Why we need Cognition: Cause and developmental disorder.

John Morton and Uta Frith

Institute of Cognitive Neuroscience, University College, London

Introduction

Our main interest here is in the nature of theories of developmental disorders and why we need to talk about the cognitive level of description. We will start by illustrating briefly why we cannot rely on purely behavioural descriptions of developmental disorders. We expand on these illustrations by showing an illustrative causal model - in this case for the "phonological deficit" model of developmental dyslexia. In this we use a notation which is neutral with respect to the theory itself, and makes explict use of three levels of description: brain, cognition and behaviour. Having illustrated the possible causal relationship between elements at the brain and cognition levels, we create a contrast by referring to two other relationships between these two levels. These we term *equivalence* and *correlation*. By equivalence we refer to elements at brain and cognitive levels which have a one-to-one relationship. By correlation, we refer to the fact that all cognitive states correspond to brain states. For example, a change in belief leads to a change in brain state but without a one-to-one equivalence. Specification of the three relationships between brain and cognitive descriptions helps in understanding the continuing need for Cognition.

Why we cannot rely on behaviour

Is it possible to adequately describe developmental disorders on the basis of behaviour alone? Symptoms (technically, what the individual in question reports about their condition) are essentially cognitive; signs (the manifestation of the disorder) usually reflect underlying problems rather than being the problem themselves. The reason for this, quite simply, is that the same behaviour can manifest itself for a wide variety of reasons and different behaviour in different people can arise for essentially the same reason. We will give one example where we should mistrust failure on a behavioural test as a guide to underlying deficits and another where the issue is success.

Take the case of a 9 year old boy with a reading age of 6. The inadmissible interpretation, without further evidence, is that of specific learning difficulty (i.e. specific brain damage leading to dyslexia). The obvious alternative explanations are that the boy has undiagnosed visual problems (correctable by glasses); that the child has undiagnosed hearing impairment which led to delay in language acquisition; that he is trying to read in a second language; that he has not attended school regularly. To clarify the diagnosis and establish the state of his cognitive processes, systematic investigations are required.

Our second example is that of a 12-year-old girl with autism who succeeds on the standard false belief task, which, according to our current theory of autism, she ought to fail. The inadmissible interpretation without further evidence would be that she is no longer autistic. She might have learned the test; she might have built up enough experience in the world to be

able to reason by analogy (she would need to have good general information processing capacity). To clarify the true extent of her disability/ability systematic investigations are required so that the state of her cognitive processes can be established.

What these two examples have in common is that they illustrate the fact that the same behaviour could arise from different cognitive preconditions. The second way of illustrating the need for the cognitive level over and above the behavioural level, in contrast, lies in the way that a particular underlying condition manifests itself differently in different individuals. We want to say that these individuals are similar in important ways, and so give them the same label. The heterogeneity of expression of all developmental disorders is legendary.

Cognition, we have argued, plays an essential role in bridging the gap between brain and behaviour. The mapping between brain and behaviour is chaotic, unless governed by principled hypotheses which are most sharply and parsimoniously expressed in terms of mental concepts. Our work over the years on both autism and dyslexia has shown how it is possible to simplify a mass of apparently unconnected observations from both biological and behavioural sources, by postulating cognitive defects.

Example of a causal model

Over the past few years, we have been concerned to establish the relationships among biological, cognitive, behavioural and environmental factors in developmental disorders (Frith, Morton, & Leslie, 1991; Morton & Frith, 1993a,b, 1995). We have done this through a framework - Causal Modelling - which is neutral with respect to any particular theory. This neutrality arises from the fact that we can represent our own causal theories or anyone else's equally well. Indeed, the only theoretical point intrinsic to the framework is the necessity of the cognitive level of representation. The point is best made with an illustration.

