# Connexions

*An online journal of cognitive science*

ISSN 1368-3233

EDITORS: GAVIN BOYCE, TOM DICKINS, KEITH FRANKISH, AND ADAM HUFFMAN

Issue 2 August 1997


## Contents

**Article**


| | | |
|---|---|---|
| PDP, Levels, and the Status of Explanation | **Andrew Bailey** | **2** |


**Reviews**


| | | |
|---|---|---|
| Elman et al., *Rethinking Innateness* | **Jill Boucher** | **8** |
| Steven Mithen, *The Prehistory of the Mind* | **T. E. Dickins** | **11** |
| The Vancouver Meta-Representation Conference | **David Byrd** | **16** |
| The Ciba Open Meeting | **Margaret Clegg** | **19** |
| The 1997 BPS Conference | **T. E. Dickins** | **21** |
| The Second Hang Seng Workshop on the Evolution of Mind | **Jennie Hawcroft** | **23** |
| The Davidson Conference at the Bolton Institute | **K. A. Wilkinson** | **24** |



*Sponsored by the Hang Seng Centre for Cognitive Studies*

# PDP, Levels, and the Status of Explanation

Andrew Bailey
*The Department of Philosophy*
*The University of Calgary*

**Abstract**: *This paper examines the status of the 'levels hypothesis' in cognitive science, and its implications for the 'parallel distributed processing' project. I consider in what respect levels of analysis are separate, and the significance of that (putative) separateness, and conclude that the most fruitful account of levels says that the levels of our account of the mind correspond to different explanatory levels. I suggest that this provides a conceptual space into which PDP might be slotted which saves it from collapse into neurophysics or a serial processing paradigm, and retains for it as many levels as seem to fill an explanatory niche. Further, on this view, PDP need not be considered an 'approximation' or 'idealisation' but could deal wholly adequately with the domain of relevant regularities and categories it has set itself.*

## Introduction

In this paper I want to examine the status of the 'levels hypothesis,' and its implications for the parallel distributed processing project in cognitive science. That is, I want to consider in what respect levels of analysis are separate, and the significance of that (putative) separateness. This, I think, impacts on the PDP project in two major ways: firstly, it addresses the question as to whether PDP is ultimately collapsible into more conventional (serial, von Neumann) approaches; secondly, it helps show what PDP theories should look like-in particular, whether they should have three levels, two, or perhaps even one.

After the usual exegetical preliminaries, I will proceed by considering various possibilities for just what is meant by separate 'levels' (in such contexts as, for example, David Marr's distinction between the levels of computational theory, representation and algorithm, and physical implementation; or Paul Smolensky's neural, subconceptual and conceptual split). I shall conclude that the status of levels is most plausibly that of picking out different equivalence classes of patterns of activity in different explanatory contexts, and that 'causal explanation' can be split from 'causal efficacy' as Andy Clark suggests in his book *Microcognition* (Cambridge, MA: MIT Press, 1989).

## I. PDP

The following 'definition' of PDP is intended to be of sufficient generality to capture most proponents of the connectionist programme, though it is mostly influenced by Smolensky and Clark.

1. Mental processing is (at bottom) parallel and distributed. At both the neural level and the computational level the mind is not a CPU but a network of simple processing units.
2. Computation is essentially numerical and does not consist in the manipulation of representations.

According to Paul Smolensky in his paper "On the Proper Treatment of Connectionism" (Cole, Fetzer and Rankin (eds.), *Philosophy, Mind and Cognitive Enquiry*, Dordrecht: Kluwer, 1990), each network unit has an 'activation value,' and each connection has a certain 'weight,' either positive ('excitatory') or negative ('inhibitory'). Connectionist networks operate by reacting to the changing activation values of 'input units' and relaxing into a new set of activation values among the 'output units' of the web. Intermediate 'hidden units' are typically also involved.

3. Models of the mind should not be 'semantically transparent'; subsymbolic explanations "rely crucially on a semantic … shift that accompanies the shift from the conceptual to the subconceptual levels" (Smolensky *ibid.*, 172). That is, there is no neat mapping between a conceptual level description of behaviour and a projectible interpretation of the internally represented objects of computational activity.

4. Connectionist networks do, however, in a sense, encode representations. "The entities in the intuitive processor with the semantics of conscious concepts of the task domain are complex patterns of activity over many units. Each unit participates in many such patterns" (Smolensky *ibid.*, 159). The constitution of a token symbol depends on its context; indeed this context is part of the symbol, and consists of subsymbols.
5. Conscious rule following-the top level-can be modelled by a virtual machine (a von Neumann, serial CPU) instantiated in a PDP network.
6. The computational level is distinct from both the neural and belief/desire levels, according to Smolensky and Clark. Smolensky calls it the 'subconceptual level' and argues that it "does not admit a complete, formal, and precise conceptuallevel description" (Smolensky *ibid.*, 160).
7. Connectionist networks are biologically more plausible than their symbolic counterparts. For example, they naturally produce shading of meaning, the ability to operate with faulty or incomplete information, and can handle evolutionarily more basic skills. In part, this is because they can operate with 'soft constraints'-that is, constraints that singly have no implications because any one can be overridden by the others.

## II. The Levels Hypothesis

This is essentially the idea that, to explain mental life, one must make reference to at least three levels:

1. The level of neuronal functioning;
2. The level of computation;
3. The level of reflective awareness-that at which the kind of mental life described in belief/desire psychology takes place.

I will pick out what I take to be the two fundamental stories about levels on a purely causal picture (supervenience and emergentism), then swiftly consider various other accounts put forward by Clark, Smolensky and Pylyshyn, some of which seem to contradict this dichotomy. I will end up putting aside the whole 'fundamental cause' picture, and suggesting that a story about explanatory contexts is more fruitful; this picture I will consider in more detail.

## III. Epistemological C onvenience: The Causal Story 1

The first, and perhaps simplest, story about levels is that they are simply a matter of epistemological convenience. This story can come in two versions: strong and weak. The weak version says that operating at the 'real' fundamental level is just too much like hard work, and that it is simply quicker and easier to operate at higher levels, which, presumably, utilise an array of operative simplifications and approximations to make things simpler. The strong flavour adds to this that it is not only easier to operate at higher levels, but it is beyond our human cognitive capacities to do otherwise: we simply cannot mentally handle the vast complexity at the bottom level.

This picture, I think, is inherently physicalist (as part of a basically causal story, anyhow): it assumes that all causation occurs at the level of physical objects and nowhere else. Further, it seems, it contains the assumption that there is only one 'physical level'-or at least some fundamentally causal stratum of the physical; perhaps subatomic particles, or maybe the four fundamental forces (weak atomic force, strong atomic force, gravity, electromagnetism). Levels are supervenient and, in a sense, 'apparent': they are of only pragmatic, contingent use, and not essential to, indeed not even part of, a full description of the universe (including the human mind). Thus this picture is eliminativist: all levels above the fundamental physical -- including the level of reflective awareness -- can (and probably should) be dispensed with in any scientific theory of the mental.

