingly.* And what did I take with me from the class-room on that auspicious day? It

was the memory that there was someone called Pythagoras, and that his proposition had something to do with right-angled triangles and the squares of their sides. A few days later, when the teacher touched on the subject again, I actually learned to recite the proposition, word perfect. It has lain in my memory ever since, and its insertion into that repository was mediated by my brain's reward system, which was steered by that lack of unfinished business.

In view of what has been said about the way the system could function, with the comparator checking reversely projected signals against those of the input, how should memories be characterized? It seems logical to conclude that they are *stored associations between stimuli and consequences*. Moreover, from the work on filling in, it seems that the comparator mechanism is functioning most effectively when attention is present. When attention is absent, the ability of the system to discern unfinished business is apparently diminished, and there is a tendency toward automation. The idea of stored associations is in tune with a more limited suggestion by Llinás and Pellionisz (1979) concerning the coupling of sensory input to muscular output. They were able to demonstrate that such coupling can be rationalized in terms of the mathematical concept known as a tensor, and they successfully applied their theory to the specific coupling of the vestibular sense to the movements of neck muscles, in the cat.

Another idea somewhat related to the one I am advocating is seen in the neural network strategy known as the *Boltzmann machine*, which was first described by Ackley, Hinton and Sejnowski (1985). The gist of their approach is that the network is able to explore the vast number of possible couplings between input and output, in a random manner, small adjustments being made to the tuning of the system (which is to say the synaptic strengths) every time a better fit is found. In practice, to stop the system getting bogged down, a trade off was permitted such that a temporary worsening of the input-output match was tolerated in order to ultimately reach a more satisfactory solution. Finally, mention should be made of the work of Zipser and Andersen (1988), in which they asked where the equivalent of the output units of a perceptron could exist in the brain. After considering several possibilities, they concluded that the final spatial output might exist only in the behaviour of the animal; it may be found only in the pointing of the eye or finger accurately to a location in space.

Mechanics of Consciousness

The most difficult problem facing anyone set on explaining brain function is to expunge the idea of what Dennett (1991) has called *the Cartesian Theatre*. One laughs, these days, at the concept of the homunculus: that little being in the head who was supposed to play the triple role of observer, arbiter and activator. But the apparent independence of the imagination still fools us into believing that mental events are somehow played out on an internal stage. This idea was particularly prominent in the writings of René Descartes, and it was for that reason that Dennett facetiously named the edifice after him, skillfully showing it to be a theatre of the absurd.

And yet the idea stubbornly persists, by implication at least, in much of the writing on the subject. It is still common to read of attempts at explaining how the line-detecting neurons discovered by Hubel and Wiesel could feed into neurons that would be able to detect more composite aspects of a visual scene. That path leads to the grandmother cell, amongst other things, and to the question of what is to observe it in action. (A better indication of the way in which the visual system handles information is provided by the caricature, which enables us to recognize the face of a familiar person, even though no facial features are drawn

in their proper physical locations. Indeed, the more grotesque the artist's distortion of the facts, the easier it is to identify the person being represented. Recognition clearly depends on relative rather than absolute position, and this must be gauged deeper in the visual hierarchy. Another important factor is cognitive compression, in which different inputs produce the same effects deeper in the hierarchy. This may play a role in the phenomenon of *deja vue*, incidents apparently experienced before merely resembling a previous episode.) In this section, in which my main task is to essay an explanation of conscious awareness, I am going to suggest that there *is* a theatre that is relevant to brain function, and that Descartes merely made an error regarding its location. It is, I would like to suggest, a theatre that has never closed.

Before saying where this arena actually is, let me cite some other work which cannot be reviewed here in detail. Firstly, there is the fascinating analysis in which Oatley and Johnson-Laird (1987) developed their cognitive theory of the emotions, showing that all our feelings can be pigeon-holed into a large but not unlimited set of categories. Secondly, there are a number of points regarding the hippocampus. There was, for example, the idea that this part of the brain provides the nervous system with a cognitive map, thereby giving the animal a sense of its personal space. It might be able to do this because of its sophisticated way of handling the timing of action potentials, perhaps in a manner roughly analogous to the time-processing techniques that underlie radar. There was also the impressive evidence that the hippocampus mediates remembering of events specifically related to conscious awareness (Squire 1987). Above all, we must note that the hippocampus is a prominent member of the limbic system, which is known to be involved in the emotions.

We are far from done with this synthesis, but let us note in passing how spatial our metaphors become when we speak of mental processes. As the reader may have noticed, I have a weakness for the word *approach*, when referring to theories, and one often speaks of *advancing* a hypothesis and *retreating* from a belief, or *position*. We talk of *grasping* the point, of *seizing* the initiative, of *holding* an opinion, and of *dropping* a hint. One has an idea at the *back* of one's mind, one considers the evidence *in front* of one, one contemplates it from all *sides*, and comes to the conclusion that it is either *above* one's understanding or *beneath* contempt. And when we are sufficiently aroused, we start to gesticulate, using bodily positions and movements that are the physical counterparts of these metaphors. We literally *throw up* our hands, or get *hopping* mad. If Johnson-Laird and Oatley ever turn their attention to our spatial metaphors, they may discover patterns that are just as systematic as those that they discovered in our emotions. And it would not be surprising if the two systems were actually connected.

I believe that there is something quite fundamental about all this, and that it is directly related to conscious awareness. It indicates, I feel, that mental events are indeed played out in what could be called a theatre, namely *the theatre of our personal space*. I believe, moreover, that this gives us a strong hint as to what conscious awareness actually is. It is simply *a mechanism which enables us to relate to our personal space*. That might sound like a rather modest attribute, until one considers how complex an environment our personal space really is. What is more, and unlike *Aplysia*, say, we have a personal space which extends well beyond our bodily surfaces. With the most highly developed of our senses, those that claim the lion's share of the cortex, we can see and hear things at considerable distances. Indeed, our personal space includes places that are not always in our immediate vicinity, such as our places of work and

leisure, the surgery at our local practitioner, the supermarket and the pharmacy, the homes of our friends and relations and, if we are religious, our churches. Even more importantly, our personal space includes people: our spouse, our children, our parents and in-laws, our friends and our colleagues. It is not surprising that all these places and people are the greatest source of stress when they change. Bereavement, divorce, change of house, change of job: these are the things the doctor looks for when a patient complains of tension.

Fast (1971) has written about subtle aspects of personal space in the workplace. An administrator will sometimes intimidate an employee by invading the latter's domain, the violation usually being no more intrusive than the placing of a hand within the victim's portion of a temporarily shared piece of furniture, such as a table or a couch.

But the aspects of personal space of greatest interest to the present discussion concern the body's more immediate environment. And here we see parts of the nervous system which go about their tasks in a quite unobtrusive manner. One of these is constantly monitoring bodily posture with respect to the surroundings. It ensures that we do not spend so much time in a given position that blood flow would be impaired. One part of this system has what could be called shock detectors in the ball of the heel, and it increases blood flow in proportion to the cadence of walking. It can be tricked by a device which administers the occasional light tap to the heel, and such gadgets are used to ensure adequate circulation in bedridden hospital patients. In another vein, I am often intrigued by my nervous system's ability to keep a check on the amount of coffee remaining in my cup. Without giving the matter any thought whatsoever, and without ever looking into the cup, I always seem to know whether or not it is empty. A related mechanism may underlie a squirrel's relationship to its hidden nuts. Finally, it might be useful to consider some of the spatial phobias, including claustrophobia and agoraphobia, as impairments of the neural machinery associated with the sense of personal space. (Claustrophobia and agoraphobia are opposites, the former being dread of enclosed spaces while victims of the latter abhor open spaces.)

This brings us back to the question of mechanism. As we have seen, the hippocampus plays the definitive role, both in the precise timing on which our personal radar is based and in the temporary storage of memories that require the participation of conscious awareness. (It cannot be emphasized too strongly, however, that this does not imply that the hippocampus participates in the *mediation* of consciousness. Far from it, because some of Squire's memory-impaired patients nevertheless conversed with him; despite their damaged hippocampuses, they were still capable of consciousness.) The importance of timing suggests that coherence in the brain's neural activity is significant. The word coherence is not used merely in its colloquial sense, however; in this case, it is meant to imply synchrony, or near synchrony, of the activities of considerable numbers of individual neurons. Both Milner (1974) and von der Malsburg (1981) have emphasized that neurons are detectors of temporal correlations, since whether or not such a cell emits a nerve impulse is determined by the instantaneous sum of the inputs it receives. Similarly, Barlow (1985) has written of neurons functioning as coincidence detectors.

Coincidence is a word that must be used with caution, however, for it requires specification of degree. Pulses spread over several milliseconds will not be coincident on the scale of microseconds, for example. This might seem like splitting hairs, but the point becomes a fundamental one when we compare learning in the neural networks of *Aplysia* and learning in our own cortico-hippocampal system. Aplysia's net-

works can capture correlations if the relevant coincidences are within half a second or so, but the synaptic modifications occur gradually, and require several repetitions of the paired stimuli. As we have seen, our own systems are capable of a rather higher degree of discrimination, and we can learn things we are exposed to just once, if they are sufficiently important. But is this difference sufficient justification for denying *Aplysia* the possibility of conscious awareness? As Nagel (1974) would remind us, there is no way of finding out directly whether or not that sea snail also possesses that attribute. We can only compare the two sets of neural circuitry and try to draw the correct inferences. Let us now attempt to do so, and start by asking whether *Aplysia* ever imagines. Lying quiescent and unprovoked by any sensory stimuli, does it ever simply think things over and change its attitude toward the outside world? Does *Alysia* possess the neural machinery that would permit it to do this?

We have strong indications as to what imagination involves, because we noted earlier that it is suspiciously accurate. We came to the conclusion that it uses the same mechanisms and even the same neural circuitry as perception. Imagination may thus be regarded as being *simulation* of the outside world, and we have seen that our version of it relies on the presence of those reverse projections in the cortex. Now in one important respect, imagination appears to be on an equal footing with normal perception. Evidence for this comes from observations similar to those made by Treisman on breakthrough. We recall her demonstration that competing trains of perception can mutually interfere if they become too similar. I once witnessed a good example of this, in a restaurant that had a resident pianist. I noticed that he could carry on a conversation without impairing his performance in any way. My curiosity aroused, I wondered if he could even do this when the conversation involved music. I went over to the piano and asked him if he knew a particular tune. Rather mischievously, I started to hum this other melody before he could react. Sure enough, it so put him off that he had to stop playing, much to the irritation of the other guests.

I apologized, of course, and we later enjoyed a discussion, during one of his rest breaks. It was particularly interesting to learn that he was not even capable of playing when merely *thinking* of a rival tune. Just imagining other melodies was sufficient to cause breakthrough. This shows, surely, that imagination and perception are not basically different from our neural networks' point of view. The competitive mechanisms that act on perceptual inputs work just as effectively on internally-generated percepts. We might close our eyes and ears to improve concentration on our thoughts, but thoughts themselves can interfere just as effectively with the perception of our environment. We have all laughed at stories of the absent-minded professor; they are not exaggerations. There appears to be complete reciprocity between thought and perception. Experimental verification of this equivalence was obtained by Segal and Fusella (1970), who found that a conjured-up image acts as an internal signal which can become confused with the external signal.

I believe that the difficulty we encounter when trying to define consciousness arises from our inability to differentiate between the rich variety of our conscious experiences and what might be a rather humble underlying mechanism. As I mentioned earlier, Crick and Koch have suggested that we temporarily shelve the issue of definition, and they cited the difficulty of saying just what a gene is, for example. By way of coincidence, Crick (1988) has written of the choice he once faced between researching into consciousness and researching into the borderline between inanimate molecules and the chemistry of life. As is well known, he opted for the latter, and discov-

ered (together with Watson) that the crux of genetic inheritance lies not in the rich variety of organisms, but rather in the humble fact that a coplanar hydrogen-bonded arrangement of adenine and thymine has exactly the same overall length as a similar combination of guanine and cytosine. These equally-long base pairs can thus be equivalently positioned as the rungs of the twisted ladder polymer that we now know as the famous double helix, and genetic variety stems from the variety of possible arrangements of those bases.

What is the gist of consciousness? What is the counterpart of those humble base pairs? I believe that it will turn out to be equally humble, and equally mechanistic. And I believe that the key factor is *time*. If the organism is to have the ability of responding to the temporal texture of its environment, on the time scale inherent in that texture, it will have to be able to retain a temporary record that spans a sufficient amount of that texture. And it will need cognitive mechanisms which extract relevant information from that texture, in the time available. Only then will a response be possible which exploits the choice implicit in the existence of that texture. Failing this, the information in that texture will be lost, and the resulting synaptic changes (if any) will merely reflect the statistics of the texture.

After such a build-up, I must admit that this idea sounds rather tame. But exploration of its implications will reveal that it could indeed give rise to those richly varied conscious experiences. Let us begin with that word *texture*. It conveys the concept of unevenness, without which there would be little information content. In the spatial domain, there is more information in a series of lines than there is in a blank page, and the line-sensitive neurons in our visual systems contribute to the extraction of that information. One can generalize the idea of space, and say that different combinations of molecules represent different coordinates in olfactory and gustatory space, and for those cases too, we have neural networks capable of extracting the relevant information. But the information of greatest potential interest to an organism concerns changes in those generalized coordinates. That series of lines will ultimately lose its interest for us, once its message has been assimilated, and smell and taste will cease to serve us if our nostrils and tongues are permanently subjected to the same molecular species. For our nervous systems, it is change which is of the essence, and because we are also seeking to explain that facet of consciousness which only involves imagination, we must extend the concept of texture to include graininess in the actual patterns of neuronal excitation.