Figure 1 is taken from British Psychological Society, (1999). This diagram represents what might be called the familiar canonical phonological deficit view of dyslexia. In the diagram, the arrows represent statements of the kind "X causes Y" or "X leads to Y" in the theory which is being represented. Note that you don't have to agree with the underlying theory to understand what the notation is intended to accomplish. The first claim of the theory is that dyslexia is a genetic disorder. This

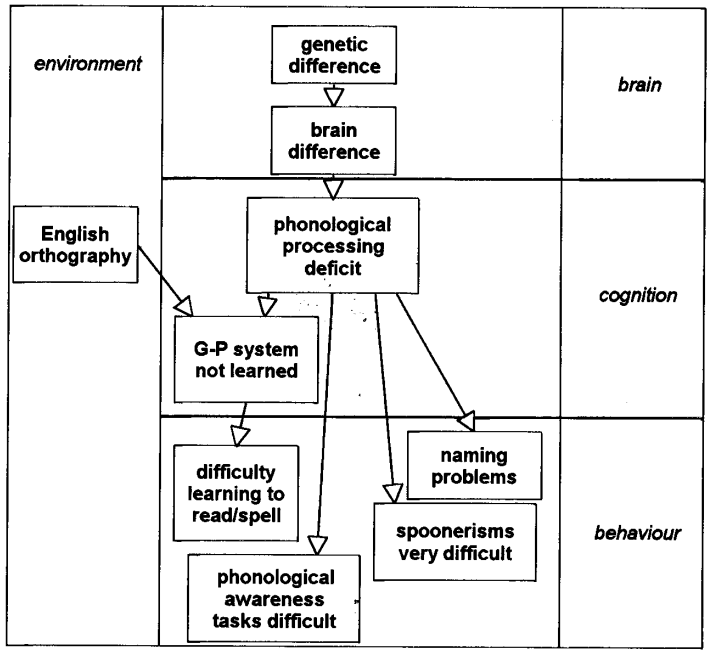

Figure 1 - the canonical phonological deficit theory of dyslexia, from British Psychological Society (1999), the arrows showing the causal relations among three levels of description - brain, cognition and behaviour

contrasts with the traditional, behavioural, view whereby dyslexia was defined as a discrepancy between mental age and reading age. We will enlarge later on the limitations of the latter viewpoint. The theory claims that there is a subset of the population which is different genetically from the rest in a particular (as yet unspecified) regard. Over the course of early development, this genetic difference leads to a difference in brain structure. The difference in brain structure, in turn, leads to a deficit in phonological processing which is revealed, in terms of behaviour, through problems in naming. In the laboratory, we also find difficulty with phonological awareness tasks and with generating Spoonerisms. By default, the theory says that these factors will be found in all environments and cultures. Additionally, the theory says that within the English-speaking culture there will be a further manifest problem - difficulty in learning to read. This problem arises because of the nature of English orthography. Specifically, the theory claims that English orthography, in conjunction with the phonological processing deficit arising from the genetic difference, makes it difficult to set up a Grapheme-Phoneme mapping system.

Note that there are indirect claims made by the theory which are quite clear from the causal model. The most important is that in the presence of a more transparent orthography than English, there will be no reading difficulty. Now, while this claim might be slightly exaggerated, it is certainly the case that dyslexia, as indexed by manifest difficulty in learning to read, is rare in Italy, where the orthography is much more transparent and consistent than English (Lindgren, De Renzi, & Richman,1985). However, unless the gene pool in Italy is unexpectedly different from that in the UK, the theory in Figure 1 leads to the prediction that there will be Italians with the same genetic difference who will have problems in naming, difficulty with phonological tasks and with generating Spoonerisms. Since the problems in naming would not have been particularly noticeable, this group of people would not have been diagnosed as having a developmental disorder, nor would they be aware of their problem. Current work suggests that Such a group has been identified. Perhaps 5% of the Italian population have the same genetic condition as the English dyslexic population, but without manifesting any significant reading problem. The reason for this is that Italian orthography is regular, and lacks the extremely complex context-sensitive mapping with characterises English orthography. Apparently, such complexity in orthographic-phonological mapping is necessary for the particular genetic configuration to manifest itself in dyslexia as characterised in reading behaviour. In terms of figure 1, the Italian group would be identical to the English apart from the environmental link from the orthography. Without that, the phonological deficit does not lead to a difficulty in learning to read. We could, then, view Italian culture as a therapeutic factor! The point of this illustration is that without a cognitive definition of dyslexia such a phenomenon could not have been predicted, and, if it had been stumbled upon, could not have been explained. The irrelevance of the discrepancy definition of dyslexia to a complete understanding of the condition should be apparent.