I suggest that this account of levels is not what PDP theorists -- or any cognitive scientists -- should want to work with. The reason, even leaving aside the knotty problem of clarifying what exactly constitutes a "fundamental causal level," is essentially that the level of computation is higher than the fundamental causal level. Thus, just as those who want to save intentional psychology as some sort of ontologically useful theory would, on this picture, be unable to evade its elimination as all but a pragmatic, approximate shorthand language; so those working in cognitive science would fail to prevent it collapsing into neurophilosophy-and ultimately particle physics.

Let me be clear about what I am and am not saying. I am not suggesting that the causal hierarchy of the sciences is untenable: I do not deny that (if this is what physicists would tell me) causes 'originate' at some subatomic level and all the phenomena we see at the highest, most everyday, level are simply aggregations of these causal interactions. Thus, I need not dispute, for example, that chemistry supervenes on physics: causal explanations in the former science are reducible to physical explanations. What I do deny-or at least, suggest the cognitive scientist must deny-is that the levels above fundamental physics are merely 'apparent,' or that they merely 'approximate' the details of lower levels. There is hopefully some sense in which analyses at the conceptual or subconceptual level are autonomous (we shall remark on plasticity in a moment) and 'real.'

## IV. Ontological Properties: The Causal Story 2

The other major story about levels is an emergentist one; it says that levels are ontologically distinct since items at higher levels have causal powers not reducible to those of items at lower levels. On this account levels are 'real' in the sense that no description of the universe would be complete if it excluded the higher levels. The properties and individuals picked out at higher levels are ontologically real, and since they are not reducible to lower level properties and entities an account only of those lower levels would miss them.

This sort of story would be ideal in formulating accounts of the reality and separateness of levels-if it were true. Unfortunately, emergence does not fit happily into contemporary scientific theory; it smacks of mystery and the immaterial. (Since, if all scientific explanation is subsumption under actual causal laws, there must presumably be either an infinite regress (mysterious!) or an inexplicable 'bottom layer' of uncaused causes (brute regularities); then in that case I'm not sure how much more spooky are irreducible causes at some 'fundamental' level compared with such causes emergent at a higher level. But still.) Also, it's not clear that the properties picked out by higher level analyses are always the sort of thing which looks emergent; we shall consider the example of 'roundness' and 'squareness' a little later.

## V. The Computer Metaph or

This account of levels runs essentially like this: lower level stories of the mind are to higher level stories as lower level computer languages are to higher level programming languages. The higher levels are 'real' in the sense that they form functional architectures-in effect, new, virtual machines. This account is to be found in Pylyshyn's book *Computation and Cognition* (Cambridge, MA.:MIT Press, 1984). Essentially, Pylyshyn seems to view the level of functional architecture as "the right level of specificity (or level of aggregation) at which to view mental processes, the sort of functional resources the brain makes available" (92). It is strongly analogous to the programming language of a computer, which 'sits on top of' the machine language (neural level?) and determines which operators are primitive, what sequences are allowed, what limitations there are on the passing of arguments, and so on.

His main argument to suggest that this is the 'correct' level from which to view mental processes is that, by virtue of being representable by the same program in some virtual machine, two programs are considered 'strongly equivalent'-that is, they represent the same algorithm/cognitive process. Hence the functional architecture level seems to be used by Pylyshyn as the best level from which to spot what he considers the relevant regularities in computation: the algorithms operated by the brain, no matter how they are physically realised. Thus, it seems, for Pylyshyn his criterion for levels (a computational strata lying between the neuronal and the reflective levels) is neither epistemological, nor ontological, but concerns the isolation of the relevant explanatory generalities. More of this anon.

## VI. The Physics Metaphor

Another analogy, found in Smolensky, is the following: lower level accounts of the mind are to higher level accounts as the theory of quantum physics is to the classical, Newtonian theory. This seems something like the epistemological view of levels considered above. In fact the universe (mind) is a quantum system (PDP network) at bottom and through and through, but at more 'macroscopic' levels it looks like Newtonian mechanics (symbolic computation). Thus these theories (classical physics and conceptual processing) are actually idealisations of reality, useful approximations which serve well enough at higher levels but are not true descriptions of the causal processes underlying the phenomena.

## VII. Modelling

Another suggestion, hinted at in Smolensky I think, might be that the different levels are simply different ways of modelling the same phenomena. In other words, the mind provides us with a set of data which, as usual, underdetermine their theoretical interpretation, and a variety of competing hypotheses and models can be constructed to try and explain them. These models can be on different levels, and might include, for example, PDP, the Language of Thought hypothesis, common sense psychology and neurology.

Two questions then arise:

- which models are merely instrumental, and which is a description of the real situation?
- which model is the most fruitful as a research programme and a tool for explanation and prediction?

The answers to these questions need not be the same. Firstly, of a set of models, that which (unbeknownst to the researchers) is the most accurate, need not be the most fruitful-in the shortterm, at least. For example, it could be said that a neurological model is most likely to be the one that models the 'real causal operations' of the brain (without being reducible to more fundamental causes), yet perhaps we will have the most success in the shortterm at limiting the broad structures of cognition by pursuing the PDP program.

Secondly, again, it seems that there is surely real explanatory value in levels of description at a higher level than the fundamental causal operating level. I will expand more on this below, but as a preliminary point, it looks like neurology actually won't be the most basic causal level within the brain -- neurology itself is almost certainly reducible to quantum physics, which in turn may even be reducible to the physics of fluctuations in 'superforce.' As we descend through more and more causally basic levels, it seems that the categories and uniformities to be observed become increasingly ironed out -- flattened into repetitive physical detail. Consequently, the explanatory vocabulary at lower causal levels becomes less rich and less useful. Even at the level of Newtonian descriptions of bodily movements, psychological prediction becomes impossible. At the level of neural interactions, or, below that, of atomic causal properties, or even further to quarks, prediction and explanation of all but the most local of effects becomes prohibitive. Further, as I shall suggest below, the explanatory vocabulary does not, and cannot, answer the relevant questions in order to serve as an adequate explanation.

## VIII. Levels Of Explanation

So finally, we come to what I suggest is the most fruitful account of levels. This, simply put, says that the levels of our account of the mind correspond to different explanatory levels. Whilst one of these gives a baselevel account of how mental phenomena and behaviour are actually caused, that does not make the others any the less explanatory. Here we leave a solely causal story behind and enter the realms of explanatory roles.

This account is a mixture of ideas to be found in Clark and Pylyshyn. Andy Clark's book *Microcognition* (*ibid.*) provides the notion of ascriptive meaning holism and a causal explanation/causal efficacy split. Pylyshyn (*ibid.*) adds talk of activity pattern equivalence classes.

### *i. Ascriptive Meaning Holism*

Clark claims, plausibly enough, that propositional-attitude ascription -- in effect, belief/desire psychology -- "is a net thrown over a whole body of behaviour and is used to make sense of the interesting regularities in that behaviour" (48). Beliefs are only ascribed in complexes, for example: the ability to do basic mathematics is required before we are prepared to ascribe understanding of the concept of 'three.' This, Clark emphasises, is not a point about the epistemic liaisons of propositional attitudes (i.e. those propositions that the agent takes to be relevant in assessing the truth value of the propositional attitude in question); rather, it is solely a point about belief/desire ascription. In other words, we are talking about the point of propositional attitude ascription -- its explanatory role.