Change will be wasted on us, however, if the system is not able to draw conclusions from it in the time available. What we need, therefore, is what could be called *a mechanics of consciousness*. We earlier saw that experimental information on transit times for signals in the cortex permitted us to think in terms of the underlying mechanics. We must now try to extend those ideas to consciousness itself. And I am going to start with a piece of conjecture regarding the manner in which the brain deals with time. My guess is that information processing cannot be carried out by the brain in a truly continuous manner. I believe that the system has to chop time up into quanta, the duration of these being governed by the requirements of that *unfinished business* mechanism. We recall that it involved feedback inhibition. Traversal of the out-and-back route involved in such inhibition consumes a certain amount of time, during which the neurons involved in that discriminatory process are unable to handle any new signals.

My guess is that the system consequently breaks time up into packets compatible with the natural period of the feedback mechanism. Given that the out-and-back route would in-

volve transit along at least two dendrites, these quanta would lie around 20 milliseconds, which corresponds to a frequency of 50 cycles per second (i.e. 50 Hertz). Interestingly, this is about the frequency of those gamma-band oscillations observed by Gray and Singer (1989) and their colleagues, and also by Eckhorn and his colleagues (Eckhorn, Bauer, Jordan, Brosch, Kruse, Munk & Reitboeck 1988). It should be emphasized that 20 milliseconds is a very brief period compared with the 500 milliseconds required for neuronal adequacy, which is to say conscious awareness (Libet, Wright, Feinstein & Pearl 1979), so we will nevertheless perceive time as being continuous.

If these ideas are correct, we can attempt a definition of that underlying mechanism. It would seem that ***conscious awareness is possible if temporal changes in sensory input (or in internal patterns of signals) can be detected while the signals resulting from that input are still reverberating in the system, in a fully-functional form.*** And we could go on to note that ***such detection will permit the system to capture correlations between cause and effect so rapidly that their significance for the organism can immediately be appreciated, and also immediately used to modify the organism's repertoire of responses.*** By this definition, detection of correlations can be exploited sequentially. It is this, I feel, which provides the basis for the continuity we invoke when we refer to the stream of consciousness. The reactions implied in the definition may be internal, external, or a combination of both. And in this respect, the equal standing of the imagination and perception, in neural terms, is of paramount importance. One consequence of this is that it enables the system to compare sensory input with previous experience, and thus to gauge novelty *immediately*. We do well to bear in mind the importance that James (1910) attached to both the continuity of consciousness and the element of surprise.

A key aspect of this definition is captured in the phrase *in a fully-functional form*. If the neural signals caused by the sensory input were already well on the way to extinction by the time the temporal changes were detected, no advantage would accrue for the organism. The sensory input might elicit an immediate response, but in terms of modifying behaviour it would merely contribute to the statistics of the environment's influence on the organism. If the signals were still fully potent, on the hand, they would be available as agents of the response to the detected correlations between cause and effect. And because this response would include the dispatch of signals to various relevant parts of the organism, followed by the return of feedback signals, there would be the wherewithal for the associated feelings, that is to say for the qualia. The preservation of the provoked signals in a fully-functional form requires a special non-dissipative form of reverberation, and this is presumably achieved by the system of forward and reverse projections.

The abilities provided by the definition make considerable demands on candidate neural networks. There must be the possibility of internal simulation, that is to say of imagination, so a system of reverse projections is presumably a prerequisite. Then again, there must be the possibility of handling simultaneous streams of information, and of keeping those streams separated. And possession of these pieces of circuitry would be futile if the corresponding signals were not in the system long enough to permit comparisons, by that *unfinished business* mechanism. The upshot of these various requirements is that there are fairly severe constraints on the underlying time constants. Conscious awareness is predicated on there being an adequate very short term memory – the neuronal reverberation time – as has indeed been stressed by Crick (1988). And we have stressed that the very short memory function must be one of high fidelity. There is also the requirement that the transit times of the

signals in the separate processing streams have the right order of magnitude.

If these ideas about the mechanics of consciousness are on the right track, they raise the obvious question about where the underlying comparator process is implemented. Where does the unfinished business mechanism operate? Does it take place in the cortex, or in a sub-cortical structure, or in a combination of both those locations? The short answer is that we simply do not know. But there is experimental evidence that appears to single out a promising candidate region of the brain. We have already considered it in some detail. It is the thalamo-cortical region. Let us evaluate its merits.

As is now well known, there is much evidence of gamma-band oscillations in the cortex. They were originally detected in the olfactory cortex, and they were later discovered in the visual areas and in other cortical regions. It has subsequently transpired that gamma-band oscillations are also observable in the neuronal structures which link the thalamus to the cortex (Steriade, Jones & Llinás 1990). As we have seen, an important structure of the thalamo-cortical region is the nucleus reticularis thalami (TRN), which acts as a sort of gate between the thalamus and the cortex. Oscillatory behaviour has also been detected in the TRN. Finally, we recall that the TRN is under the control of the brainstem reticular formation (BRF), which appears to exert its influence by the process of disinhibition; the TRN inhibits the thalamo-cortical route, and the BRF serves as a sort of gain control by varying the inhibitory strength of the TRN. But suppose that the BRF actually varies the *frequency* of the TRN's bursts of inhibition. This could effectively tune the thalamo-cortical route's frequency so as to match that of the unfinished business mechanism in the relevant areas of the cortex.

It will be worth our while to linger a little longer on these questions of oscillations and time constants. We can, for example, compare the approximate period of 25 milliseconds that I have been estimating for the unfinished business mechanism with the approximately 18 milliseconds that Bundesen and Shibuya found was the minimum time required for a symbol to make an impression on the visual system. Their work was considered earlier, and it may be related to the texture-discrimination studies of Julesz (1981) and others. And of particular interest amongst the texture-discrimination investigations is that observation by Treisman, that detection of an O amongst Qs is difficult, whereas detection of a Q amongst Os is easy. This is readily explained by the unfinished business mechanism. Both parts of the Q, that is to say the circle and the line, set up their own regions of feedback inhibition. A subsequently viewed O will therefore run into the quenching effect of that circle. But if a preponderance of Os are viewed first, there will only be feedback inhibition corresponding to a circle, and the line portion of a subsequently-viewed Q will not be inhibited.

Finally, and most subtly, there is the fact that we can vary at will the degree of focus of our attention. I have argued that this degree is controlled by the BRF, through the variation of its disinhibition of the TRN, the latter, in turn, inhibiting the relevant part of the thalamus. As we noted, the BRF thereby acts like a gain control. Could it be that the unfinished business mechanism admits of a similar, and possibly related, variability? Perhaps the mismatch which can be detected by that process is actually a variable parameter, the sensitivity being somehow related to the fineness of the focus of attention. There are various ways in which the system could accomplish this, one being an unevenness in the packing of the inhibitory neurons of the TRN and another being a similar heterogeneity in the distribution of the axons linking the BRF to the TRN.

If there really is such unevenness, and if it varies systematically, it could be that my distinc-

tion between automatic and negotiable processes is too sharp. Perhaps negotiability simply diminishes with increasing distance from the center of gaze (in the case of vision), gradually to be replaced by an increasing degree of automation. This is pure speculation, however, and we will need to know more about the detailed anatomy of the TRN, and its interaction with the BRF, before the merits of this suggestion can be properly judged. Meanwhile, we ought to consider what happens on another time scale. These speculations about the unfinished business mechanism have been relevant to gamma oscillations. But the important frequencies in the hippocampus lie in the theta band. They are in the vicinity of 5 cycles per second, which indicates a time constant of about 200 milliseconds. This is of the same order as that round-trip transit time for signals travelling between a primary sensory area of the cortex to an association area. In order to understand how these processes could be related to consciousness, we will require some extra pieces of information. Let us consider them immediately.

We earlier considered Goldman-Rakic's suggestion that the prefrontal cortex might function as what could be called the mind's blackboard. And building on Broadbent's earlier investigations of bandwidth, we saw that Baddeley was able to argue for a tripartite blackboard, in which the actual job of marshalling the traffic of signals is carried out by a central processor. We also noted that the other two components of this system, namely the visuospatial sketchpad and the phonological loop, probably involve quite extended stretches of the cortex. It could be, therefore, that the prefrontal cortex is merely the site of the central processor, not the entire blackboard.

There is an independent piece of evidence which endorses this view. Although it is related to something that will not be discussed until later, it will be useful to briefly consider it here. I am going to be presenting what I believe to be a strong case for the idea that the underlying cause of autism is a pathologically low threshold against breakthrough. By this, I mean that autistic patients probably suffer from a confusion of signals not encountered by normal people. This breakthrough probably even applies to competing signals within a single sensory modality. But the intriguing thing is that there is strong evidence of inter-modal breakthrough in the case of some autistics. These patients appear to experience difficulties in handling multiple sensory inputs. They will listen to television, for example, but avoid watching the screen at the same time. There is even a recorded example of an autistic person (Williams 1992) being quite oblivious of a cat jumping into her lap if she was listening to someone speaking (Grandin, private communication).

Now whereas attention to someone speaking must have involved the phonological loop, the detection of the cat jumping into her lap would have required vision, and thus the visuospatial sketchpad. (It is true that knowing *something* was in her lap could have been managed by the tactile sense alone, but let us limit ourselves to competition between hearing and seeing.) The inter-sensory interference experienced by this patient indicates that Baddeley's central processor is not merely a place where signals from different sensory modalities are added. On the contrary, the possibility of this interference shows that the processor can also serve as an arbiter. There is no reason, indeed, why arbitration should not be its primary function, or even its *only* function; it may indeed function like a set of traffic lights.

In the case of the simple type of blackboard one encounters in the classroom, lines and symbols are drawn on its surface, and they remain there until they are erased. They are subsequently replaced by other lines and symbols, either immediately or later. One could liken the

situation to that in the brain when only one sensory modality is operative. But in the case of signal handling by the brain, the "writing" on the blackboard does not remain for an arbitrary amount of time. Instead, it persists for a well defined period that must be related to the temporal characteristics of the cortical region implicated, irrespective of whether it is the phonological loop or the visuospatial sketchpad that is involved. Earlier, when we considered what I loosely called cortical mechanics, we saw that these characteristic persistence times will be of the order of 200 milliseconds. They are governed by the time it takes signals to travel from the primary input area of a particular sense to the association area, and back again.

When two senses are operative, the equivalent blackboard is more complicated, and thus more difficult to envisage. But we can salvage the simple analogy used above by thinking in terms of action potentials rather than lines and symbols. Such action potentials will impinge upon the central processor, irrespective of whether they have come from the phonological loop or the visuospatial sketchpad. In this respect, the various senses all use the same symbolic elements, and this is presumably the reason why they can mutually interfere. And the analogy with writing on the two-dimensional surface of a blackboard is probably still useful, because the existence of a columnar structure in the cortex effectively makes the latter two-dimensional.

The time scale over which this traffic-controlling process operates would almost certainly be longer than the 25 milliseconds envisaged for the comparator mechanism described earlier. It would be marshalling signals that travel between a series of cortical areas, as well as to and from the frontal lobe. Although there is no supporting experimental data, one could suppose that the times involved would be similar to those invoked above, during the discussion of cortical dynamics, that is to say of the order of 200 milliseconds. This corresponds to the theta rhythm, which is of course clearly detectable in the hippocampus. Seen in this light, the participation of the hippocampus in memory processes that require consciousness is not particularly surprising. As was emphasized earlier, however, this does not mean that the hippocampus actually mediates consciousness.

Given compliance with the constraints implicit in what has been discussed in the preceding few paragraphs, the system will produce the results desired of conscious awareness. Just how reliable those results are will also be determined by the underlying anatomy and physiology. Oscillations may serve to provide coherence between signals emanating from sensory elements belonging to the same percept; they help a given sense to know what goes with what. As Damasio (1989) has stressed, there is an even more severe *binding problem* between elements belonging to *different* sensory modalities; different senses must also know what goes with what. Damasio's solution was the postulated existence of neural systems with the special task of integrating functional regions of the brain (or, to be more precise, the telencephalon).

From this definition, it would seem highly likely that animals close to our own species also possess conscious awareness. Monkeys certainly have reverse projections, and although there have not been measurements of the relevant characteristic times, it is at least a fact that cortical oscillations have been recorded in their visual areas. The same is true of cats. Just how deep the powers of reason are in these other animals is another issue, of course, but based on the known anatomy and physiology it does not seem too far-fetched to conclude that monkeys, cats, and similar creatures do think. *Aplysia*'s anatomy, on the other hand, does not seem to have the right network configuration. There appears to be nothing that could unequivocally be identified as reverse projections. Neither is there evidence of the sort of machin-

ery that would be required for the simultaneous handling of two parallel processing streams. It seems doubtful that *Aplysia* would be capable of thought.

The question of intelligence is only peripheral to what we are considering here, but some comment will not be out of place, because this attribute is presumably related to how much one can be simultaneously aware of. Let us again recall those contributions of Broadbent and Treisman. Broadbent's filter can be thought of as determining the number of items that can simultaneously be contributing to the metaphorical jig-saw. This idea was described in terms of pictures, because visual concepts are relatively easy to illustrate. The aperture synthesis concept should be generally applicable, however, and it would be particularly useful when considering abstract concepts. Now from what has just been postulated concerning the mechanism of conscious awareness, the anatomy and physiology of the situation permits two aperture syntheses to be accommodated in the system simultaneously. Treisman's breakthrough mechanism, as I see it at least, would be equivalent to one synthesis interfering with the other, presumably because of too much similarity between some of their elements.