Sub-Levels of description in the brain. In figure 1, we go from gene to brain state without defining the latter. Gallese, (in press) made distinctions among a number of levels in the biological domain within which he embeds his work, which is mainly at the level of neurons. These levels are:

membrane proteins
receptors

neurons and synapses
neural assemblies
cortical circuits
behaviour/cognition

He discussed the possible independence of these levels, pointing out that while complexity at the receptor level is not relevant for behaviour/cognition, action potentials, arising from activity in neurons, can have such meaning. Assemblies of neurons are even more likely to have meaning. Functional imaging studies currently relate to cortical circuits. We can note that a full causal model of developmental disorder could start with a problem at the membrane protein level (preceded by a genetic disorder, perhaps) and move down through the five biological levels before leading to a cognitive problem.

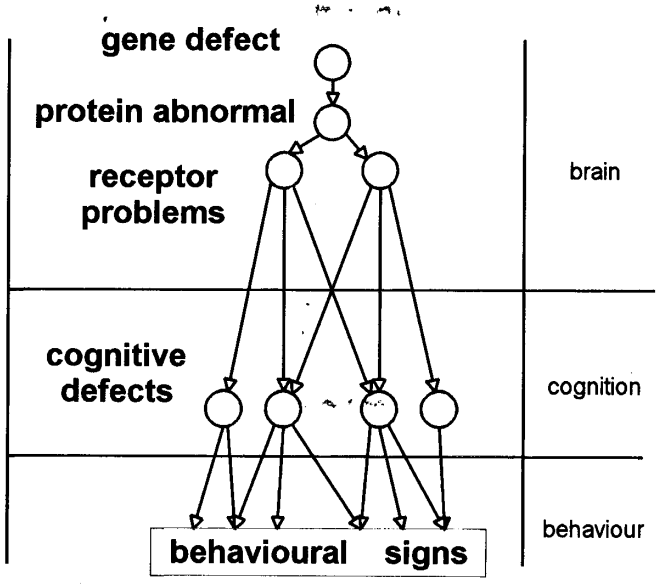

Figure 2 - sketch of a generic causal model of a developmental disorder with biological origin. For reasons of simplicity, we omit the environmental factors from here.

Three relations between the cognitive and biological levels

In our previous writing on causal modelling, causal arrows have always gone from the biological to the cognitive level. This is because the developmental disorder that we wish to explain are known to have a biological origin before birth. There could be other disorders which may be best explained by postulating a *causal* influence from the cognitive to the biological level.[1] In contrast to the causal relationship between biological and cognitive levels, we need to discuss two other relationships which we will call *equivalence* and *realisation*. To appreciate these distinctions we can start with the causal model sketched in figure 2. We have no problem in saying that the gene defect *causes* the various cognitive defects and behavioural signs. Equally it is straightforward to say that receptor problems (assuming that they are specified) *cause* a cognitive defect - such as, in autism, the lack of a Theory of Mind Mechanism (Leslie, 1985). However, it is clear that there is an explanatory gap between these two elements which would have to be filled before the theory could be considered satisfactory. To start with, the receptor problem could lead to a defect at the level of neurons and synapses, which could, in turn, lead developmentally to an abnormality best described in terms of neural assemblies. These levels would have to be specified causally.