Common sense psychology aims to make sense of behaviour, and is gauged to that end. Its principles and categories "need not correspond to any natural, projectible divisions in whatever underlying physical or computational structures make possible the behaviours" (49).

***ii. Causal Explanation vs. Causal Efficacy***

In essence, Clark argues that the increase in generality of explanations, bought at the cost of becoming a less full account of the particular causal chain responsible, is an explanatory virtue. What he calls (following Jackson and Pettit) 'program explanations' abstract away from the specific causally active features of a particular case and highlight a common feature of a range of cases.

One of the examples Clark cites is the peg and hole explanation, whereby we explain the fact that a square peg won't pass through a circular hole by the general fact that squareness won't pass through an equivalent area of roundness. We do not provide the 'real' causal story at the subatomic level -- we don't give the full story of all the actual, causally efficacious factors. And the reason for this is not primarily that we have epistemological problems dealing with all those quarks, but that the subatomic facts are not the ones we wish to group together. At the subatomic level, roundness and squareness are not visible -- but it is square pegs and round holes we wish to formulate a principle for. Further, there might be some Twin Earth where matter is not mesons and bosons but Xs and Ys; there, the causally efficacious microstructural facts will be different, but still higher level descriptions and explanations in terms of roundness and squareness will be operative.

Consider, also, the old functionalist dogma of alternative realisability or plasticity of mind: mind and its properties, that which we are interested to explain, could in principle remain constant while its causally efficacious makeup radically altered. So, the argument runs, explanations of mind cannot be phrased in terms of organic, grey stuff; at the wet level, mind is invisible. 'Process explanations,' by contrast, narrate the very features which are causally efficacious in a particular *explanandum*.

***iii. Activity Pattern Equivalence Classes***

Pylyshyn argues that "explanations are relative to particular vocabularies. … Explanations attempt to capture generalisations, and different vocabularies reveal different generalisations" (2). He claims that we are thus required to resort to a 'cognitive vocabulary' in order to do psychology. His example is the 911 problem: someone witnesses a serious accident, rushes to a phone box, and dials 9 and 1. Why? What will she do next? The only way to give an adequate explanation of this scenario is with the intentional vocabulary of common sense psychology.

This is not because there is no other, more fundamental, causal story to be given; we need not deny causal closure in order to insist on intentional explanation. The problem is that a causal explanation would capture the wrong sort of generalisations -- generalisations which are much less useful than our holistically ascribed psychological explanation and its categories. That is, for certain types of explanation, there are especially relevant generalisations; further, these are actually more predictive in practice. This is fairly obvious -- a mere list of behavioural evidence would be unlikely to allow even the most powerful of computers to predict the behavioural sequence of dialling a further one. And there are an infinitude of possible physical situations which could instantiate our telephone example.

More importantly, physical descriptions of the situation miss the relevant generality; each instance would count as a different sequence, and the predictively significant regularity would be invisible. There is something in common between all such 91(1) situations, we surely want to say, but this regularity is not an actual causally efficacious regularity and so is not evident at the fundamental causal level. The generality is of another sort: not an approximation of aggregated causal interactions, but a genuine, non-fundamentally causal way of grouping events in the world.

***iv. Explanation***

Perhaps the standard philosophical account of explanation, typified by the deductivenomological model of Carl Hempel, is the covering law model. This talks of explaining phenomena by subsuming them under projectible (from observed to unobserved instances), counterfactual-sustaining universal statements. A proper explanation, on this view, consists of a set of statements in roughly the form of modus ponens (though the explanation may be probabilistic rather than deductive): a statement of the *explanandum* (B), a set of initial conditions (A), and a law connecting the initial conditions with the *explanandum* (if A then [probably] B).

However nothing about this model of explanation requires that it deals only with the fundamentally and actually causally efficacious entities involved in a particular event. As I have suggested above, there can be projectible universal statements which subsume categories of events without those laws dealing with the fundamental physical causal layer: the layer forbidding square pegs to enter round holes, for example; or the law which states that if one desires that P above all else, and believes that P only if Q, then one will do Q. That is, genuine, useful causal explanations need not describe the causally efficacious.

## IX. Conclusion

Where does this leave the PDP project? I suggest that it provides a conceptual space into which PDP might be slotted which saves it from neurophysics, and, if this route is taken, retains for it as many levels as seem to fill an explanatory niche. (Clark, for example, identifies five possible levels of analysis of explanatory value: the numerical, the subsymbolic, cluster analysis, symbolic AI, and common sense psychology.) The freefall into the 'fundamentally causally efficacious' is averted by demonstrating that projects such as PDP have an explanatory value of their own, autonomously from the project of uncovering basic physical causation. Further, PDP need not be considered an 'approximation' or 'idealisation' but could deal wholly adequately with the domain of relevant regularities and categories it has set itself.

# Reviews

## Elman et al., *Rethinking Innateness: a Connectionist Perspective on Development*

**1996 Cambridge Ma.: MIT Press ISBN: 0 -262-05052-8 487pp. £34.95 (hb)**

**Jill Boucher**
*Department of Human Communication Sciences*
*University of Sheffield*

As its title indicates, the topic of this book falls within the nature-nurture debate, a theoretical area which has often involved arguments generating more heat than light. This book is conspicuous for the light it sheds and for the care taken to minimise heat generated.

The book consists of seven chapters. In the opening chapter 'New perspectives on development' the authors suggest that recent advances in developmental and cognitive neuroscience on the one hand, and in connectionist modelling on the other, can help us to understand how nature and nurture interact in human development. In particular, the authors suggest, these advances may help to answer a number of basic questions posed by developmentalists, including:

- What does it mean for something to be innate? What is the nature of the 'knowledge' contained in the genome?
- Why does development occur in the first place?
- What are the shapes of change? What can we infer from the shape of change about the mechanisms of change?
- Can we talk meaningfully about 'partial knowledge'?
- How does the environment affect development, and how do genetic constraints interact with experiences?

Chapter 1 also includes some introductory information on how genes work, and justifications for what the authors term their 'biologically oriented connectionism'. And there is an extended section discussing the first question posed above, namely 'What does it mean to be innate'. Here they introduce some of the key themes of the book, which include a move away from the notion of genes as blueprints for specific neural systems and related behaviours, and towards a notion of genes as catalysts for developments at all levels from molecular to cellular to neural to behavioural. Related to this is a move away from the notion of innate representational knowledge, and towards seeing behavioural development as constrained by neural architecture and its computational properties, and on the timing of different phases of development. The minimal role given to representational knowledge within the authors' theory relates in turn to a move away from the notion of the innate modularity of brain and mind, and towards a broader interpretation of what might be meant by specificity and how it is arrived at in development.