What, in this admittedly qualitative picture, would be the determinant of intelligence? Perhaps it is simply the number of elements that can be present in either synthesis at a given time. This would allow us to retain the possibility that an individual's intelligence differs for different types of task, because some faculties might find it harder to insert appropriate pieces into the jig-saw, as it were. Treisman's breakthrough might influence things in a more subtle manner, by making some classes of item more susceptible to mutual interference than others. In any event, if the disturbances due to breakthrough were frequent and persistent, potential might not be fulfilled, and the innate intelligence might not be able to realize itself. This is highly speculative, of course, but the outcome could be retardation even though capacity was not lacking.

Finally, there is the question of *qualia*, which is to say what it actually *feels like* to be consciously aware. This is the thorniest issue of them all, of course, and some would say that a theory of conscious awareness is worth little without a satisfactory explanation of these concomitants. It would be difficult to gainsay that attitude, given that qualia are all that we have to go on, as subjective witnesses of consciousness. Feelings, if the above speculations have any substance to them, would have to come from that ability to detect correlations between cause and effect. The work of Libet and his colleagues shows that one particular type of touch sensation requires about 500 milliseconds to develop, the term *neuronal adequacy* being applied to this. From what was stated earlier, 500 milliseconds would be sufficient time for signals to *twice* make the round trip from peripheral area to association area and back to peripheral area. (It is important that the signals return to the peripheral area, because it seems that this is where images are mediated, as discussed at various points earlier in the text.) Although details of the mechanism are lacking, it does not seem implausible to suggest that the two round trips are required to establish the difference between cause and effect, if the *unfinished business mechanism* is responsible for making the comparisons.

In suggesting that this is the origin of qualia, I feel support from a certain type of observation, even though its source is decidedly unpleasant. There is a cruel type of torture known as *sensory deprivation*. The general idea is to deny the victim any sensory input whatsoever. This is actually a difficult technical challenge – thank goodness. It is easy enough to cut out all visual stimulation, of course, and total sound-proofing is attainable if one is pre-

Figure 9 This classic pair of figures, drawn by Minsky and Papert (1969), exposes the inadequacy of the typical perceptron. Apart from their different orientations, the figures appear to be quite similar, and would be classified as being so by the typical artificial neural network. In reality, the figures differ in a fundamental respect: that on the right comprises two topological components whereas that on the left has only one. The human cognitive system experiences a similar difficulty, and (sequentially operating) conscious attention has to be used in order to discern the difference. This effort involves short-term memory, in order to keep track of the lines and the spaces.

pared to go to sufficient expense. Taste and smell stimuli can be eliminated relatively easily. To cut out all tactile stimuli, the perpetrators of this brand of horror have devised moulds that fit the body so closely that even the splayed-out fingers are unable to touch one another. Whether or not there is even provision for isolating the tongue, I do not know.

My reason for citing sensory deprivation is that the victim is denied all possibility of detecting *correlations between cause and effect.* Had it been possible to rub one finger with another, such correlation would still be apparent, but as I say, even this primitive (but perhaps salvaging) possibility is removed. So what happens to the victim? It is found that the mind literally begins to fragment. Hallucinations begin to impose themselves within an hour or so, and prolonged abuse of the victim can bring on insanity. One might feel that detection of correlations between things imagined would be adequate to maintain mental tone. Perhaps this is the case, and approximately an hour might be as long as one can keep things going in one's thoughts.

My only reason for describing such a nefarious treatment of one's fellow humans is that it lends strong endorsement to what I have been suggesting. Detection of the correlation between cause and effect, rapidly enough to permit reaction, is postulated to be the defining characteristic of conscious awareness. When a person is prevented from making such detection, and is furthermore prevented from reacting to it, consciousness is disturbed.

Returning to the question of how the organism would benefit from the above-defined mechanism of consciousness, we could hardly do better than to contemplate the pair of diagrams used by Minsky and Papert, to expose the shortcomings of the neural network known as the perceptron (Minsky & Papert 1969). These are reproduced in figure 9, and we see that they appear to be very similar. Indeed, it looks as if the one might simply be a rotated version of the other. The perceptron must carry out its cognitive task in a simultaneous and parallel fashion, and it is unable to detect the quite subtle difference between the two. Neither can we, unless we trace out the course taken by the worm-like object, or the inter-worm space, in each version. Only then do we find that there is a quite fundamental difference.

But let us not overlook what is being demanded of our systems, in order to carry out this discriminatory task. As we follow the curved spaces, we have to continuously detect the bounding lines, and systematically avoid them. And we have to continuously keep track of what we have experienced. Indeed, we have to bear in mind what our overall strategy was to be: are we following the course of a worm or an inter-worm space; have we encountered a blockage, and so on? It is true that we could augment our short term memories with pencil and paper, to help us keep track of events, but this will merely replace one task by another. This accomplishment might appear to be a modest one, but it is the type of elementary process that lies at the heart of more complicated cognitive achievements.

Embedding

Considerable significance has been attached to the concept of bandwidth in mental processing, and I have been allying myself with the views of Broadbent on that issue. Indeed, I essayed an explanation of bandwidth in neural terms, describing it through analogy with a constantly-fading jig-saw, and used the term aperture synthesis to formally describe the process. It does not seem too farfetched to assume that bandwidth determines the depth of reason that an individual is capable of, greater bandwidth permitting greater depth. It might even be the case that bandwidth is somehow related to intelligence itself. This view would take as its justification the idea that intelligence ultimately stems from an ability to notice the relationships between things, and such relationships can be discerned only between the things that the mind can simultaneously accommodate. It has been said that the difference between a juggler and a buffoon is just one extra ball or club; if the system is overloaded it simply breaks down. I believe that bandwidth limitation may be a contributing factor in autism; autistic people may not be capable of the depths of reasoning accomplished by normal people, simply because their minds have smaller-than-normal capacities.

Although this is something of an aside, I am reminded of a scene in Peter Ustinov's play *Romanov and Juliet*, in which an obsequious character in an imaginary country has a series of meetings with the American and Russian ambassadors, alternating between them. When informing the American that the Russians have a secret weapon, he is told that they, the Americans, already *know* this. He then meets with the Russian ambassador and tells him that the Americans know about the secret Russian weapon, only to be informed that the Russians *know* that the Americans know about it. At his next meeting with the American ambassador, where he again attempts to ingratiate himself by exposing something he believes to be valuable, he is dismayed to discover that the Americans already *know* that the Russians know that the Americans know. The series continues a few more times, the story getting more complicated on each occasion, and it ends with one of the ambassadors exclaiming *Oh! We didn't know, that they knew that we knew that they knew that we knew* – and so on. The trouble is that once the story contains more than a few links, it becomes impossible to keep track of the significance of what is being revealed. My point in relating this fascinating piece of theatre is that it nicely illustrates the temporal counterpart of the aperture synthesis mechanism. Once things exceed a certain logical depth, we cannot fathom them. The psychologist speaks of *embedded inference*, and human limits on embedding are not impressive; only about four steps make our species uncomfortable.

In a subsequent paper on the theory of mind theme, Frith, Morton and Leslie published a diagram comprising data on 47 autistic people, only 10 of whom passed the Sally-Anne test (Frith, Morton & Leslie, 1991). They used this diagram to make the point that the great majority of those 10 people had mental ages in excess of five and biological ages greater than eight. This well-documented work was most valuable, because of the statistical material it contained. And although it clearly established that lack of a theory of mind is a characteristic of the majority of autistics, the lack was nevertheless being shown to be not universal. I believe that this paper can provide further valuable insight into the question of autism. The point is that mental age divided by biological age, and

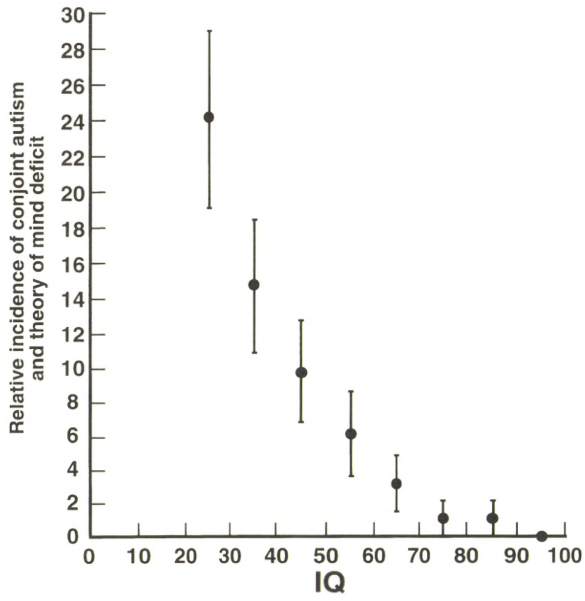


Figure 10. This plot of the relative incidence of conjoint autism and theory of mind deficit, as a function of intelligence quotient, was derived from data published by Frith, Morton and Leslie (1991). (The error bars are standard deviations.) It shows that the distribution of IQ for such individuals sharply contrasts with what is observed for normal individuals, the (Gaussian) distribution for the latter having a peak at IQ=100. From the evidence of this plot, it appears that the autistic peak is below IQ=20, and it might even be at or close to zero.

multiplied by 100, gives the IQ. It was thus a straightforward task to draw in the iso-IQ lines (i.e. lines of constant IQ) on the Frith-Morton-Leslie diagram, for the 47 children investigated. Even though there was no data for mental age lower than four, nor any for biological age below eight, it is a straightforward matter to make corrections for this, and thereby obtain a relationship between incidence of autism and IQ that extends right down to the point where IQ is 20. Although the testing of mental age is not without its difficulties, it is to be emphasized that all the children in these tests had been subjected to the *same* probing of mental age.

The resulting graph was remarkably smooth, and it revealed a systematic increase in the incidence of autism with decreasing IQ, as shown in figure 10. The systematic increase was no surprise; it had previously been noted by many researchers in this field that autistics tend to have sub-normal IQs. It must be emphasized that this is a statistical trend, however. As we saw earlier, there are autistic people with relatively high IQs (although not one of the above 10 had an IQ in excess of 100; none of them was *better* than the norm), and there are people with low IQs who are not autistic. But that is not the point. The importance of this graph was its demonstration that the *probability* of being autistic does increase as the IQ falls. And the real surprise lay in the suggestion that the *distribution* for autistics is shifted downward compared with normal people.

The distribution of IQ amongst normal people in a typical population is of the normal (Gaussian) type, the peak lying at IQ=100. The data of Frith, Morton and Leslie was indicating that the distribution for autistic people may well also be normal, but that the entire curve is shifted downward toward lower IQ. Indeed, there was no sign of a maximum being reached in the autistic distribution, even for the IQ=20 mark. The data are consistent with a maximum lying at IQ=0. The other noteworthy feature of the work was the one pointed out by the authors themselves, namely that those autistics who *did* have a theory of mind tended to be those with higher IQs.

Now given that lack of a theory of mind was being put forward as a major, and possibly *the* major, autistic characteristic, one could wonder why these higher-IQ members of the group were being classified as suffering from the infirmity. In what respect were they nevertheless autistic? To shed more light on that aspect of the issue, one would like to know how mental ability with respect to a given task, such as a theory of mind test, is actually related to the IQ

level. A full solution to this problem will presumably not be possible until we know more about the neuronal determinants of IQ, which we will be considering later. Meanwhile, let us see how far we can get with qualitative arguments. Different tasks vary in the demands they make on our mental powers. I believe that it is safe to assume that the number of easy things that our brains can manage far outstrips the number of difficult things. One can imagine the totality of things that demand a given level of mental proficiency to occupy a sort of *mental plane*. (This would no doubt have a great number of dimensions, so it would be a mental *hyper*plane.) My assumption thus suggests that the area of such a plane diminishes as the mental level becomes deeper. A key point is that this should apply to all individuals, irrespective of their IQs. All people, autistics included, can manage tasks that make different mental demands; they are able to operate on different planes.

Two questions arise with respect to these mental planes, in a given individual: how does the *state* of a given plane vary with IQ, and how is the mental level of the *deepest plane* related to IQ. That there is a deepest level is indicated by our readiness to accept that *certain things are beyond us*. In this respect, I happily agree with McGinn's concept of cognitive closure (McGinn, 1991), and will hereafter actually use his term. (McGinn's formal definition of *cognitive closure* states that a mind M is cognitively closed with respect to a property P (or theory T) if and only if the concept-forming procedures at M's disposal cannot extend to a grasp of P (or an understanding of T).) What is implied by the state of a mental plane? It is meant to convey the degree of knowledge of things requiring a given depth of reasoning. The neonate's various mental planes must be rather sparsely populated; just after birth, there are numerous *holes in the mind*. As learning proceeds, the various planes begin to fill out, but there will clearly be a large environmental contribution to this process; except for the occasional flash of insight, we tend to know what we have been told. Although it is possible for an individual to work things out for himself, the general situation will be that holes will exist where informational input has been lacking.

To progress further, we must ask how the situation changes as we shift the focus from a shallow mental plane to a deeper one. I believe that the difference lies in the demands made on cooperativity. We colloquially speak of *coherence*, using it in the sense of having one's wits about one. Perhaps this is related to actual physical coherence, by which I mean the cooperative functioning of different parts of the brain. Let us now recall that aperture-synthesis process. It too could be said to provide a measure of coherence, in that the instantaneous size of the composite picture is determined by the number of snapshots that can be present in the system at any one time. This is obvious when the aperture is actually being synthesized in the visual domain. But suppose that a similar process is at work whenever various items stemming from several sensory modalities are contributing to what is instantaneously in the mind. In that case, the phrase *getting the picture* could take on a deeper significance.

We have accumulated a number of observations which clearly require resolution, so let us take stock for a moment. To begin with, there is the fact that the distribution of IQ for autistics is shifted toward lower values, compared with normal people. Then we have the observation that autistics with higher (and even normal) IQs can actually *have* a theory of mind. Finally, there is the fact that normal people with low IQs can have a theory of mind. The inescapable conclusion is that lack of a theory of mind is not an absolutely necessary condition for having autism. It seems, rather, that it should be regarded as a secondary symptom; it appears to be a consequence of something

more fundamental rather than the actual root of the trouble.