Let us now look at the cognitive level using the example of autism. The lack of a Theory of Mind Mechanism, in the case of autistic children, has been attributed in one version of the theory to the absence of a computational device called an Expression Raiser (or EXPRAIS). Morton (1986) supposed EXPRAIS to be a cognitive primitive - that is, EXPRAIS is

[1] It is clearly the case that all cognitive events are accompanied by changes in brain state. This is not what is normally meant by cause.

supposed to be irreducible. What this means is that there is no other *cognitive* element above it in the causal chain accounting for its absence in this particular theory of autism.[2] The final step in the argument is to suppose that such cognitive primitives are invariably instantiated in the brain in exactly the same way, and that there is a one-to-one mapping. A particular cognitive primitive is always instantiated by a particular, identifiable neuron, type of synapse, or assembly of neurons, N. Further, each occurrence of activity in N can be taken as evidence for the operation of this cognitive primitive.

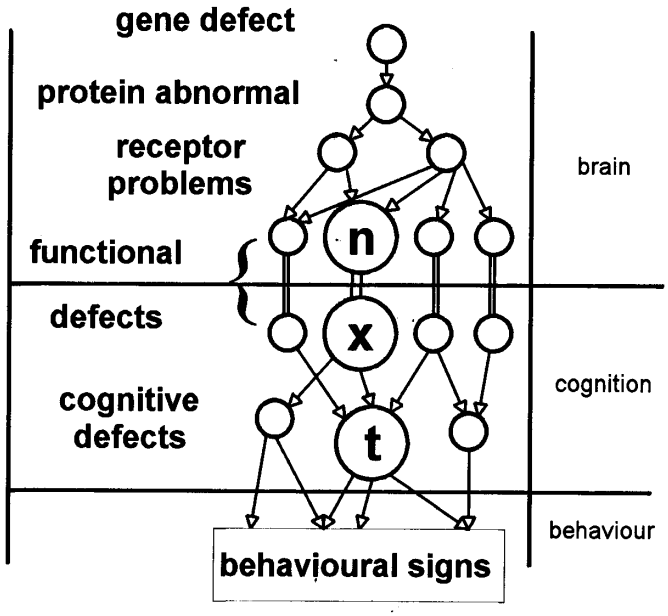

Figure 3 - an expansion of the model in figure 2, showing the equivalence relation between elements at the biological and cognitive levels. n is the neural equivalent to x (EXPRAIS), while t represents the Theory of Mind Mechanism.

1. ***equivalence*** - this is the sense in which elements of the mind and elements of the brain are identical. We apply the term to the relationship between a cognitive primitive and its neural instantiation - the mind-brain units.

equivalence: cognition to brain. On our interpretation of the Leslie theory, EXPRAIS, as a cognitive primitive, will have a simple equivalent in neural circuitry. Once we have identified it, this neural circuitry will be given a name. This name would be equivalent to EXPRAIS because it would refer to the identical element. We can now contrast the equivalence relationship with the causal relationship in a causal model. In figure 3, ***x*** represents EXPRAIS and ***n*** is the equivalent neural circuitry while ***t*** represents the Theory of Mind Mechanism. We could equally say that a neural deficit ***(n)*** causes a deficit in the Theory of Mind Mechanism or say that a deficit in EXPRAIS causes the deficit in the Theory of Mind Mechanism. What we would not want to say is that the neural deficit causes the deficit in EXPRAIS. A deficit in one is *equivalent to* a deficit in the other. The two claims are identical and so cannot have a causal relationship.[3] Cause is an irrelevant concept here.

---

[2] Note that you do not have to agree with the specific illustrative theory in order to appreciate the logic of the point being made. It could turn out that EXPRAIS was analysable into two parts, X & Y, each of which would either be a primitive or analysable. And so on. Equally, if you don't like EXPRAIS, substitute your own theory - of equal specificity, of course. Then apply the same logic. Note that if you postulate a primitive that is the result of learning, then its absence could be caused through a problem in some learning mechanism.