Chapter 2 constitutes an excellent introduction to connectionism for readers (of whom I am one) who know rather little about the subject. Chapters 3 and 4 give detailed accounts of simulations of various aspects of cognitive development, in particular language learning and the development of knowledge about the physical world. The work described in these chapters has been selected so as to support the authors' claim that linguistic and other learning resembling that which occurs in normal children can be simulated given quite simple computational architectures and learning algorithms, and also to illustrate how the number of mechanisms required to explain complex patterns of learning can be reduced within a connectionist framework. Dynamical systems theory is briefly outlined in Chapter 4, providing another eminently clear mini-tutorial for the uninitiated.

Chapter 5 is entitled 'Brain development' and, like Chapter 2 on connectionism, constitutes a highly informative introduction to concepts, theories and research in its specialist area. However the main focus of this chapter is on the origins of specificity in the cerebral cortex. The authors open their argument by stating that

> We find no evidence that our species possesses unique brain cells or circuitry. Rather, humans simply seem to have a greater extent of cerebral cortex, with a delayed timetable of postnatal development. (p. 240).

They approve a theory which suggests that particular thalamic afferents normally innervate particular regions of cortex not because of genetically coded 'instructions', but from a combination of timing and spatial constraints. To back this argument up, they cite evidence of the considerable variation in localisation of functions which results under normal as well as under abnormal conditions of development, while stopping short of arguing for unconstrained brain plasticity.

Chapter 6 is entitled 'Interactions all the way down', and covers a range of cases which illustrate interactions at all levels during development, from genes to language:

"In each instance, we try to show that a complete understanding of the phenomenon can only be achieved by understanding the complex web of interactions which are necessary for producing the phenomenon. In our view, the role of the genetic component - call it the innate component, if you will - is to orchestrate these interactions in such a way as to achieve the desired outcome." (p 321).

In this chapter, also, the authors consider why human development takes so long, in view of the fact that a prolonged period of development would seem to be highly maladaptive.

In the final chapter, the authors summarise what they consider to be the endpoints of the evidence and arguments presented in the book. These are:

- Mechanism and content are not the same thing, and there need not be a one-to-one mapping between the two. Most domain-specific outcomes are probably achieved by domain-independent means.
- Relationships between mechanisms and behaviours is frequently nonlinear. Dramatic effects can be produced by small changes.
- Knowledge ultimately refers to a specific pattern of synaptic connections in the brain. In this very precise sense, there is probably no innate higher-level knowledge.
- The developmental process itself lies at the heart of knowledge acquisition. Some complex behaviours may not be acquirable without passing through a developmental pathway.
- Connectionism provides a framework for understanding emergent form and the interaction of constraints at multiple levels; connectionism should not be thought of as radical empiricism.
- Development is a process of emergence.

Up to this point the authors have remained steadfastly non-confrontational, choosing to make their case without reference to the fact that if they are correct, a lot of very well known cognitive scientists are going to have to radically change (in some cases) or at least to modify (in other cases) their current theories. On p. 367, having summarised their main conclusions, the authors mildly enquire "Does anyone disagree?", and to show that they have not been dismantling a straw man they quote some strong claims for representational nativism made by Carey and Spelke, and by Leslie, amongst others. They then present 'Twelve arguments about innate representations, with special reference to language' which begins *"We think it unlikely that Universal Grammar could be encoded directly in the genotype*" and which concludes

> - grammars may constitute the class of possible solutions to the problem of mapping nonlinear thought onto a highly constrained linear channel. If children develop a robust drive to solve this problem, and are born with processing tools to solve it, then the rest may simply follow because it is the natural solution to that particular mapping process (p. 390).

Finally, a few thoughts on 'Where do we go from here?' concludes the book.

Inevitably this book will be seen as 'anti-nativist' and 'anti-modularist' (especially by those who only come to know of it by hearsay, rather by reading it). However, the authors repeatedly point out that their purpose is to reach a better understanding of what is innate and of how domain-specific neural and behavioural systems develop, rather than to deny the major role of biological endowment in development, or to deny the fact that the adult mind and brain are highly systematised.

What they do argue against is the likelihood that infants are biologically endowed with cortically instantiated representational knowledge. If their arguments are correct, then it is hard to see how their position could be reconciled with there being an innate symbolic language of thought, such as Mentalese. In addition, if 'innate knowledge' entails innate representations, then knowledge of natural kinds, the principles of Universal Grammar, and innate knowledge in other domains would also be inconsistent with the book's central claim. However, if Carey and Spelke, Chomsky, Pinker, Leslie and others do not insist on innate representational knowledge, and are happy to talk instead in terms of innate computational (neurally instantiated) architectures which reliably establish cortical representations of natural kinds, or of the principles of Universal Grammar, then the area of difference between the so-called 'representational nativists' and the authors of this book reduces.

Suppose that representational nativists became 'mechanism nativists', would that produce harmony? Almost certainly not, since the erstwhile representational nativists would almost certainly join forces with those evolutionary psychologists who posit the existence of many highly specific innate mechanisms, eg. for mate selection, cheater detection, food selection etc etc, and this would be at odds with the position taken in the book, viz:

> To the extent that realistic developments can be modeled in a (connectionist) system of this kind, there is a prima facie case against the notion that domains like language, music, faces or mathematics must be carried out by dedicated, innate and domain-specific neural systems. (p.41).
>
> The dramatic results of experiments involving transplanting and rewiring bits of cortex to to other regions...argue strongly against solutions which depend upon innately specified populations of neurons prewired for complex cognitive functions such as language. (p.19).
>
> Existing evidence for the localisation of language (such as it is) provides little support for the idea that children are born with neural mechanisms that are prespecified for and dedicated solely to language processing.

The authors' favoured solution to the key issue of the domain-specificity of adult knowledge is that the structure of certain domain-specific tasks or problems to be solved (for example, acquiring a language for communication) leads inevitably to certain domain-specific and to some extent universal solutions. The hexagonal shape of the honey cells made by bees is quoted as one example of a universally occurring, species-specific phenomenon which is not specified in the genome, but rather results from the problem of maximising the use of limited space, combined with the physical properties of beeswax. Spandrels provide another example. An obvious gap in the book (which the authors would be the first to recognise: they admit that they know about 5% of what they would like to know and that 95% of what they think they know may be wrong) is a full demonstration and explanation of how the problems posed by, for example, the drive to communicate or the need to mindread combine with domain-independent (not necessarily domain-general) architectures and processes to produce language and a theory of mind.

I was not happy with all of the arguments proposed in the book. For example, I think the case for specificity at the level of architectural constraints is understated, and the case for brain plasticity correspondingly overstated. No doubt other readers, and especially any lingering representational nativists, and 'special-mechanism nativists', will find plenty of individual points to take issue with. However, this is a groundbreaking book, packed with clearly presented information and stimulating argument, and written with a combination of authority and modesty which is quite unusual and a very winning combination. The book's value will, I suspect, lie precisely in the arguments which it provokes, so criticism and controversy will testify to its importance rather than to its weaknesses. So I advise readers of Connexions to read it! And if you baulk at the book's length (nearly 400 pages), just read the final chapter, which brings together most of the main arguments. If you do this, you will miss a great deal, but at least not be quite out of the swim.