I am strengthened in this conclusion by the autobiographies of autistic individuals. It is uncommon for an autistic person ever to reach the stage of development at which book writing can be undertaken. There have been a few notable exceptions, however, and their cases are clearly very significant; their accounts provide rare opportunities for delving into the autistic mind. Two such autobiographies were recently produced by Williams (1992) and Grandin (see Grandin & Scariano 1986). Williams was initially labelled deaf, abnormal, retarded, spastic, crazy and insane, amongst other things. By her own admission, she existed in a state of dreamlike recession. And yet she yearned to become normal. She had reached the age of 25 before hearing the word autism for the first time, and she had the insight to appreciate that the description of its symptoms fitted her own case. Against the odds, Williams achieved a place at university, lived independently, and ultimately wrote about her experiences and, very significantly, her thoughts. Grandin has had a similar uphill struggle in life, and she too has achieved academic qualifications. With a vengeance, indeed, for she holds a Ph.D. in veterinary science.

There is no mention, in either of these two books, of Williams or Grandin ever having been subjected to the Sally-Anne experiment, or indeed any other theory of mind test. And this is hardly necessary, because they both reveal through their writing that they are aware of other people's thoughts. Williams makes the first reference to this on the second page of her book, when she recounts that her parents thought that she had leukaemia. Grandin's first reference to the thoughts of others is on her third page, and it refers to her mother's belief that her daughter was merely somewhat slow in developing.

It appears, then, that we must look beyond appreciation of thought in others, in order to get at the real origins of this distressing syndrome. We are looking for something which can lead to a lowering of the IQ, even though this is not an essential aspect, and we are looking for something which is not in itself a consequence of low IQ. We are naturally looking for an explanation of the theory of mind deficit, in particular, as well as to Wing's triad of deficits in general. Returning to the descriptions of the symptoms that were enumerated at the beginning of this paper, we ought also to be looking for an explanation of those things in what I dubbed the second category, for they were not obviously related to a theory of mind deficit. To make any further progress, we are going to have to turn to the bottom-up approach and ask what anatomy and physiology are able to expose.

Genetics of Autism

There are strong indications of a genetic origin of autism. This can be seen in the statistics of its occurrence. When one child of a given pair of parents is autistic, the probability that another of their progeny will also be autistic is much higher than the overall incidence in the population would indicate. We noted earlier that the latter figure is 4 in every 10,000, which is to say 0.04 per cent. In fact the probability of another autistic child being born to parents who already have one is 2 per cent, which exceeds the random-chance rate by a factor of about 50. Twin studies have been a particularly fruitful source of information. In one such investigation, carried out by Folstein and Rutter, 4 out of 11 pairs of identical twins showed concordance, that is to say that both twins suffered from the syndrome (Folstein & Rutter, 1977). Of the remaining 7 pairs, the non-autistic twin was nevertheless found to show some autistic traits (such as language impairment and truncated intellectual development) in 5 of the cases. There was no concordance amongst the 10 pairs of non-identical same-sex twins, although here too there were autistic tendencies in the other twin in one of the cases. (Unless there was a factor related to the specific pregnancy, one would not expect the concordance rate for non-identical (fraternal) twins to be any higher than for ordinary siblings.) The data on non-identical twins are not surprising, because we have just seen that the probability of two autistics being born into in a single family is 2 per cent. The sample of 10 non-identical twins was thus too small by a factor of about 100, to produce statistically significant results.

There have been several other genetic investigations of autism since the classic paper by Folstein and Rutter, and Smalley and colleagues found strong evidence for a genetic origin of the affliction in their recent survey (Smalley, Asarnow & Spence 1988). In particular, the twin studies carried out by Ritvo and his colleagues (Ritvo, Freemen, Mason-Brothers, Mo & Ritvo 1985) revealed a 95.7 percent concordance for monozygotic twins, and a 23.5 percent concordance for dizygotic twins. Comparing this with the incidence of autism in the population as a whole, namely 0.02 to 0.05 percent, one sees that the evidence for a hereditary mechanism is very strong indeed. Smalley and colleagues note that the concordance for dizygotic twins indicates an autosomal recessive mechanism, and this was also the conclusion of Ritvo and his colleagues (Ritvo, Spence, Freeman, Mason-Brothers, Mo & Marazita 1985).

It is one thing to demonstrate that there is a hereditary component to autism, but quite another to say just where the error lies. There are 23 chromosome pairs in humans, and each of these caries the biochemical instructions for the synthesis of several thousand different types of protein molecules, the information for each being contained in what is known as a gene. These molecules come in different classes, including examples specialized for metabolic processes, signalling in the immune system, and signalling in the nervous system, while others serve structural functions. A single mutation in a gene is sufficient to cause the corresponding molecule to be produced in an incorrect form, and the result can be malfunction that can be more or less severe.

Mutations occur at the level of the bases in the DNA molecules, which are one of the two constituents of the chromosomes (the other being histone molecules). These bases consist of a couple of dozen atoms, so a mutation would not normally show up when the chromosomes are viewed under the microscope. An

exception occurs in the case of what has come to be known as the *fragile-X syndrome* (Fish, Cohen, Wolf, Brown, Jenkins & Gross, 1986). The X refers to the X chromosome, and when the syndrome is present this appears thinner along a portion of its length, making that stretch look easy to break. The condition was clinically recognized before the discovery of its association with the fragile-X site, and it was known as the Martin-Bell syndrome (Martin & Bell, 1943). It is frequently accompanied by mental retardation. Gillberg has reported on a case in which identical triplets were all autistic and all bearers of the fragile-X site (Gillberg, 1983.). The syndrome is observed far more often in males, presumably because these have only the one X chromosome. Females have two, and if their other copy is normal, the influence of the faulty member may be negligible. Several macroscopic consequences of fragile-X syndrome are impressively similar to some of the symptoms seen in autism, including aversion to eye contact, language difficulties, and aversion to physical contact. These must stem, in some way or other, from microscopic abnormalities. But before moving on to that side of the question, let us consider other clinical evidence.

A particularly common observation concerns perinatal complications associated with the syndrome. More than one in every three births of autistic children involves problems such as delayed breathing (Kolvin, Ounsted & Roth, 1971). This had led to the suggestion that lack of oxygen, however brief, could be a contributing factor in the development of autism. This would seem to indicate something other than a genetic origin for the impairment. It might be more relevant to regard such birth difficulties as consequences rather than possible causes, however. If the fundamental fault lies in the (genetically-determined) development of the nervous system, difficulties such as delayed breathing could be one of the outcomes. There is recorded evidence of faults in neuronal formation in autistic patients (Gillberg & Forsell, 1984).

Anatomy and Physiology of Autism

The ultimate products of neuronal development are the individual brain structures, of course, and there is ample evidence of malformation in several of these. A recent report by Bauman (1991), following up an earlier postmortem investigation by herself and Kemper (Bauman & Kemper, 1985), describes abnormalities in the hippocampus, subiculum, entorhinal cortex, septal nuclei, mammillary body, and selected nuclei of the amygdala. These are all components of the limbic system, and the aberrations point to impaired emotional function in autism. Deformities were also discovered in the cerebellum (specifically, the neocerebellar cortex and the roof nuclei). Because there was no observable loss of neurons in the inferior olive, which is the last port of call for signals before they impinge on the cerebellum, Bauman was able to conclude that the observed abnormalities began or occurred before birth. Arin, Bauman and Kemper (1991) have established that there is a reduction in the number of Purkinje cells in the autistic cerebellum, the deficit lying in the range

of 50-95 per cent. Evidence of Purkinje cell depletion in autism has also been reported by Ritvo and his colleagues (Ritvo, Freemen, Scheibel, Duong, Robinson, Guthrie and Ritvo 1986). Given the regulatory role in movement that the cerebellum plays, this loss could underlie the motility deficits in autism discussed in the first section of this paper.

Modern brain imaging techniques are playing an increasingly important part in the investigation of living autistic people. A recent review of the results obtained by computed tomography (of X-ray images – also known as CT scanning), magnetic resonance imaging (MRI), and positron emission tomography (PET) has been reported by Courchesne. He noted that one of the most frequently found abnormalities in autism, irrespective of whether it was accompanied by mental retardation, is reduction of cerebellar tissue (Courchesne 1989, Courchesne 1991). He and his colleagues encountered cases in which some parts of the cerebellum were only half their normal size. In agreement with what was mentioned earlier, the evidence indicated that the reduction of tissue was a result of developmental abnormalities rather than damage following full development. The statistical significance of this work was beyond reproach; the MRI study alone involved 283 autistic patients. Not all parts of the cerebellum are affected, the regions most depleted being the cerebellar hemispheres and the vermis (Murakami, Courchesne, Press, Yeung-Courchesne & Hesselink 1989).

There has been a study of glucose metabolism in the cerebella of (seven) autistic patients which complicates the picture somewhat. It was carried out by Heh, Smith, Wu, Hazlett, Russell, Asarnow, Tanguay and Buchsbaum (1989), and it showed that the metabolic rate is not lower in autistics, despite their decreased number of Purkinje cells. (We should recall that the Purkinje cell is the pivotal neuron in the cerebellum; it is the ultimate focus of all the afferent axons, and it is the sole source of efferent axons.) Heh and his colleagues actually found that there was a slightly *higher* metabolic rate in several cerebellar regions, including the cerebellar hemispheres, compared with normal people of the same age. They suggested that these places of elevated glucose turnover might represent islands of inefficient metabolism, which could be sites of early neurodevelopmental abnormality, the upshot being redundant and poorly integrative neural circuits.

This idea is very much in harmony with the recent changing of attitude toward cerebellar function. That it is important to motor coordination is beyond doubt, but there is a growing feeling that it also plays a role in cognitive processes, as was suggested by Leiner, Leiner and Dow (1986). This applies in particular to the neocerebellum, which is the evolutionarily newest region of that major brain component (Ivry & Baldo 1992). Wallesch and Horn (1990) found that patients with lesions in the neocerebellum show deficits in cognitive operations connected with spatial perception (see also Bracke-Tolkmitt, Linden, Canavan, Rockstroh, Scholz, Wessel & Diener 1989). Particularly significant, for what we will be discussing a little later, was the finding that memory was *not* affected, even though there were clear indications of impairments in verbal IQ, associative learning and spatial reasoning. This latter result was corroborated in a most direct manner by a PET investigation undertaken by Decety, Sjöholm, Ryding, Stenberg and Ingvar (1990). Observations were made on normal people who were performing various controlled tasks, such as the mere *imagination* of hand movements. The technique was used to monitor the blood flow in various cerebellar regions, this giving a good indication of the metabolic activity. There were clear indications that the cerebellum is indeed involved in such processes.

As has been stressed by Akshoomoff and

Courchesne (1992), the involvement of the cerebellum in malfunction could easily implicate other brain components. They pointed out that the sensory modulation properties of the cerebellum stem in part from its interaction with areas known to be important for selective attention, such as the pulvinar, the superior colliculus, and the parietal and frontal cortices. Courchesne himself, in collaboration with Press and Yeung-Courchesne, has very recently demonstrated parietal lobe abnormalities in autistic patients, using magnetic resonance imaging (Courchesne, Press & Yeung-Courchesne 1993). This is an exciting development, because there had been prior work on non-autistic people which revealed a most intriguing consequence of damage to the parietal lobe. Posner, Walker, Friedrich and Rafal (private communication) discovered that in patients with such lesions there are difficulties in the *disengagement* of attention. The usual course of events, when a visual object is being scanned, for example, is disengagement, followed by a shift of attention, followed by re-engagement on the new location. In parietal lobe patients, apparently, there is a tendency to stay locked onto one point in the visual scene (say) for excessive amounts of time. The time intervals involved are quite brief, of course, but the effect was nevertheless pronounced. It is as if people with this type of damage become temporarily fixated. Courchesne and his colleagues have gone on to establish that similar effects are indeed seen in autistic patients.

Although we are specifically dealing with humans, here, there is a strong temptation to cite some relevant work on monkeys. If these creatures have sustained damage to the parietal region, there is a reduction in their level of conscious awareness, and they seem to withdraw into self-isolation. They lose the directed, information-seeking behaviour typical of normal primates (Mountcastle 1978).