[3] Consider someone claiming that a fracture in the tibia caused the shin bone to break, recalling that the anatomical name for the shin bone is *tibia*.

equivalence: brain to cognition. In the previous paragraph, our starting point was a cognitive primitive which, at the moment, lacks a neural instantiation. Next, we take as our starting point a structure identified on the biological side. One candidate as a biological primitive (in the sense we are developing) is "mirror neurons" (Rizzolatti et al, 1996; Gallese et al, 1996). These are neurons in the pre-frontal lobe of the monkey cortex which respond selectively both when the animal sees another animal pick up an object in a particular way and when the animal picks up an object itself in the same way. The observed actions which most commonly activate mirror neurons are grasping, placing and manipulating. Other mirror neurons are even more specific, responding to the use of a particular grip - such as a precision grip or a power grip - as well as to the animal itself picking up an object using the same kind of grip. It is important to understand what these neurons are not for. To start with they cannot be interpreted in terms of the preparation for an impending movement since the neurons still fire when the monkey sees a specific movement while it is engaged in an action unrelated to the observed movement. Neither can the response be interpreted in terms of the anticipation of a reward, since there is no response from the neuron when the experimenter picks up a piece of food with a pair of pliers prior to the monkey picking it up, whereas if the experimenter picks the food up with his fingers, the target neuron does fire (Rizzolatti & Fadiga, 1998). According to Rizzolatti & Fadiga (1998), the most likely interpretation of the mirror neurons is that their discharge generates an internal representation of the observed action. Since there are a number of neurons which specialise in the same action, we can assume that, in respect of "generating an internal representation", these collections can be considered as functional units. What we wish to propose is that these collections can be seen as *equivalent to* cognitive elements. To understand this, let us take a two collections of these mirror neurons, one specialising in grasping - *MNg* - and one specialising in placing - *MNp*. Let us further note that while these collections of neurons could, in principle, play a role in imitation, their activation cannot be seen as being equivalent to imitation. The reason for this is that, while apes and humans imitate the actions of others, it seems that monkeys do not. According to Rizzolatti & Fadiga (1998), "monkeys, although endowed of a mechanism that generates internal copies of actions made by others, are unable to use them for replicating those actions" (p. 91). We might hypothesise that the neural circuits involved in the imitation of action in the apes included collections of neurons such as *MNg* and *MNp*.

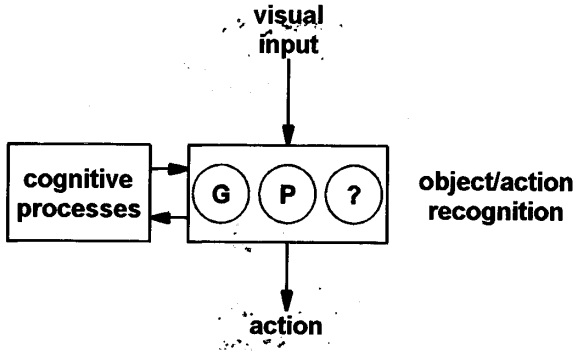

Figure 4 - cognitive model of grasping action meanings. Whether an action follows output of the units will depend on other factors.

Let us now consider a possible *cognitive* theory about the recognition of actions which could be seen as equivalent to the biological theory outlined above. This model, illustrated in figure 4, is based on the early Logogen Model (Morton, 1969). It considers that there are units, equivalent to logogens, which fire when there is enough evidence in the input to conclude that a grasping action, G, or a placing action, P, etc. is occurring. Such firing would have the effect of "generating an internal representation" of that action, which would be interpreted by cognitive processes. In addition, such units could be stimulated by inputs from the cognitive

processes (an internal action goal), as a result of which instructions would be sent to the effector systems to use the specified action on an object of current attention. Under this, or a similar theory, the unit labelled *G* would be equivalent to the neural collection *MNg*, and the unit labelled *P* would be equivalent to the neural collection *MNp*. In either a brain theory or a cognitive theory, the sets of terms could be used interchangeably without any change in meaning. That is, a statement about *MNg* would be identical to a statement about *G*, and vice versa.