# *Generality, Modularity and then Fluidity: A Review of Stephen Mithen, The Prehistory of the Mind*

**1996, Thames & Hudson, ISBN 0 -500-05081-3, pp. 288, Hbk £16.95**

**Thomas E. Dickins**
*Department of Human Communication Sciences*
*University of Sheffield*

## (1) Introduction:

Steven Mithen's new book, The Prehistory of the Mind, is a beautifully written investigation into the evolution of cognition. Mithen uses a number of metaphorical devices to navigate the literature. The book opens with a summary account of the history of our species as a four act play. Mithen's aim is to carefully write the stage notes in order to get the players to just the right point in the story (pp. 17 - 23). The acts correspond to time slices and it is Acts 3 and 4 that are the most exciting with regard to humans; Act 3 spans from 1.8 million to 100,000 years ago, Act 4 from 100,000 years ago to the present day.

Mithen has also adopted the metaphor of the mind as a cathedral. The contemporary cathedral of the mind has resulted from three building phases spanning a cross section of evolutionary time. Each phase denotes a response to the adaptive demands facing our ancestors and Natural Selection is to be regarded as the overseeing architect:

> Phase 1. Minds dominated by a domain of general intelligence - a suite of general purpose learning and decision-making rules.
> Phase 2. Minds in which general intelligence has been supplemented by multiple specialized intelligences, each devoted to a specific domain of behaviour, and each working in isolation from the others.
> Phase 3. Minds in which the multiple specialized intelligences appear to be working together, with a flow of knowledge and ideas between behavioural domains. (p. 64)

Phase 1 represents a simple architecture with a central nave dominating the structure within which slow and error prone learning occurs. Phase 2 sees the addition of a number of chapels in which very specific, and more complex functions are fulfilled. The learning within these chapels is rapid and makes fewer errors. Mithen suggests that three chapels will dominate - those serving to grant social intelligence, natural history intelligence and technical intelligence, which can be otherwise termed intuitive psychology, biology and physics. Mithen stress that none of these chapels inter-connect - any interaction between domains must rely on the slow and error prone processes of the central nave and will be rare. Finally, phase 3 sees the addition of direct access between chapels that allows for the seamless integration of all the chapels' special services. Such integration allows for much speedier sharing of information and grants a more flexible grasp upon the world.

## (2) Psychological Arguments:

Mithen arrives at this structure of evolutionary change after considering the work of developmental theorists such as Susan Carey, Patricia Greenfield, Annette Karmiloff-Smith and Elizabeth Spelke. He sees the phases of evolutionary construction mirrored in ontogenetic development:

> The first (phase) is paralleled by the domain-general learning processes identified as critical to the very young infant; the second parallels the modularization of the mind with domain specific-thought and knowledge; and the third parallels what Karmiloff-Smith describes as 'representational redescription' and Carey and Spelke describe as 'mapping across domains' - when knowledge becomes available for use in multiple domains of activity. (p.64)

Earlier in the text Mithen also discusses the modular model of Jerry Fodor and the "multiple intelligences' idea of Howard Gardner (pp. 37 - 42). Mithen is not concerned to give a detailed critique of each position held by this array of theorists. Instead he wants us to realise that underlying all of these models is a common commitment to a certain degree of modularity but also to a degree of integration between these domains. It is this integration which is the defining mark of modern minds and this is what evolutionary stories must explain.

Mithen's use of ontogenetic evidence to formulate putative evolutionary stages may concern some readers as it appears to be a recapitulation argument - it is now widely accepted that ontogeny does not precisely mirror phylogeny. Mithen is sensitive to this, fully admitting his doubts about it, and simply argues that his three phases are an approximate measure to get theory started.

> I am going hesitantly to adopt the notion of recapitulation... . (The) reason for my hesitation is that I have no theoretical conviction that recapitulation of the evolution of mind during (ontogenetic) development necessarily occurs. If it does, I am sure that it is likely to be manifest in some broad parallels rather than any strict correspondence of phylogenetic and ontogenetic stages.
>
> Whether or not recapitulation of the mind is correct, it provides a means to establish the framework of hypothetical architectural phases which is needed to continue with my study. (p. 63)

Mithen goes on to suggest ignoring data from ontogenetic work would be academically negligent. If it meshes with data from other sources he feels that it has to be taken seriously.

## (3) Comparable Structures ?

Mithen's model is not dissimilar to that of Dennett's (1995, 1996), which is expressed through the equally structural metaphor of the "Tower of Generate and Test". Dennett is motivated by philosophical arguments and may subsequently constitute a support for Mithen's scheme. (For related comments on Dennett's evolutionary viewpoint, see Dickins and Frankish, 1997, from which the following sections on Dennett and Clark are adapted.)

Dennett's Tower consists of four floors, ascended through evolutionary time, each of which represents a more efficient way of solving day-to-day survival problems. Each progressive solution is a 'better move' than the one before. The ground floor is inhabited by Darwinian Creatures that are blindly generated by natural selection and possess different hardwired phenotypes. Their responses to survival problems are determined by their genetic inheritance and are quite inflexible. Although Mithen does not mention such creatures it is clear that he regards them as a necessary precursor to his generalised-intelligence-only organisms.

The second floor is inhabited by Skinnerian Creatures that can vary their behavioural response to the environmental contingencies they encounter. Skinnerian Creatures have hardwired reinforcement mechanisms that bias them to make what Dennett terms 'Smart Moves'. Such a creature will vary its response to stimuli until something good comes of it, whereupon it will become conditioned to produce the same response again should similar stimuli be encountered. Such conditioning is possible, of course, only if the initial response is not fatal. Skinnerian Creatures are akin to those found during the first phase of Mithen's cathedral construction. Dennett has committed himself to associative learning and Mithen adopts both "general" learning and decision making rules as a part of his architecture. Presumably, Mithen's notion of decision making rules is merely a short hand for the discriminative behaviour of an associative learner, although he does not make this clear.

Popperian Creatures, who inhabit Dennett's third floor, run less risk of making fatal first moves. These creatures have an inner environment - a mental representation of the external world - and can run internal simulations of various courses of action. In this way, they can calculate the likely effect of candidate actions and eliminate the ones likely to have undesirable consequences - thus 'permitting their hypotheses to die in their stead' as Karl Popper puts it.

Popperian Creatures are much smarter than their Skinnerian cousins. However, their ability to form and test hypotheses is still limited by their genetic endowment. Their representational abilities, in particular, may remain relatively encapsulated, so that information from one domain is not routinely made available for the solution of problems occurring in others. This is obviously analogous to phase 2 of Mithen's building programme. Gregorian Creatures, who live on the next floor, are smarter yet. They supplement their innate problem-solving abilities with "mind tools" acquired from their peers. They have learned Richard Gregory's lesson that tools not only display intelligence, but create it too. A well-designed tool meshes with our native abilities and extend them in new and far-reaching ways. (Think, for example, of how a pair of scissors extends our ability to manipulate and shape artefacts.) Mind tools are culturally transmitted tricks, shortcuts, and strategies which enable creatures to arrive more swiftly at Smart Moves for solving problems. These are phase 3 creatures.