Autism at the Molecular Level

Having considered abnormalities at the level of brain components, each with their millions of neurons, let us now focus right down to the molecular level. There is a strong connection between autism and epilepsy, and there have been suggestions that part of the trouble with the latter syndrome is rooted in biochemistry. But first, we should look at the statistics. As Minshew (1991) has noted, epilepsy is estimated to be present in 20-35 per cent of autistic individuals (see also Volkmar & Nelson 1990, and Rutter 1984). Of the seizure types observed in these cases, major motor seizures are the most easily recognized, and they account for the majority of documented cases. Most autistics who also have epilepsy develop it either in early childhood or in adolescence. It is very interesting to note that there is a connection between the cerebellar vermis and the hippocampus (and septum). This is in the form of a feedback loop, and the fascinating thing is that the loop has been circumstantially linked to epilepsy and emotional disorders. Unfortunately, this data was obtained from monkeys and cats, but we saw above that the monkey data, at least, ought to be reasonably applicable to our own

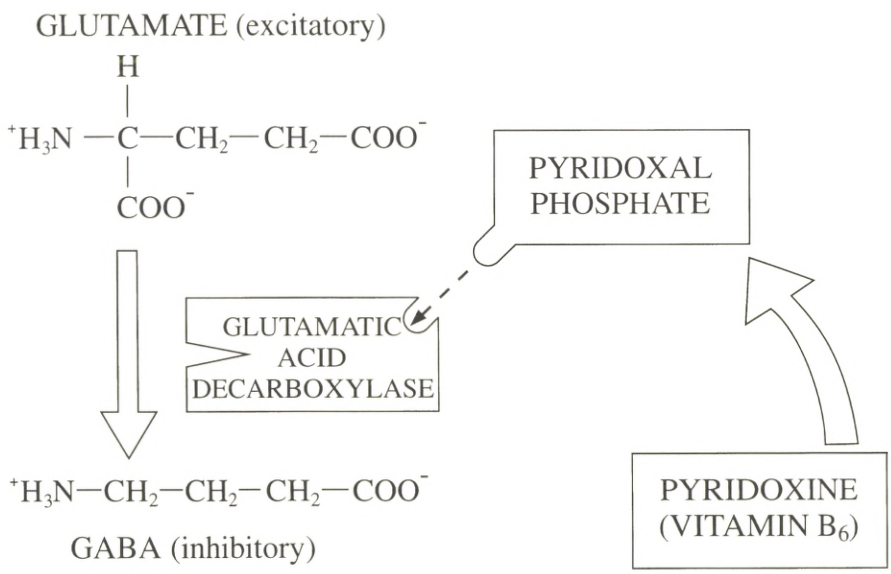


Figure 11 The inhibitory neurotransmitter GABA is derived from the excitatory neurotransmitter glutamate, the conversion being effected by glutamic acid decarboxylase. This enzyme requires pyridoxal phosphate as a co-factor, molecules of that substance being formed from pyridoxine, better known as vitamin B6. Heavy doses of that vitamin have been found to produce amelioration of the symptoms of autism, to a much more pronounced degree than that seen with any other substance yet used in such treatment.

species. The study was carried out by Heath, Dempesy, Fontana and Fitzjarrell (1980), and they used the EEG as a means of monitoring the effects of stimulating various regions. It turns out that provocation of the cerebellar vermis inhibits epileptic seizures emanating from the hippocampal region, and it was also found that there is a feedback route from the hippocampus to the cerebellar vermis.

Turning now to the molecular level, as promised, there has long been the suggestion that epilepsy stems from a biochemical fault. The classic paper on this topic was written by Meldrum (1975), and in it he reviewed the link between the inhibitory neurotransmitter GABA and seizure; insufficient GABA in vulnerable regions of the brain promotes kindling of epilepsy. Now it so happens that the production of GABA in the appropriate cells uses glutamate as a starting point (Stryer 1981, Cooper, Bloom & Roth 1991). It is quite fascinating that molecules of the major inhibitor should be produced by modifications to one of the most prominent excitatory neurotransmitters. Both these molecules are rather small, each containing a dozen or so atoms, but the modification from the one to the other requires the aid of a much larger molecule, namely the (protein) enzyme known as glutamic acid decarboxylase (GAD). This difference in sizes is quite usual; the molecule making the changes needs a lot of muscle. Now in order to do its job, GAD needs the assistance of a smaller molecule by the name pyridoxal phosphate, and this molecule is, in turn, produced from pyridoxine, better known as vitamin B6 (see figure 11).

Let me now bring this little journey to an end by stating that one of the substances found to be most effective in ameliorating the symptoms of *autism* is that same vitamin B6 (Lelord, Muh, Barthelmy, Martineau, Garreau & Callaway 1981, Martineau, Barthelmy, Garreau & Lelord 1985, Martineau, Barthelmy & Lelord 1986, Rimland 1987, Martineau, Barthelmy, Cheliakine & Lelord 1988). I believe that this leads to the following picture. The tendency toward epilepsy in autistic patients may be due, in part, to a faulty metabolic pathway between glutamate and GABA, and this leaves certain cells with a deficit of that inhibitory neuro-

transmitter. The upshot, I believe, is neurons which are too easy to excite. This, in turn, is conjectured to lead to kindling, with the resultant full blown epileptic fit, if the deficit of GABA is sufficiently pronounced. (The view that autism might be linked to a lowering of the threshold for excitability had earlier been expressed by Ornitz (1987).)

From what I have described thus far, it might seem that we have merely been considering a side issue, of an admittedly disturbing nature, but a side issue nevertheless. Now let me bring this story right into center stage in the context of autism. Suppose that the majority of autistic patients, rather than just the one in three who develop the full version, actually have latent epileptic tendencies. This would mean that most of them would have nervous systems that are too easy to excite. Let me then make a further jump and say that this could mean that breakthrough, of the type discovered by Treisman (which we discussed in detail earlier), could be a common occurrence in the autistic brain. The actual neuronal mechanism of breakthrough is not yet known, but it does not seem too farfetched to suggest that it is linked to the exceeding of a threshold, and that this could result from a failure of inhibition. We have already contemplated what deleterious effects breakthrough has on attention. Perhaps this is a burden that the typical autistic person may be saddled with. I will return to this theme again in the next section, after bringing some other facets of the issue into focus.

Roots of Intelligence

There are quite pronounced differences between the EEG patterns of autistics and normal people. The normal pattern, as we have seen, continues out to about 500 milliseconds, and Libet and his colleagues have explained this in terms of what he and his colleagues termed neuronal adequacy (Libet, Wright, Feinstein & Pearl 1979). The latter, in turn, was seen to be related to conscious awareness. In the EEG trace of the typical autistic person, the amplitudes of the various peaks are less, this tendency increasing for longer times (within the 500 milliseconds, that is). The indication is thus that whatever causes the various swings in signal, in the normal person, is truncated in the case of the autistic. (The diminished signal strengths in the EEG of autistic people was reported in Dawson, Finley, Phillips and Galpert's paper entitled "P300 of the auditory evoked potential and the language abilities of autistic children", which was cited by Courchesne (1987) as being in preparation). It is as if signals are having a harder time getting through to their destinations, and they therefore decay away faster than is normal.

This result is interesting in its own right, of course, but it takes on added significance when one notes that a marked relationship has been observed between EEG characteristics and intelligence. This was established by Ertl and Schafer (1969). If one compares the evoked potentials of people with high IQ to those recorded from people with low IQ, the former are seen to have more troughs and valleys, for a given time interval (of about 250 milliseconds). This indicates that people with low IQs

lack certain high frequency components in their response characteristics, the missing frequencies lying in the vicinity of 40 cycles per second. (The actual data are for 10 people with IQs in the range 120-142, compared with 10 people having IQs in the range 62-89.) This result is rather suggestive, because that was the frequency of those oscillations which were observed to be related to coherence in visual scenes, and which were conjectured to be associated with feature binding. If, as has been suggested by Crick and Koch (1990, 1991), the 40 cycle per second oscillations are a ubiquitous feature in conscious awareness, the suggestion could be that people with low IQ may have a harder time capturing correlations in sensory input.

Miller (1956) has referred to a magical number of information processing. In fact, this was not so much a single number as a *range* of numbers, stretching from 5 to 9; he called it *the magical number seven plus or minus two.* He had discovered that normal people can repeat between five and nine digits that have just been read out to them, and this number is now referred to as a person's digit span (Miller 1956). Digit span was the first thing that Broadbent employed when he was carrying out the investigations on attention that led to his filter theory. One can feel the circle of evidence beginning to close, because digit span has been found to correlate with the evoked potential. Polich, Howard and Starr (1983) have established this, by showing that the normal 300 millisecond peak in the EEG trace tends to be delayed and diminished in people with lower spans.

With IQ being related to EEG characteristics, and EEG characteristics being related to digit span, one might guess that digit span would have to be related to IQ. This is the case, as was established by Ellis (1963). I have tried digit span tests on a number of autistic people. These results showed a high degree of mutual

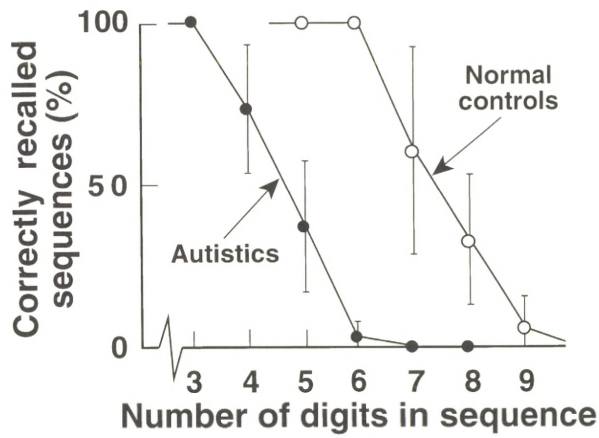


Figure 12. The results of digit-span tests on five institutionalized autistic patients and five age and sex matched normal individuals. Despite the large error bars, caused by the smallness of the sample, the data reveal a clear difference between the two groups. Because the tests were performed aurally, the poorer performances of the autistic individuals might indicate deficiencies in their phonological loops. If the ideas put forward in the present paper regarding the underlying cause of autism are correct, the intelligence levels of the autistic probands would also show truncation provoked by breakthrough.

consistency, for the five patients I studied, and they revealed a marked contrast to what Miller observed for normal people, as is shown in figure 12. For sixe digits, virtually all normal people have a 100 per cent score. I found that for autistic people the chance of recalling six digits is virtually zero; for the average patient, the magical number is four and a half, plus or minus one and a half.

What are all these numbers and EEG characteristics a measure of, however? I believe that they are indeed related to those pioneering studies of Broadbent, and that they also owe much to Treisman's work. Let me sum up the main points of this admittedly long story. Firstly, we saw that a deficit of GABA may be a major autistic feature, with the concomitant diminution of inhibition and a tendency toward too much excitation. This was then suggested to make the autistic vulnerable to breakthrough.

It can now be added that even Grandin, who *has* a theory of mind (and a Ph.D.) but is nevertheless autistic, lists as one of her main troubles the inability to handle more than one information stream. And Williams has even reported that she has been unaware of a cat jumping into her lap if she is listening to somebody talking (Grandin, private communication). (Rimland (1964), noting that there is a trade-off between fidelity and bandwidth in cognitive processing, has written of individuals who were so prone to become lost in thought that they were prohibitively dangerous behind the wheel of a car.) Let us also recall the isolation of the typical autistic person. Could it be that breakthrough is the root of the trouble? If that is so, the autistic person's inner life may be a veritable cacophony of signals, and concentration may be extremely difficult unless information is arriving in only a limited fashion. If it is not, the system might automatically decrease its sensitivity level, though the cortex-BRF-TRN-cortex loop discussed earlier, thereby exacerbating the bandwidth-limiting effect.

Turning to the question of intelligence, and its correlation with the above numbers and EEG characteristics, one could speculate that IQ may be related to the aperture-synthesis principle. This came from Broadbent's ideas, we recall, combined with the observed function of the pulvinar. To be specific, the more items that can simultaneously be accommodated in the aperture, the greater will be the chance of noticing correlations, and this would indeed appear to make for greater mental power. But there is more to mental power than the mere gathering of information. This information must be used for making predictions, because in default of forecasts there can be no surprises. This surely is a prerequisite for common sense.

We have nearly reached our goal. Miall (to be published) has suggested that the cerebellum may function as a predictor-corrector, and we have seen that the cerebellum of the autistic person is invariably impaired. Then there was that suggestion of an excess of excitation in the neural networks of the autistic brain. Perhaps this means that the autistic suffers from too many winners in those winner-take-all processes. In the particular PDP network known as the perceptron, such excess can deny the network its capacity for generalization, essentially converting it to a look-up table (Rummelhart, Hinton & Williams 1986). The brain is not a perceptron (Crick 1989), but a similar shortcoming may nevertheless be one of the results of the above excess of signals penetrating the inhibitory cloud. By pigeon-holing everything as if it were a special case, the autistic brain's networks might get cluttered up with the particular, at the expense of the general. It might become, in other words, bogged down with rigidly-held impressions, and it might not be able to read between the lines. It might, we could say, become obsessed with sameness. As we saw at the outset, this is one of the characteristics of that second group of symptoms, which were not explained by a lack of a theory of mind.

Although these are mere indications, rather than a coherent theory, I believe that the consequences for the sufferer of these deficiencies are reasonably clear. And here I would like to turn to the work of Boden (1990). Her analysis of creativity led her to suggest that the key attribute is *the ability to vary the variable.* Indeed, one could say that the first hurdle is recognition *of* variability. Einstein's path to the *theory of relativity* was through mentally letting the speed of an object *vary.* It does not seem too far-fetched to suggest that Leslie's ideas about the origin of a *theory of mind* could likewise be perceived as an ability to imagine *variability* (Leslie 1987). Is this not what we do when we pretend that one thing represents another?

It seems possible that breakthrough could also underlie another autistic symptom which still awaits an explanation, namely those ritua-

listic and bizarre behaviours. In his biography of his autistic son Raun, Kaufman (1976) described the bouts of hand flapping brought on by a plate spun on its edge. If the plate's influence stemmed from the rapid optical periodicity caused by the spinning, this would be reminiscent of the effect of rhythmically flashing lights on epileptics. And we have seen that there may be certain similarities between autism and epilepsy at the microscopic level; the exceeding of a pathalogically low threshold might occur in both afflictions. Thresholds can be exceeded even in normal people, of course, a relevant example being the unconscious drumming of one's fingers on the table during periods of irritation or impatience. Nail biting and head scratching are related phenomena. These things serve no constructive purpose, even though they might help to relieve tension. McFarland (1966) has analyzed such out-of-context activities, and his proposed mechanism for their occurrence invokes activation of irrelevant motor commands as a result of a process that resembles the exceeding of certain thresholds. Although this is an admittedly more tenuous conclusion, it seems possible that unusually low thresholds for certain motor programmes could underlie the impaired motility this is a common feature of autism.