To summarise, then, we are claiming that in some cases there may be equivalence between biological and cognitive descriptions of functional entities, and that the relation between the two is to be contrasted with the causal relationship which is the primary concern of causal modelling.[4] We may next see that the equivalence relationship also contrasts with the realisation relationship.

2. ***Realisation*** refers to the way in which cognitive elements (such as processes or beliefs) are matched (embodied, represented) in the brain in cases other than those covered by the definition of equivalence. To illustrate, consider that any change at the cognitive level leads to a change at the brain level. However, such changes are not usually part of a causal model. Thus take the case that we come to believe something new: that this article has to be completed by the end of March, for example, rather than by the beginning of March. This welcome change in belief will, of course, be accompanied by some change in brain state. However, the causal consequences of the change of belief cannot be traced or predicted from the change in brain state. We have to consider the cognitive (meaning) representation to do this. Indeed, although the change in belief is identical for both of us and the causal consequences are similar, the accompanying changes in brain are very unlikely to have any relationship.

To help understand this point, we can contrast this example, where there is no causal relationship from brain to cognition, with a case where the change in brain state might be directly involved in a causal chain. Consider someone involved in a traumatic event such as a traffic accident. According to some theories (e.g. Metcalf & Jacobs, 1998), the fear (cognitively mediated) engendered by the event leads to the memory record being stored without the involvement of the hippocampus. The consequence of this is that the memory becomes intrusive and the person suffers from flashbacks. This sequence can be

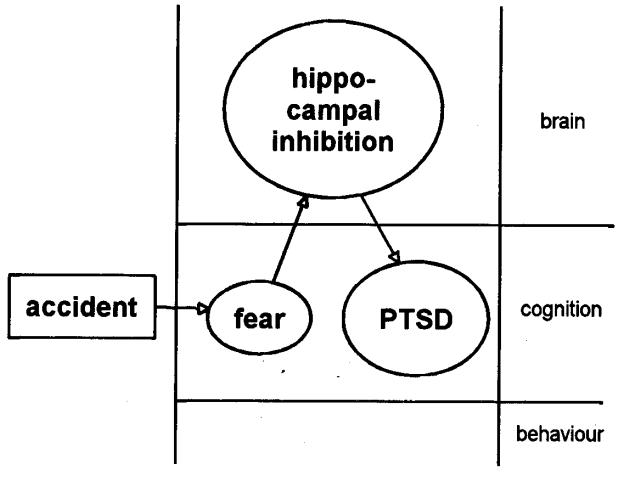

Figure 5 - a model for PTSD in which a cognitive change leads to a change in brain state with further cognitive consequences

[4] In considering the lack of constraints from biological facts onto cognitive theory, Mehler, Morton & Jusczyk (1984) made an exception of cases of a tight relationship between the levels - effectively what we mean by equivalence in this article.

seen as a cognitive state having brain consequences which lead to further cognitive outcomes and is illustrated in figure 5.

A further example is in relation to pain. The traditional theory of pain, as proposed by Descartes, was that pain fibres relay information to a 'pain centre' in the brain. Activity in the pain centre is then experienced as pain. This is the usual brain-to-cognition causal relationship. However, as Derbyshire (1997) points out, it has become clear that there is no direct relationship between the amount of stimulation of pain fibres and the experience of pain. Rather, with something like dental pain, psychological factors such as the circumstances of an operation, the patient's understanding of and attention to the trauma and levels of experienced anxiety intervene to produce the final perception of pain. Even for the apparently simple case of dental pain, then, a model is required where the social context, and psychological factors mediate at the biological level in the perception of pain. Derbyshire also summarises current work on the experience of pain by sufferers of rheumatoid arthritis. The theory specifies feedback from the psychological to the biological level, such that some aspects of a 'negative coping strategy' can lead to an increase in inflammation.