## (4) A possible role for language?:

Both Dennett and Mithen argue that language may well be the key development in hominid cognition that allowed the integration of more modular mechanisms (see also Dennett, 1991). For Dennett, words are the principal mind tool that potentially allowed existing computations, or modular operations to be augmented. Mithen endorses the Dunbar (1993) model of natural language evolving to fulfil social coalition needs, however, Mithen's linguistic mechanisms are under-specified. At one point he suggests information travel between modules might have been caused directly by different types of language mixing up technical, social and other domains. Social utterances would predominate due to the evolutionary heritage of language. However, certain items that were commonly used would also be named As a result social utterances and the odd "name" for technical items, such as stone tools, etc. could have co-occurred in continuous "talk". At another point, Mithen hints that language may be a module in its own right and that modules were perhaps never completely sealed. Information, presumeably in non-linguistic form, from various domains may have leaked into the language module and the use of such information within utterances would somehow see to it that such a "mechanism" became fixed (pp. 185 - 187).

Mithen's loose hypothesising does serve to nudge the reader toward some interesting speculations about the role of public language in thought. Some of these have been made elsewhere by Dennett, as indicated, and recently by Andy Clark. Clark has argued that language augments existing computational abilities by externalising and recombining the information used by pre-linguistic computations in several ways. Clark sees language as fulfilling a Vygotskian scaffolding function. For instance, language aids us in mastering new skills. When learning to drive, for instance, we receive a set of linguistically encoded instructions from our tutor. These instructions can then be "replayed" or rehearsed by us when we are practising- either aloud or "internally" - such that we can recreate the teaching situation. Language, in this case, permits the mental rehearsal of self-directed commands and exhortations in order to focus behaviour, aid recall, and reinforce learning. Clark writes:

> The role of public language and text in human cognition is not limited to the preservation and communication of ideas. Instead, these external resources make available concepts, strategies and learning trajectories which are simply not available to individual, un-augmented brains. Much of the true power of language lies in its under appreciated capacity to reshape computational spaces which confront intelligent agents. (1995)

This conception of augmentation might give Mithen what he wants. It does not seem too far fetched to argue that even the most simple of protolanguages would act to externalise and extend existing cognitive resources. This might be sufficient mechanism to grant Mithen the integration, or "cognitive fluidity" he sees as demarcating human minds from all others.

## (5) Evolutionary Psychologists:

Mithen's support for cognitive fluidity is motivated not only by his condensation of the developmental literature but also by doubts about the Swiss Army Knife thesis notably expounded by Cosmides and Tooby (CT) in "The Adapted Mind" (Barkow, Cosmides and Tooby, 1992).

CT argue that the modern mind is a reflection of the evolutionary pressures brought to bear upon our Pleistocene ancestors. These pressures led to specific adaptive solutions and each adaptive solution is in essence a computationally specific module. A Swiss Army Knife metaphor is invoked to capture the central thrust - the modern mind is a tool for dealing with problems but this tool has specific blades for specific problems. It is not a general purpose machine. In other words, by Mithen's metaphor CT's modern mind is a phase two cathedral with isolated chapels about a minimal, and relatively unimportant nave.

Mithen's criticism of CT's model is first made in a subsection entitled "Hunter-Gatherers and Cambridge dons versus the evolutionary psychologists". He begins by noting that CT-modularism apparently fails to account for the extraordinary achievements of scientists such as Stephen Hawking and linguists such as John Lyons (all known to Mithen when he was a Junior Research Fellow at Cambridge). How can it be, for instance, that a mind with reasoning mechanisms adapted for the problems facing Pleistocence hunter-gatherers, is able to learn and execute mathematical operations?

What Mithen is pointing to is that CT have overlooked the issue of how personal level processes interact with subpersonal computations. For CT the whole of the modern mind consists of subpersonal modules solving a finite array of problems. Their classic example is that of logical reasoning. They have shown that some simple logical operations are more successfully solved when part of a social context. From this they argue for specific social modules that can only inadequately solve more abstracted tasks based on the same logic. Yet, CT fail to account for the success of those people who have been trained in logical analysis - how can they seemingly "reprogram" their subpersonal processes? (For related commentary see Dickins, 1997.)

CT's evidence for their modular model is based on strong evidence from reasoning task experiments and also from other areas within psychology (see "The Adapted Mind"). However, Mithen's points seem valid. His conception of cognitive fluidity, possibly mediated by language, seems to allow a compromise position to be taken. CT's modules are very likely in place and, when presented with problems nearer to those ancestral problems that helped to form these modules, they produce rapid and accurate responses. In other words, the archaic modular architecture may give cognitive biases. However, a result of cognitive fluidity is that information can be pooled and modular functions can be extended into more abstract areas. This compromise will only get us a certain distance before we need to introduce some higher order thinking processes too, in order to explain the ability to learn various logics etc., which at the very least will necessitate some computational changes to existing modules. Mithen has not provided these specifications despite his foray into language and CT appear averse to any higher order, more generalised thinking machinery. In short, this is a problem in need of resolution if evolutionary psychology wishes to progress.

Mithen continues with his criticism of CT by briefly discussing modern day hunter-gatherer culture. In CT's account hunter-gatherers in the Pleistocene will have had an encapsulated module for social reasoning and another for dealing with food. However, Mithen indicates a potential difficulty with this thesis:

> A girl choosing fruit using the same reasoning devices she uses for choosing a mate is likely to end up with severe stomach ache because she will choose unripe fruit - fruit which seems to have good muscle tone. Yet as soon as we look at modern hunter-gatherers this seems to be precisely what they do; not get stomach ache from eating unripe fruit, but reason about the natural world as if it were a social being. (p. 47)

Mithen expands this comment with examples from various extant cultures. For instance, the Inuit regard polar bears as food and hunt them. Yet, they endow the bears with human qualities that even result in performing the ceremonial activities associated with the death of a human where necessary. Indeed, the Inuit see polar bears as ancestors. This smearing of social thinking across domains is typical of all hunter-gatherer populations in the world today and leads Mithen to wonder:

> If - and it is a very big if - these modern hunter-gatherers are indeed a good analogy for those of the Pleistocene, how could selective pressures have existed to produce a Swiss Army Knife for the mind? (p. 49)

Whether or not it is a good analogy, Mithen has given us reason to think hard about how selective pressures will have operated in Pleistocene times. The exact adaptive problems facing Pleistocene humans may have differed from those of modern hunter-gatherers but the multiplicity and specificity of them will not have: many specific problems is not a necessary and sufficient condition for many specific cognitive modules.

## (6) Archeological Evidence:

Despite Mithen's worries about CT-modularism it is apparent from his phases of cathedral construction that he regards our more distant ancestors as having had minds structured like a Swiss Army Knife. Indeed, Mithen does not see any archaeological evidence for fluidity until the rise of the Modern Humans (Homo sapiens sapiens; this is Act 4) and critically sees the sort of modularism CT argue for only in Homo erectus, Homo heidelbergensis and Homo neanderthalensis - the Early Humans, as Mithen calls them (this is predominantly Act 3, although the Neanderthals and Homo heidelbergensis linger into Act 4). This dates purely modular thinkers earlier than CT's Pleistocene hypothesis.