We have been following a long and tortuous path, and as I say, these are sketchy ideas rather than a well-rounded theory. It is encouraging, however, that all those other things about autism that were not explained by the theory of mind concept may now be on the threshold of clarification. If that is so, there would seem to be one firm hint as to what environment would be helpful to sufferers of this devastating affliction: they might fare best under conditions in which the amount of information impinging upon their senses is carefully controlled. If the simultaneous demands of sight and hearing really are threatening them with sensory overload, as I am suggesting, it might prove worthwhile to try communicating with them verbally in darkened rooms, and visually in total silence. They will not be able to avoid situations which make multi-sensory demands, of course, but it might prove helpful to introduce these in a controlled fashion, following such mono-modal priming. And the transition to the multimodal situation could be made less burdensome if the amount of information being given to the young autistic is strictly limited. In view of the (conjectured) truncation of what has here been called the aperture synthesis mechanism, it might be a good idea to initially limit this information to a pairing of monosyllables with simple iconic diagrams.

Some Predictions

In this paper, I have suggested that the underlying cause of autism is inappropriate interference between neural signals caused by breakthrough, either within a single sensory modality or between different modalities. And I have suggested that such breakthrough is a consequence of failure of inhibition, specifically citing the possibility that insufficient glutamate is being converted to GABA. When this occurs, it could of course be due to a number of different factors. One obvious possibility is malfunctioning of the enzyme glutamic acid decarbox-

ylase (GAD), and I would like to suggest that the GAD genes of autistic patients be examined for possible mutations. There are other excitation-inhibition pairings in the nervous system, however, and it seems likely that some autistic patients could be suffering from similar imbalances in those systems. This underlines the fact that autism is an affliction which can be caused by a range of different aetiologies rather than a single one.

The second prediction is related to something considerably more circumspect. Saermark and his colleagues (Saermark, Lebech, Bak & Sabers 1989, 1990) have recorded magnetoencephalographic (MEG) traces from a number of epilepsy patients, and have subsequently subjected these amplitude-time responses to a statistical analysis which reveals what is known as the attractor dimension (Amit 1989, Basar 1980). Although this is something of a simplification, one can regard the attractor dimension as indicating how many different regions of the brain are simultaneously interacting, during the conversion of sensory information into cognitive responses. For normal people, this dimension characteristically lies in the range 10-12. The suitably-processed traces of epileptic patients also display this dimension, but they comprise a second dimension as well. The fascinating thing about this second dimension is that it is invariably just half of the first (normal) value, that is to say 5-6. (These measurements are naturally not made during epileptic fits.)

It does not seem too farfetched to suppose that the appearance of that second value, in the case of epileptics, is due to incipient breakthrough, and that the tendency of the attractor dimension to collapse to half its normal value is one consequence. And in view of the circumstantial connection between epilepsy and autism, which I have been stressing in this paper, I am led to predict that MEG measurements on autistic patients might reveal a similar tendency toward two attractor dimensions, rather than one.

Conclusions

Although the main aim of this paper has been to present a new theory of autism, it has been necessary to include a discussion of intelligence, because the majority of autistic people have impaired levels of that mental attribute. And an adequate discussion of intelligence required, in turn, consideration of the broader issue of consciousness.

It was emphasized that one needs to differentiate between the (richly varied) products of consciousness and the underlying mechanism. The conclusion was that a description of that mechanism probably provides the best definition of consciousness presently available. Such a description was attempted, the resulting definition being that ***conscious awareness is possible if temporal changes in sensory input (or in internal patterns of signals) can be detected while the signals resulting from that input are still reverberating in the system, in a fully-functional form.*** Moreover, it was noted that ***such detection will permit the system to capture correlations between cause and effect so rapidly that their significance for the organism can immediately be appreciated, and also immediately used to modify the organism's repertoire of responses.*** The requirement that the reverberating signals still be in their fully-functional form guaranteed that they would be available

to produce the subjectively experienced concomitants known as qualia, these being mediated by the loops in the nervous system that run from the cortex to the various sub-cortical structures, and ultimately back to the cortex.

Subject to compliance with these requirements, an organism can be conscious, but the content of that consciousness will depend upon its intelligence. In accord with the discoveries of Broadbent (1958) regarding limitations in the capacity of the system, it was proposed that the level of intelligence of any individual is determined by the temporal and (generalized) spatial extents of its consciousness mechanism. The temporal factor was assumed to be related to the extent of very short memory, this being taken to last just a few seconds, while the (generalized) spatial factor was believed to be a (complicated) product of the attentional mechanism. This latter, in turn, was conjectured to be determined by what is happening in such sub-cortical structures as the pulvinar, the nucleus reticularis thalami and the brainstem reticular nucleus, as well as in the frontal and parietal regions of the cortex. Intelligence was perceived as being dependent on the performance of all these brain components, and the scope for individual differences was therefore seen as being considerable.

Similarly, some truncations of intelligence might be more important than others regarding an individual's ability to interact with the human environment. This was taken to be the case in those autistic people who lack a theory of mind, and who thus display the classic triad of autistic symptoms. But it was emphasized that these symptoms do not, in themselves, provide an explanation of the underlying trouble. It was suggested that the real cause of autism is the interference between neural signals known as breakthrough, which had been extensively studied in normal subjects. This was conjectured to cause what is essentially a short-circuiting of either the temporal or (generalized) spatial components of intelligence, or indeed both of these. At the molecular level, the breakthrough was postulated to be the consequence of an imbalance between excitation and inhibition at the neuronal level. In those autistic patients who show an amelioration of the symptoms when treated by vitamin B6, it was suggested that the syndrome might be caused by mutations in their genes for glutamic acid decarboxylase (GAD). It was further suggested that peculiarities in the magnetoencephalographic (MEG) characteristics of epileptic patients might also be seen in autistic patients.

Acknowledgements

I have enjoyed the benefit of discussions with Igor Alexander, Christen Bak, Claus Bundesen, Francis Crick, Andreas Engel, Hans Eysenck, Peter Foldiak, Uta Frith, Temple Grandin, Christof Koch, Merab Kokaia, Povl Krogsgaard-Larsen, Anders Lansner, Hans Liljenstrom, Sara Linse, Birgitte Michelsen, Margareta Mikkelsen, Claus Nielsen, Anne Nørremølle, Clara Claiborne Park, Steffen Petersen, Knud Saermark, Arne Schousboe, Hitomi Shibuya, John Taylor, Berthe Willumsen, and Semir Zeki, during the course of this work, and am happy to have this opportunity of formally thanking them.

Bibliography

Ackley, D.H., Hinton, G.E. and Sejnowski, T.J., 1985: A learning algorithm for Boltzmann machines. Cognitive Science, 9: 147-169.

Akshoomoff, N.A. and Courchesne, E., 1992: A new role for the cerebellum in cognitive operations. Behavioral Neuroscience 106: 731-738.

Amit, D.J., 1989: *Modeling Brain Function – The World of Attractor Neural Networks* (Cambridge University Press, Cambridge).

Arin, D.M., Bauman, M.L. and Kemper, T.L., 1991: The distribution of Purkinje cell loss in the cerebellum in autism. Neurology 41: (Supplement 1) 307.

Asperger, H., 1944: Die autistichen Psychopathen im Kindesalter. Arkiv für Psychiatrie und Nervenkrankheiten 117: 76-136.

Baddeley, A.D., Logie, R., Bressi, S., Della Sala, S. and H Spinnler, H., 1986: Quarterly Journal of Experimental Psychology 38A: 603.

Baddeley, A.D., 1986: *Working Memory* (Oxford University Press, Oxford).

Baddeley, A.D., 1990: *Human Memory: Theory and Practice* (Allyn and Bacon, Needham Heights, Mass).

Baddeley, A.D., 1992: Working Memory. Science 255: 556-559.

Baddely, A.D. and Hitch, G.J., 1974: G A Bower (ed) *The Psychology of Learning and Motivation* 8: 47-89.

Barlow, H.B., 1980: Nature's joke: A conjecture on the biological role of consciousness. B.D. Josephson and V.S. Ramachandran (eds) *Consciousness and the Physical World* (Pergamon, Oxford).

Barlow, H.B., 1985: Perception: what quantitative laws govern the acquisition of knowledge from the senses? C.W. Coen (ed) *Functions of the Brain* (Clarendon Press, Oxford).

Baron-Cohen, S., Leslie, A.M. and Frith, U., 1985: Does the autistic child have a "theory of mind"?. Cognition 21: 37-46.

Bartak, L. and Rutter, M., 1976: Differences between mentally retarded and normally intelligent autistic children. Journal of Autism and Childhood Schizophrenia 5: 109-120.

Basar, E., 1980: *EEG-Brain Dynamics* (Elsevier Biomedical Press, Amsterdam).

Bauman, M.L., 1991: Microscopic neuroanatomic abnormalities in autism. Pediatrics 87: 791-796.

Bauman, M.L. and Kemper, T.L., 1985: Histoanatomic observations of the brain in early infantile autism. Neurology 35: 866-874.

Becker, J.T., 1987: R.J. Wurtman, S.H. Corkin and J.H. Growdon (eds) *Alzheimer's Disease: Advances in Basic Research and Therapies* (Center for Brain Sciences and Metabolism Charitable Trust, Cambridge) 343-348.

Bennett, J., 1978: Some remarks about concepts. The Behavioral and Brain Sciences 4: 557-560.

Boden, M.A., 1990: *The Creative Mind – Myths and Mechanisms* (Abacus, London).

Bracke-Tolkmitt, R., Linden, A., Canavan, A.G.M., Rockstroh, B., Scholz, E., Wessel, K. and Diener, H.C., 1989: The cerebellum contributes to mental skills. Behavioral Neuroscience 103: 442-446.

Broadbent, D.E., 1958: *Perception and Communication* (Oxford University Press, Oxford).

Broadbent, D.E., 1982: Task combination and selective intake of information. Acta Psychologica 50: 253-290.

Bushnell, M.C., Goldberg, M.E. and Robinson, D.L., 1981: Behavioral enhancement of visual responses in the monkey cerebral cortex, I: Modulation in posterior parietal cortex related to selective visual attention. Journal of Neurophysiology 46: 755-772.

Cherry, E.C., 1953: Some experiments on the recognition of speech, with one and two ears. Journal of the Acoustical Society of America 25: 975-979.

Coltheart, M., 1983: Iconic memory. Philosophical Transactions of the Royal Society, London B 302: 283-294.

Cooper, J.R., Bloom, F.E. and Roth, R.H., 1991: *The Biochemical Basis of Neuropharmacology* (Oxford University Press, New York).

Cooper, L.A. and Shepard, R.N., 1973: W.G. Chase (ed) *Visual Information Processing* (Academic Press, New York) 75-176.

Cotterill, R.M.J. and Nielsen, C., 1991: A model for 40 Hz oscillations invokes inter-area interactions. NeuroReport 2:289-292.

Courchesne, E., 1987: A neurophysiological view of autism. E. Schopler and G.B. Mesibov (eds) *Neurobiological Issues in Autism* (Plenum Press, New York).

Courchesne, E., 1989: Neuroanatomical systems involved in infantile autism: The implications of cerebellar abnormalities. G. Dawson (ed) *Autism* (Guildford Press, New York).

Courchesne, E., 1991: Neuroanatomic imaging in autism. Pediatrics 87: 781-790.

Courchesne, E., Press, G.A. and Yeung-Courchesne, R., 1993: Parietal lobe abnormalities detected with MRI in patients with infantile autism. American Journal of Roentgenology 160: 387-393.

Crick, F., 1984: The function of the thalamic reticular complex: The searchlight hypothesis. Proceedings of the National Academy of Sciences, USA 81: 4586-4590.

Crick, F., 1988: *What Mad Pursuit: A Personal View of Scientific Discovery* (Basic Books, New York).

Crick, F., 1989: The recent excitement about neural networks. Nature 337: 129-132.

Crick, F. and Koch, C., 1990: Towards a neurobiological theory of consciousness. Seminars in Neurosciences 2: 263-275.

Crick, F. and Koch, C., 1991: Some reflections on visual awareness. Cold Spring Harbour Symposia on Quantitative Biology 55: 953-962.

Crick, F. and Koch, C., 1992: The problem of consciousness. Scientific American 267: 110-117.

Damasio, A.R., 1989: The brain binds entities and events by multiregional activation from convergence zones. Neural Computation 1: 123-132.

Daneman, M. and Carpenter, P.A., 1980: Journal of Verbal Learning and Verbal Behaviour 19: 450.

Decety, J., Sjöholm, H., Ryding, E., Stenberg, G. and Ingvar, D.H., 1990: The cerebellum participates in mental activity: tomographic measurements of regional cerebral blood flow. Brain Research 535: 313-317.

Dennett, D.C., 1978: Beliefs about beliefs. The Behavioral and Brain Sciences 4: 568-570.

Dennett, D.C., 1991: *Consciousness Explained* (Allen Lane, The Penguin Press, London).

Deutsch, D., 1992: Paradoxes of Musical Pitch. Scientific American 267: 70-75.

Deutsch, D., 1992: Some new pitch paradoxes and their implications. Philosophical Transactions of the Royal Society of London, Series B 336: 391-397.

Duncan, J., 1980: The demonstration of capacity limitation. Cognitive Psychology 12: 75-96.

Eckhorn, R., Bauer, R., Jordan, W., Brosch, M., Kruse, W., Munk, M. and Reitboeck, H.J., 1988: Coherent oscillations: a mechanism of feature linking in the visual cortex. Biological Cybernetics 60: 121-130.

Ellis, N.R., 1963: The stimulus trace and behavioral inadequacy. N.R. Ellis (ed) *Handbook of mental deficiency* (McGraw-Hill, New York) pp 134-158.

Ertl, J.P. and Schafer, E.W.P., 1969: Brain response correlates of psychometric intelligence. Nature 223: 421-422.

Fast, J., 1971: *Body Language* (Souvenir Press, London).