The two examples we have just discussed relate to acquired conditions. It is, of course, possible that some developmental disorders may originate at the cognitive level and would then need to be discussed along the same lines.

Why we need Cognition

As Rugg (2000) has pointed out, without cognition, localisation of function reverts to phrenology, with such human traits as "cautiousness", "conjugality" and "veneration" being assigned specific locations in the brain. The slight difference is that nowadays the labels correspond to redescription of patterns of data rather than terms drawn from intuitive theories of the mind.

In conclusion, let us add some history. Cognition has thrived in the years since 1972 and the reasons are clear to us. Initially it enabled people to develop the science of the mind without being crippled by the behaviourist empiricism of referees. Later, with the growth of cognitive neuropsychology, Cognition has allowed biologists and those who study behaviour to talk to one another. It is a role which suits Cognition and it is what cognition can uniquely accomplish.

References

British Psychological Society (1999). *Dyslexia, Literacy and Psychological Assessment: Report by a Working Party of the Division of Educational and Child Psychology*. Leicester: B. P. S.

Derbyshire, S.W.G. (1997) Sources of variation in assessing male and female responses to pain. *New Ideas in Psychology*, **15**, 83-95.

Frith, U., Morton, J. & Leslie, A.M. (1991). The cognitive basis of a biological disorder: autism. *Trends in Neurosciences*, **14**, 433-438.

Gallese, V. (in press). Actions, faces, objects and space: How to build a neurobiological account of the self. In *Proceedings of the III Conference of the Association for the Scientific Study of Consciousness (ASSC).*

Gallese, V., Fadiga, L., Fogassi, L. & Rizzolatti, G. (1996). Action recognition in the premotor cortex. *Brain*, **119**, 593-609.

Lindgren, S. D., De Renzi, E. & Richman,L. C. (1985) "Cross-national comparisons of developmental dyslexia in Italy and the United States. *Child Development*, **56**, 1404-17.

Mehler, J., Morton, J. & Jusczyk, P.W. (1984). On reducing language to biology. *Cognitive Neuropsychology*, **1**, 83-116.

Metcalf, J. & Jacobs, W.J. (1998). Emotional memory: The effects of stress on "cool and hot memory systems. *The Psychology of Learning and Motivation.* 38, 187-222. Academic Press: NY.

Morton, J. (1986). Developmental contingency modelling. In: Van Geert, P.L.C. (Ed.), *Theory Building in Developmental Psychology*. North Holland: Elsevier Science Publishers B.V., 141-165.

Morton, J. & Frith, U. (1993). What Lesson for Dyslexia from Down's syndrome? Comments on Cossu, Rossini, and Marshall (1993). *Cognition*, **48**, 289-296.

Morton, J. & Frith, U. (1993). Approche de la dyslexie développementale par la modélisation causale. In: J.-P. Jaffré (Ed.), *Les Actes de la Villette*. Paris: Nathan.

Morton, J., & Frith, U. (1995). Causal Modelling: A Structural Approach to Developmental Psychopathology. In: Cicchetti, D. & Cohen, D.J. (Eds.), *Manual of Developmental Psychopathology. Volume 1.* New York: Wiley, 357-390.

Rizzolatti, J. & Fadiga, L. (1998). Grasping objects and grasping action meanings: the dual role of monkey rostroventral premotor cortex (area F5). In *Novatis Foundation Symposium 218: Sensory Guidance of Movement.* Chichester: Wiley. (Pp 81-103)

Rizzolatti, G., Fadiga, L., Fogassi, L. & Gallese, V. (1996) Premotor cortex and the recognition of motor actions. *Cognition & Brain Research,* **3**, 131-141.

Rugg, M. (2000) Human memory: a perspective from cgnitive neuroscience. Inaugural professorial lecture, 8 march, 2000. University College London.