At times one becomes uneasy about the archaeological evidence Mithen has used. For instance, one of the claims he makes is that we can see evidence of modularism in Early Humans from the spatial relationships between various material assemblages. It would appear that technical intelligence-using activities, such a tool-making, were carried out separately from other activities. Traditionally, the small clusters of tool fragments associated with Early Human sites were thought to be indicative of small social groups, which conflicts with suggestions of larger groups based upon brain size data etc.. Mithen suggests that modular cognition meant that tool activities were carried out separately from, say, social activities. Therefore, the fragments of tools found may not provide any information about overall group size as only a fraction of the groups might be making tools at any one time. Modern humans will have undertaken many, if not all, of their activities in the one place such that they socialised whilst making tools etc.. The larger clusters of tool fragments, and other items, found on Modern Human sites might indicate that this is the case, and these assemblages might also give a better key to group size as a result. Mithen claims that this is due to an integrated and fluid mind.

I am not sure that this particular argument supports the notion of modularism in Early Humans. Why would an Early Human physically move itself to a different site for different activities and why is Mithen so sure that no social interaction went on at the tool making sites? Indeed Mithen seems to weaken his own argument:

> ... the complex social behaviour and large social aggregations of Early Humans took place elsewhere in the landscape, perhaps no more than a few metres away - and are archaeologically invisible to us today. (p. 137)

If archaic social behaviours are archaeologically invisible how can he be certain that they happened elsewhere? But, even if Early Humans did move about, as Mithen suggests, why does this indicate modularity? Modern Humans have specific sites for tool making, for butchery of animals, for sexual behaviour etc. but no one would argue that the geography of these events indicates a cognitive domain governing each one. None the less, the massive changes in technology that we see in Act 4 do suggest a radical cognitive change which might be due to modular integration if we can find support for Early Human modularity.

## (7) Conclusion:

There are a number of reasons for recommending this book. First, Mithen has provided cognitive scientists with a new method of analysing archaeological finds which shall prove very fruitful, despite the above reservations. Second, Mithen has presented scholars, within human evolutionary disciplines, with a tantalising, and therefore provocative, structure for the shifts in cognitive evolution based on a number of different disciplines. Third, if any cognitive scientist still have doubts about the importance of evolutionary modelling for theories about the mind, this book is one of the best remedies available thanks to its lucid style and bold attempt to tell the whole story. I am inclined to treat the detail of this structure with caution but it deserves serious attention and will generate much future research. Finally, there is much in this book I have not written about. Notably the sections on culture, art and farming, and how a fluid mind allows these things to occur, are fascinating explorations and take one far beyond the drier aspects of cognitive theory. In short, Mithen has shown us that the mind is both mysterious and explicable.



## References:

Barkow, J., Cosmides, L. & Tooby, J. (Eds) (1992) *The Adapted Mind*. Oxford University Press.

Clark, A. (1995) 'Language: The Ultimate Artifact.' Precirculated MS of a paper given to the Hang Seng Centre for Cognitive Studies, University of Sheffield.

Dennett, D.C. (1991) *Consciousness Explained.* Penguin Press.

Dennett, D.C. (1995) *Darwin's Dangerous Idea.* Penguin Press.

Dennett, D.C. (1996) *Kinds of Minds.* Weidenfeld & Nicolson

Dickins, T.E. (1997) 'The State of the Evolutionary Art in Psychology.' *Connexions*, (2),

Dickins, T.E. & Frankish, K. (1997) Review of 'Dennett, D.C. (1996) "Kinds of Minds"'. *History and Philosophy of Psychology Newsletter, British Psychological Society*, Spring (24), 36 - 40.
Dunbar, R. I. M. (1993) 'Coevolution of neocortical size, group size and language in humans.' *Behavioural and Brain Sciences*, 16, 681 - 735.

## *A report on the Conference on Metarepresentation presented by the Cognitive Science Program of Simon Fraser University at Harbour Centre*

## David Byrd

The cognitive turn in philosophy and psychology has led some to wonder "if cognitive science is the answer, what was the question?" It has been claimed that theories of mental representation are misdirected if the goal is to reach an explanation of higher level phenomena such as consciousness or personal identity. Even if we could formulate an account of how it is that a cognitive system represents its surroundings, for example, such that it successfully identifies and navigates about objects, how could this help to explain the more sophisticated elements of human cognition?

One attempt to address this issue is provided by recent work on metarepresentation. It is this avenue that the Vancouver Cognitive Science Conference, organised by the Cognitive Science Program At Simon Fraser University, chose to investigate this year.

The role metarepresentation plays in accounting for human cognition has been neatly summarised by Norman Swartz:

> Cognitive systems are characterised by their ability to construct and process representations of objects and states of affairs. Representations are themselves objects in the world and therefore potential objects of (second-order or meta-) representations. However, humans seem to be nearly unique in their ability to represent representations. This meta-representational ability plays a central role in reflective thinking and in human interaction. (Swartz, personal communication, 14 October 1996.)

The conference brought together philosophers of mind, philosophers of science, psychologists, linguists, and anthropologists. The talks given were arranged into seven sessions, each consisting of two talks.

The first session was titled *"Meta-Representation: The Broad View,"* and **Alvin Goldman** led the discussion with his talk, *"The Mentalising Folk."* Goldman was, for the most part, concerned with the role that metarepresentations play in simulation theories (that Goldman takes to be the best account for third person understanding). According to this account our brains/minds contain representations that represent other people's representations. If we want to understand what someone else may infer (how someone else "thinks") we "simulate" their thinking in our brain. An example might be that if I know you like to feed cats that come to your back door (i.e. I have a representation that represents your representation of liking to feed cats) and I know that you know there currently is a cat at your back door (this is a further representation of one of your representations), I can infer that you want to feed the cat currently at your back door by simulating your own thinking. As long as no other interfering factors come into play this simulation of your thoughts will be correct. So, according to Goldman, since the ability to represent the representations of others is essential to understanding others, metarepresentation plays a pivotal role in understanding human cognition.

The next talk in the "Broad View" session was **Daniel Dennett**'s *"Making Tools for Thinking."* Apart from raising general questions, such as "can chimpanzees metarepresent?", Dennett tried to give us an intuitive feel for why we evolved the ability to metarepresent. He did this by using an example where three lions are stalking a small group of gazelle. Before the ambush one of the lions takes a group of sticks and rocks and, like a plan scrawled in the dirt by players in a huddle, communicates to the other lions its intent. Now obviously we do not observe animals doing this, but, if they did it would be an immensely valuable tool for survival. Dennett suggests this is precisely what we are able to do in our brains with representations. We developed the ability to use and manipulate representations and metarepresentations because those species that could do it would have an adaptive advantage over those that could not.