Felleman, D.J. and Van Essen, D.C., 1991: Distributed hierarchical processing in the primate cerebral cortex. *Cerebral Cortex* 1: 1-47.

Fish, G.S., Cohen, I.L., Wolf, E.G., Brown, W.T., Jenkins, E.C. and Gross, A., 1986: Autism and the fragile-X syndrome. American Journal of Psychiatry 143: 71-73.

Folstein, S. and Rutter, M., 1977: Infantile autism: a genetic study of 21 twin pairs. Journal of Child Psychology and Psychiatry 18: 297-321.

Freund, T.F., Martin, K.A.C., Soltesz, I., Somogyi, P. and Whitteridge, D.I., 1989: Arborization pattern and postsynaptic targets of physiologically identified thalamocortical afferents in striate cortex of macaque monkey. Journal of Comparative Neurology 289: 315-336.

Frith, U., 1989: *Autism: Explaining the Enigma* (Basil Blackwell, Oxford).

Frith, U., Morton, J. and Leslie, A.M., 1991:The cognitive basis of a biological disorder: autism. Trends in neurosciences 14: 433-438.

Fukurotani, K., 1982: Color information coding of horizontal-cell responses in fish retina. Color Research and Applications 7: 146-148.

Funahashi, S., Bruce, C.J. and Goldman-Rakic, P.S., 1989: Mnemonic coding of visual space in the monkey's dorsolateral prefrontal cortex. Journal of Neurophysiology 61: 331-349.

Gattass, R., Fiorani, M., Rosa, M.G.P., Pirion, M.C.G., Sousa, A.P.B. and Soares, J.G.M., 1992: Changes in receptive field size in V, in relation to perceptual completion. R. Lent (ed) *Visual system from genesis to maturity* (Birkhauser, Boston).

Geldard, F.A. and Sherrick, C.E., 1972: The cutaneous rabbit: A perceptual illusion. Science 178: 178-179.

Geldard, F.A. and Sherrick, C.E., 1983: The cutaneous saltatory area and its presumed neural base. Perception and Psychophysics 33: 299-304.

Geldard, F.A. and Sherrick, C.E., 1986: Space, Time and Touch. Scientific American 255: 85-89.

Georgopoulos, A.P., Schwartz, A.B. and Kettner, R.E., 1986: Neuronal population coding of movement direction. Science 233: 1416-1419.

Georgopoulos, A.P., Lurito, J.T., Petrides, M., Schwartz, A.B. and Massey, J.T., 1989: Mental rotation of the neuronal population vector. Science 243: 234-237.

Gilbert, C.D. and Wiesel, T.N., 1985: Intrinsic connectivity and receptive field properties in visual cortex. Vision research 25: 365-374.

Gilbert, C.D. and Wiesel, T.N., 1992: Receptive field dynamics in adult primary visual cortex. Nature 256: 150-152.

Gillberg, C., 1983: Identical triplets with infantile autism and the fragile-X syndrome. British Journal of Psychiatry 143: 256-260.

Gillberg, C. and Forsell, C., 1984: Childhood psychosis and neurofibromatosis – More than a coincidence?. Journal of Autism and Developmental Disorders 14: 1-8.

Gillberg, C., 1989: *Diagnosis and Treatment of Autism* (Plenum Press, New York).

Gillberg, C. and Coleman, M., 1992: *The Biology of the Autic Syndromes* (Cambridge University Press, Cambridge).

Goldman-Rakic, P.S., 1991: Prefrontal cortical dysfunction in schizophrenia: The relevance of working memory. B.J. Carroll and J.E. Barrett (eds) *Psychology and the Brain* (Raven Press, New York).

Goldman-Rakic, P.S., 1992: Working memory and the mind. Scientific American 267: 72-79.

Grandin, T. and Scariano, M., 1986: *Emergence: Labelled Autistic* (Arena Press, Navato, California).

Gray, C.M. and Singer, W., 1989: Stimulus-specific neuronal oscillations in orientation columns of cat visual cortex. Proceedings of the National Academy of Sciences, USA 86: 1698-1702.

Gross, C.G., Bender, D.B. and Rocha-Miranda, C.E., 1969: Visual receptive fields of neurons in inferotemperal cortex of the monkey. Science 166: 1303-1306.

Harman, G., 1978: Studying the chimpanzee's theory of mind. The Behavioral and Brain Sciences 4: 576-577.

Heath, R.G., Dempesy, C.W., Fontana, C.J. and Fitzjarrell, A.T., 1980: Feedback loop between cerebellum and septal-hippocampal sites: Its role in emotion and epilepsy. Biological Psychiatry 15: 541-556.

Hendry, S.H.C. and Jones, E.G., 1989: Synaptic organization of GABA and GABA/tachykinin immunoreactive neurons in layer IVC-beta of monkey area 17. Society of Neurosciences Abstracts 14: 1123.

Heh, C.W.C., Smith, R., Wu, J., Hazlett, E., Russell, A., Asarnow, R., Tanguay, P. and Buchsbaum, M.S., 1989: Positron emission tomography of the cerebellum in autism. American Journal of Psychiatry 146: 242-245.

Ivry, R.B. and Baldo, J.V., 1992: Is the cerebellum involved in learning and cognition? Current Opinion in Neurobiology 2: 212-216.

James, W., 1910: *The Principles of Psychology* (Macmillan, London).

Jones, E.G., 1993: GABAergic neurons and their role in cortical plasticity in primates. Cerebral Cortex 3: 361-372.

Julesz, B., 1981: Textons, the elements of texture perception, and their interactions. Nature 290: 91-97.

Kanner, L., 1943: Autistic disturbances of affective contact. Nervous Child 2: 217-250.

Kaufman, B.N., 1976: *Son-Rise* (Harper and Row, New York).

Kolvin, I., Ounsted, C. and Roth, M., 1971: Studies in the childhood psychoses. V. Cerebral dysfunction and childhood psychoses. British Journal of Psychiatry 118: 407-414.

Kyllonen, P.C. and Christal, R.E., 1990: Intelligence 14: 389.

Lee, C., Rohrer, W.H. and Sparks, D.L., 1988: Population coding of saccadic eye movements by neurons in the superior colliculus. Nature 332: 357-360.

Leiner, H.C., Leiner, A.L. and Dow, R.S., 1986: Does the cerebellum contribute to mental skills. Behavioral Neuroscience 100: 443-454.

Lelord, G., Muh, J.P., Barthelmy, C., Martineau, J., Garreau, B. and Callaway, E., 1981: Effects of pyridoxine and magnesium on autistic symptoms – Initial observations. Journal of Autism and Developmental Disorders 11: 219-230.

Leslie, A.M., 1987: Pretense and representation: The origins of the "Theory of Mind". Psychological Review 94: 412-426.

Leslie, A.M. and Frith, U., 1988: Autistic children's understanding of seeing, knowing and believing. British Journal of Developmental Psychology 4: 315-324.

Lewis, D.K., 1969: *Convention: A Philosophical Study* (Harvard University Press, Cambridge).

Libet, B., Wright Jr, E.W., Feinstein, B. and Pearl, D.K., 1979: Subjective referral of the timing for a conscious sensory experience: A functional role for the somatosensory specific projection system in man. Brain 102: 193-224.

Libet, B., 1982: Brain stimulation in the study of neuronal functions for conscious sensory experience. Human Neurobiology 1: 235-242.

Libet, B., 1989: Conscious subjective experience vs. unconscious mental functions: Theory of the cerebral processes involved. R.M.J. Cotterill (ed) *Models of Brain Function* (Cambridge University Press, Cambridge).

Llinás, R.R. and Pellionisz, A., 1979: Brain modeling by tensor network theory and computer simulation. The cerebellum: distributed processor for predictive coordination. Neuroscience 4: 322-348.

Martin, J.P. and Bell, J., 1943: A pedigree of mental defect showing sex-linkage. Journal of Neurology and Psychiatry 6: 154-157.

Martineau, J., Barthelmy, C., Garreau, B. and Lelord, G., 1985: Vitamin B6, Magnesium, and Combined B6-Mg: Therapeutic effects in childhood autism. Biological Psychiatry 20: 467-478.

Martineau, J., Barthelmy, C. and Lelord, G., 1986: Long-term effects of combined vitamin B6-magnesium administration in an autistic child. Biological Psychiatry 21: 511-518.

Martineau, J., Barthelmy, C., Cheliakine, C. and Lelord, G., 1988: Brief report: An open middle-term study of combined vitamin B6-magnesium in a subgroup of autistic children selected on their sensitivity to this treatment. Journal of Autism and Developmental Disorders 18: 435-447.

McCulloch, W.S. and Pitts, W., 1943: A logical calculus of the ideas immanent in nervous activity. Bulletin of Mathematical Biophysics 5: 115-133.

McFarland, D., 1966: On the causal and functional significance of displacement activities. Zeitschrift für Tierpsychologi 23: 217-235.

McGinn, C., 1991: *The Problem of Consciousness* (Basil Blackwell, Oxford).

McWhirter, N. and McWhirter, R., (eds), 1963: *The Guinness Book of World Records* (Bantam Books, New York).

Meldrum, B.S., 1975: Epilepsy and gamma-aminobutyric acid-mediated inhibition. International Reviews of Neurobiology 17: 1-36.

Miller, G.A., 1956: The magical number seven, plus or minus two: Some limits on our capacity for processing information. The Psychological Review 63: 81-97.

Milner, P.M., 1974: A model for visual shape recognition. Psychological Review 81: 521-535.

Minshew, N.J., 1991: Indices of neural function in autism: Clinical and biologic implications. Pediatrics 87: 774-780.

Minsky, M.L. and Papert, S.A., 1969: *Perceptrons: An Introduction to Computational Geometry* (MIT Press, Cambridge, Mass).

Mitrofanis, J. and Guillery, R.W., 1993: New views of the thalamic reticular nucleus in the adult and the developing brain. Trends in Neurosciences 16: 240-245.

Moran, J. and Desimone, R., 1985: Selective Attention Gates Visual Processing in the Extrastriate Cortex. Science 229: 782-784.

Moray, N., 1960: Broadbent's filter theory: Postulate H and the problem of switching time. Quarterly Journal of Experimental Psychology 12: 214-220.

Mountcastle, V.B., 1978: Brain mechanisms for directed attention (Sherrington Memorial Lecture). Journal of the Royal Society of Medicine 71: 14-28.

Murakami, J.W., Courchesne, E., Press, G.A., Yeung-Courchesne, R. and Hesselink, J.R., 1989: Reduced cerebellar hemisphere size and its relationship to vermal hypoplasia in autism. Archives of Neurology 46: 689-694.

Nagel, T., 1974: What is it like to be a bat? Philosophical Review 83: 435-445.

Naruse, H. and Ornitz, E.M., 1992: *Neurobiology of Infantile Autism* (Excerpta Medica, Amsterdam).

Oatley, K. and Johnson-Laird, P.N., 1987: Towards a cognitive theory of emotion. Cognition and Emotion 1: 29-50.

Ornitz, E.M., 1983: The functional neuroanatomy of infantile autism. International Journal of Neuroscience 19: 85-124.

Ornitz, E.M., 1987: Autism. G.Adelman (ed) *Encyclopedia of Neuroscience* (Birkhäuser, Boston) 92-93.

Park, C.C., 1967: *The Siege* (Little Brown, Boston).

Perky, C.W., 1910: An experimental study of imagination. *American Journal of Psychology* 21: 422-452.

Perret, D.R., Rolls, E.T. and Caan, W., 1982: Visual neurones responsive to faces in the monkey temporal cortex. Experimental Brain Research 47: 329-342.

Polich, J., Howard, L. and Starr, A., 1983: P300 latency correlates with digit span. Psychophysiology 20: 665-669.

Posner, M.I., Walker, J.A., Friedrich, F.J. and Rafal, R.D., 1984: Effects of parietal injury on covert orienting of attention. Journal of Neuroscience 4: 1863-1874.

Posner, M.I. and Driver, J., 1992: The neurobiology of selective attention. Current Opinion in Neurobiology 2: 165-169.

Premak, D. and Woodruff, G., 1978: Does the chimpanzee have a theory of mind? The Behavioral and Brain Sciences 4: 515-526.

Ramachandran, V.S. and Gregory, R.L., 1991: Perceptual filling in of artificially induced scotomas in human vision. Nature 350: 699-702.

Ramachandran, V.S., 1992a: Filling in the blind spot. Nature 356:115.

Ramachandran, V.S., 1992b: Blind spots. Scientific American 266: 44-49.

Ramachandran, V.S., 1993: Filling in Gaps in Logic: Some Comments on Dennett. Consciousness and Cognition 2: 165-168.

Rimland, B. 1964: *Infantile Autism: the syndrome and its implications for a neural theory of behavior* (Appleton-Century-Crofts, New York).

Rimland, B., 1987: Vitamin B6 (and magnesium) in the treatment of autism. Autism Research Review International 1: 3.

Ritvo, E.R., Spence, M.A., Freeman, B.J., Mason-Brothers, A., Mo, A. and Marazita, M.L., 1985: Evidence for autosomal recessive inheritance in 46 families with multiple incidences of autism. American Journal of Psychiatry 142: 187-192.

Ritvo, E.R., Freemen, B.J., Mason-Brothers, A., Mo, A. and Ritvo, A.M., 1985: Concordance for the syndrome of autism in 40 pairs of afflicted twins. American Journal of Psychiatry 142: 74-77.

Ritvo, E.R., Freemen, B.J., Scheibel, A.B., Duong, T., Robinson, H., Guthrie D. and Ritvo, A., 1986: Lower Purkinje cell counts in the cerebella of four autistic subjects: Initial findings of the UCLA-NSAC autopsy research report. American Journal of Psychiatry 143: 862-866.

Robinson, D.L. and Petersen, S.E., 1992: The pulvinar and visual salience. Trends in Neurosciences 15: 127-132.

Rolls, E.T., 1992: Neurophysiological mechanisms underlying face processing within and beyond the temporal cortical visual areas. Philosophical Transactions of the Royal Society, London B 335: 11-21.

Rubin, E., 1915: *Synsoplevede Figurer: Studier i psykologisk Analyse* (Nordisk Forlag, Copenhagen).

Rumelhart, D.E., Hinton, G.E. and Williams, R.J., 1986: Learning internal representations by error propagation. D E Rumelhart and J L McClelland (eds) *Parallel Distributed Processing: Explorations in the Microstructure of Cognition* (MIT Press, Cambridge, Mass., 1986) 1: 318-364.

Rutter, M., 1978: Diagnosis and definition of childhood autism. Journal of Autism and Childhood Schizophrenia 8: 139-161.

Rutter, M., 1984: Autistic children growing up. Developmental Medicine and Child Neurology 26: 122-129.

Saermark, K., Lebech, J., Bak, C.K. and Sabers, A. 1989: Magnetoencephalography and attractor dimension : Normal subjects and epileptic patients. E.Basar and T.H.Bullock (eds) *Brain Dynamics – Progress and Perspectives* (Springer-Verlag, Berlin).

Saermark, K., Lebech, J., Bak, C.K. and Sabers, A. 1990: Magnetoencephalography and attractor dimension : Normal subjects and epileptic patients. E.Basar (ed) *Chaos in Brain Function* (Springer-Verlag, Berlin).

Sakrison, D.J., 1968: *Communication Theory: Transmission of Waveforms and Digital Information* (Wiley, New York).

Schopler, E. and Mesibov, G.B., 1987: *Neurobiological Issues in Autism* (Plenum Press, New York).

Searle, J.R., 1980: Minds, Brains and Programs. The Behavioral and Brain Sciences 3: 417-457.

Searle, J.R., 1984: *Minds, Brains and Science* (Penguin Books, London). 1984 BBC Reith Lectures.

Searle, J.R., 1992: *The Rediscovery of the Mind.* (MIT Press, Cambridge).

Segal, S.J. and Fusella, V., 1970: Influence of imaged pictures and sounds on detection of visual and auditory signals. Journal of Experimental Psychology 83: 458-464.

Shallice, T. and Warrington, E.K., 1970: Independent functioning of verbal memory stores -a neuropsychological study. Quarterly Journal of Experimental Psychology 22: 261.

Shepard, R.N., 1964: Circularity in judgments of relative pitch. Journal of the Acoustical Society of America 36: 2346-2353.

Shepard, R.N. and Metzler, J., 1971: The mental rotation of three-dimensional objects. Science 171: 701-703.

Shibuya, H. and Bundesen, C., 1988: Visual selection from multielement displays: measuring and modeling effects of exposure duration. Journal of Experimental Psychology 14: 591-600.

Singer, W., 1977: Control of thalamic transmission by corticofugal and ascending reticular pathways in the visual system. Physiological Review 57: 386-420.

Smalley, S.L., Asarnow, R.F. and Spence, M.A., 1988: Autism and Genetics. Archives of General Psychiatry 45: 953-961.

Squire, L., 1987: *Memory and Brain* (Oxford University Press, Oxford).

Steriade, M. and Llinás, R.R., 1988: The functional states of the thalamus associated neuronal interplay. Physiological Reviews 68: 649-742.

Steriade, M., Jones, E.G. and Llinás, R.R. 1990: *Thalamic oscillations and signaling* (Wiley, New York).

Stroop, J.R., 1935: Studies of interference in serial verbal reactions. Journal of Experimental Psychology 18: 643-662.

Stryer, L., 1981: *Biochemistry* (Freeman, San Francisco).

Treffert, D., 1989: *Extraordinary People* (Bantam, London).

Treisman, A.M., 1960: Contextual cues in selective listening. Quarterly Journal of Experimental Psychology 12: 242-248.

Treisman, A.M., 1964: Verbal cues, language, and meaning in selective attention. American Journal of Psychology 77: 206-219.

Treisman, A.M. and Gelade, G., 1980: A feature-integration theory of attention. Cognitive Psychology 12: 97-136.

Troscianko, T., 1987: Colour Vision: Brain Mechanisms. R.L. Gregory (ed) *The Oxford Companion to the Mind* (Oxford University Press, Oxford) 150-152.

Van Essen, D.C., Anderson, C.H. and Felleman, D.J., 1992: Information processing in the primate visual system: An integrated systems perspective. Science 255: 419-423.

Velmans, M., 1991: Consciousness from a first-person perspective. Behavioral Brain Science 14: 702-726.

Volkmar, F.R. and Nelson, D.S., 1990: Seizure disorders in autism. Journal of the Academy of Child Adolescent Psychiatry 1: 127-129.

von der Malsburg, C., 1981: The correlation theory of brain function. Internal Report Number 81-2. Max-Planck Institute for Biophysical Chemistry, Göttingen.

Wallesch, C.W. and Horn, A., 1990: Long-term effects of cerebellar pathology on cognitive functions. Brain and Cognition 14: 19-25.

Williams, D., 1992: *Nobody Nowhere: The remarkable autobiography of an autistic girl* (Doubleday, London).

Wimmer, H. and Perner, J., 1983: Beliefs about beliefs: representation and constraining function of wrong beliefs in young children's understanding of perception. Cognition 21: 103-128.

Wing, L., 1971: *Autistic children: a guide for parents* (Constable, London)

Wing, L. and Gould, J., 1978: Systematic recording of behaviour and skills of retarded and psychotic children. Journal of Autism and Childhood Schizophrenia 8: 79-97.

Wing, L. and Gould, J., 1979: Severe impairments of social interaction and associated abnormalities in children: epidemiology and classification. Journal of Autism and Developmental Disorders 9: 11-30.

Wing, L., 1981: Language, social, and cognitive impairments in autism and severe mental retardation. Journal of Autism and Developmental Disorders 11: 31-44.

Wurtz, R.H., Goldberg, M.E. and Robinson, D.L., 1980: Behavioral modulation of visual responses in the monkey: Stimulus selection for attention and movement. Progress in Psychobiology and Physiological Psychology 9: 43-83.

Wurtz, R.H., Goldberg, M.E. and Robinson, D.L., 1982: Brain mechanisms of visual attention. Scientific American 246: 124-132.

Yarbus, A.L., 1967: B.Haigh (trans) *Eye Movements Vision* (Plenum, New York).

Zeki, S. and Shipp, S., 1988: The functional logic of cortical connections. Nature 335: 311-317.

Zeki, S., 1992: The visual image in mind and brain. Scientific American 267: 42-50.

Zipser, D. and Andersen, R.A., 1988: A back-propagation programmed network that simulates response properties of a subset of posterior parietal neurons. Nature 331: 679-684.

Submitted to the Academy February 1994.
Published April 1994.

Title. – Titles should be kept as short as possible and with an emphasis on words useful for indexing and information retrieval.

Abstract, Summary. – An abstract in English is compulsory. It should count 10-15 lines, outline main features, stress novel information and conclusions, and end with the author's name, title, and institutional and/or private postal address. – Papers in Danish may be provided with a summary in another language by agreement between author and Editor.

Typescript. – Page 1 should contain title, author's name and the name of the Academy. Page 2: Abstract, author's name and address. Page 3: Table of contents if necessary. Captions should be supplied on separate sheets. Footnotes should be avoided if at all possible; if indispensable, they, too, should be typed on separate sheets. Consult a *recent* issue of the series for general layout.

Typewrite with double space throughout and leave a 4 cm margin *right.* Indicate desired position of illustrations and tables with pencil in margin *and repeat it in the galley proof.*

Use three or fewer grades of heading unless more are indispensable. Avoid long headings. Indicate clearly the hierarchy of headings.

Figures. – Please submit two copies of each graph, map, photograph, etc., all marked with the author's name. Whenever possible all figures will be placed within the text; the nature of the illustrations will govern the Editor's choice of paper quality.

All figures, also line drawings, must be submitted as glossy, photographic prints suitable for direct reproduction. Prints fitting the indicated printed area are preferred, but the final size is the responsibility of the Editor. The scale should be indicated in the caption or, preferably, on the illustration itself.

Fold-out figures and tables should be avoided. Use distinct (but not dominant) capital letters for the items in composite figures. For transfer lettering use simple, semi-bold typefaces. The size of the smallest letters should not be less than 1.5 mm. Intricate tables are often more easily reproduced from line-drawings or from technically perfect original computer or type processor output.

References. – In general, the Editor expects all references to be formally consistent and in accordance with accepted practice within the particular field of research. Bibliographical references should preferably be given as, e.g., Shergold 1975, 16, the latter figure indicating the page number unless misunderstandable.

Correspondence

Manuscripts should be sent to the Editor, Det Kongelige Danske Videnskabernes Selskab, H. C. Andersens Boulevard 35, DK-1553, Copenhagen V, Denmark (tlf. +45 33 11 32 40). Questions concerning subscription to the series should be directed to the publishers.

Publisher

Munksgaard Export and Subscription Service
Nørre Søgade 35, DK-1370 Copenhagen K, Denmark

Editor: Poul Lindegård Hjorth

© (Year). Det Kongelige Danske Videnskabernes Selskab. All rights reserved. No part of this publication may be reproduced in any form without the written permission of the copyright owner.

Biologiske Skrifter

Biol. Skr. Dan. Vid. Selsk.

Priser excl. moms/Prices abroad in Danish Crowns

Vol.

25 Six Papers in the Biological Sciences, being Part Two of *Sixteen Research Reports* by the Niels Bohr Fellows of the Royal Danish Academy of Sciences and Letters, published on the Occasion of the Centenary of Niels Bohr. 1985 ..... 200.-
(*Sixteen Research Reports* Part One is identical with: Ten Papers in the Exact Sciences and Geology, *Matematisk-fysiske Meddelelser* 41. 1985, 400.-)

26 JENSEN, HANS ARNE: Seeds and other Diaspores in Soil Samples from Danish Town and Monastery Excavations, dated 700-1536 AD. 1986 ..... 200.-

27 NILSSON, JYTTE R.: The African Heterotrich Ciliate, *Stentor andreseni* sp.nov., and *S. amethystinus* Leidy. A Comparative Ultrastructural Study. 1986 ..... 100.-

28 WUNDERLIN, RICHARD; LARSEN, KAI; and LARSEN, SUPEE SAKSUWAN: Reorganization of the *Cercideae (Fabaceae: Caesalpinioideae)*. 1987 ..... 80.-

29 JENSEN, HANS ARNE: Macrofossils and their Contribution to the History of the Spermatophyte Flora in Southern Scandinavia from 13000 BP to 1536 AD. 1987 ..... 200.-

30 DYCK, JAN: Structure and Light Reflection of Green Feathers of Fruit Doves *(Ptilinopus spp.)* and an Imperial Pigeon *(Ducula concinna)*. 1987..... 100.-

31 FRIIS, ELSE MARIE; CRANE, PETER R.; PEDERSEN, KAJ RAUNSGAARD: Reproductive Structures of Cretaceous Platanaceae. 1988 ..... 100.-

32 WINGSTRAND, K. G.: Comparative Spermatology of the Crustacea Entomostraca 2. Subclass Ostracoda. 1988..... 250.-

33 MIKKELSEN, VALD. M.: The Commons of Rejnstrup, Denmark. 1989 ..... 50.-

34 ØLLGAARD, BENJAMIN: Index of the *Lycopodiaceae*. 1989..... 150.-

35 SRINIVASAN, VIJAYALAKSHMI; FRIIS, ELSE MARIE: Taxodiaceous conifers from the Upper Cretaceous of Sweden. 1989 ..... 100.-

36 FRIIS, ELSE MARIE: *Silvianthemum suecicum* gen. et sp. nov., a new saxifragalean flower from the Late Cretaceous of Sweden. 1990..... 70.-

37 MOESTRUP, ØJVIND; THOMSEN, HELGE A.: *Dictyocha speculum* (Silicoflagellata, Dictyochophyceae), studies on armoured and unarmoured stages. 1990..... 100.-

38 MIKKELSEN, VALD. M.: Borrelyngen on Bornholm, Denmark. 1990..... 100.-

39 JØRGENSEN, C. BARKER: Water Economy in the Life of a Terrestrial Anuran, the Toad *Bufo bufo*. 1991..... 70.-

40 HANSEN, MICHAEL: The Hydrophiloid Beetles. Phylogeny, Classification and a Revision of the Genera (Coleoptera, Hydrophiloidea). 1991 ..... 700.-

41 FRIIS, ELSE MARIE; PEDERSEN, KAJ RAUNSGAARD; CRANE, PETER R.: *Esqueiria* gen. nov., fossil flowers with combretaceous features from the Late Cretaceous of Portugal. 1992 ..... 90.-

42 HENNINGSEN, KNUD W.; BOYNTON, JOHN E.; WETTSTEIN, DITER VON: Mutants at *xantha* and *albina* Loci in Relation to Chloroplast Biogenesis in Barley (*Hordeum vulgare* L.). 1993..... 700.-

43 Brain and Mind. Symposium on the Occasion of the 250th Anniversary of The Royal Danish Academy of Sciences and Letters August 17-20, 1992. Ed. by RODNEY M. J. COTTERILL. In proof.

44 MIKKELSEN, VALD. M.: Borrelyngen on Bornholm, Denmark. 2. 1994..... 100.-

45 COTTERILL, RODNEY M. J.: Autism, Intelligence and Consciousness. 1994..... 150,-

Printed in Denmark by Special-Trykkeriet Viborg a-s. ISSN 0366-3612. ISBN 87-7304-254-4