Following Dennett's talk came the session on evolutionary perspectives. In the talk, *"Meta-representations in Evolutionary Perspective,"* **Leda Cosmides** and **John Tooby** gave an account of why the type of metarepresenting that Goldman discussed evolved in us. They suggested that we developed ToMM's ("theory of mind" modules) capable of metarepresentation so that we could "read minds." Having evolved as social creatures, these capabilities were essential in order to predict the actions of others. Those individuals that could "link observable cues (such as direction gaze) to representations of unobservable mental states (such as wants and beliefs)," would accrue an adaptive gain. Evidence for this, claim Cosmides and Tooby, is seen in cases of schizophrenia where the ability to metarepresent seems to break down.

**Andrew Whiten** gave a talk, *"Chimpanzee Cognition and the Question of Mental Re-presentation,"* that responded to Dennett's earlier chimp speculation with a positive "Yes, chimpanzees do metarepresent". Whiten argues for this on the strength of observed chimpanzee abilities of imitation, "mindreading," and pretence. The most interesting of these abilities, "mindreading," was displayed to the audience in a video of a chimpanzee that could differentiate between someone knowing one thing and someone knowing another (i.e., it could tell whether someone had a certain representation instead of other representations). The video showed a situation where a chimpanzee had the opportunity to point to a key (visible but inaccessible to the chimpanzee) that would open a locked box containing the chimpanzee's food that the keeper could then get for the chimpanzee. Repeated experiments showed that the chimpanzee would reliably point to the key when the chimpanzee knew its keeper did not know where the key was, and not point the key when the it knew that its keeper did know where the key was. Whiten takes this to be proof of the chimpanzee's ability to successfully represent the representations of others.

The next session was concerned with developmental perspectives. The first talk was **Alan Leslie**'s *"Specialisation in a Higher Cognitive Function: Neuropsychological Studies of Meta-representation"*. Leslie believes children attend to mental-representations not by acquiring, or constructing, a theory that does so, but by being born with a ToMM. His argument arises from evidence of autistic children that are in the same environment as normal children, yet the autistic children simply lack the machinery to metarepresent.

**Susan Carey**'s talk, *"Distinguishing Conceptual Enrichment from Conceptual Change,"* argued that ToMM's, as systems of metarepresentations, change during a child's development. She first made the distinction between mere accretion of knowledge and genuine conceptual change by analogy to a similar distinction made by historians of science (Kuhn). Carey used Williams Syndrome case studies (where children have the ability to increase their knowledge, yet cannot undergo conceptual change) as evidence for the distinct abilities of normal children to use systems of metarepresentations that can undergo conceptual change.

The concluding session of the first day was on the linguistic perspectives of **Raymond W. Gibbs** and **Deirdre Wilson**. Gibbs' talk, *"Metarepresentations In Staged Communicative Acts,"* simply illustrated the pervasive role of metarepresentations in our communication with each other, especially in our use of metaphor and irony. He showed a video clip from a documentary of an American family in the 1970's. A mother disapprovingly says to her son, "You don't do anything around here. You're a nothing!" And the son responds, "Yeah, you're right. I'm a nothing!" Here the son is expressing his belief that his mother's belief of him is false. If representations play the foundational role of constituting our thoughts that cognitive scientists think they do, even our simple everyday conversations contain metarepresentations of great complexity.

In addition to giving even more examples of metarepresentations in natural language use, **Deirdre Wilson**'s talk, *"Linguistic Meta-representation,"* showed how we use metarepresentations in linguistic comprehension (or utterance interpretation). One strategy for utterance interpretation might be Naive Optimism: the listener looks for an interpretation that satisfies the listener's expectation of conversational relevance. Or using Cautious Optimism, a listener might look for an interpretation that the speaker might have thought would satisfy the listener's expectation of relevance. Or even Sophisticated Understanding: the listener looks for an interpretation that the speaker might have thought the listener would think would satisfy the listener's expectation of relevance. The different strategies obviously make heavy use of metarepresentations.

The next day began with a session on higher order cognitive processes. In his talk, *"The Metaloop of Exemplaric Representation,"* **Keith Lehrer** suggested that metarepresentations play a large role in evaluating the meaning of representations or, analogously, statements. For example, when we are given two statements that are apparently inconsistent, such as "this is green" and "this is not green", we may use the larger statements: "Jim believes 'this is green.'" and "I believe 'this is not green.'" to help us resolve the apparent inconsistency. Likewise, Lehrer suggests we do this with representations. First level representations are "quoted" and then used in higher level metarepresentations. Lehrer calls this process metamental ascent, but he believes it suffers a possible regress such that in order to understand a word or concept, it has been alleged that we must quote or represent it. But to understand that, we have to quote or represent that, and so on. To avoid this regress, he suggests we "obtain a minimal representation of meaning by using a word [or concept] as our exemplar of a class of words [or concepts] including itself." He calls this process "exemplarisation" and says that it allows one to ascend to quotation and descend to disquotation, thus ending the regress in a "harmless referential loop." An example would be if we wanted to represent (and have access to the content of) the song "Twinkle, Twinkle, Little Star." We would represent it by quotation: "I am thinking of 'Twinkle, Twinkle, Little Star.'" But in its quoted form the representation's meaning (content) is opaque. To understand its content we simply disquote it: "Oh, you want to know what it means? Listen: 'Twinkle, twinkle, little star, how I wonder...'"

**David Rosenthal** gave a talk entitled, *"Conscious Metacognition,"* that claimed we can use metarepresentations to explain how we are conscious of some mental states or representations and not others. Taking memory as an example, out of all the memories (mental states) that we possess, and of which we could be conscious, why is it that we are conscious of some (perhaps only one) rather than all of our memories at once? Rosenthal suggests a mental state is conscious if a meta-mental "pointer" happens to be indicating that state. This pointer is a higher order thought (or HOT) that selects, out of all possible states, the one of which the cognitive system will be conscious. Rosenthal thinks this mechanism allows us to explain why consciousness seems so spontaneous; we are conscious only of the states to which the mental "pointer" points, and not of the pointer itself.

The next session on metarepresentations and social communication, was somewhat disappointing to me because I failed to understand the argument for the first talk, and I did not see how the second talk addressed the notion of metarepresentation at all. In outline, the first talk by **Francois Recanati**, *"The Iconicity of Meta-representations,"* argued that even though there is something metarepresentational about deferential representation, namely that a deferential representation contains a metarepresentational component in its "character," Recanati believes the content of a deferential representation is not actually metarepresentational at all.

The second talk, **Steven Davis**'s *"Deference and Conversation,"* investigated the nature of linguistic deference and the sorts of conversational exchanges that experts and non-experts must be prepared to engage in to make it the case that non-experts can refer to, for example, the differences between elms and beeches, even though the non-experts may not actually know the differences.

Fortunately the next and final session was much clearer. **Robert Wilson**'s talk, *"The Mind Beyond Itself,"* discussed the external nature of metarepresentations following from an externalist account of representations. He explored this in the areas of memory, external symbols, culture and evolution, and theories of mind in general. An example of this approach would be that of defining a cognitive system as a monkey plus a symbol manipulating apparatus external to the monkey. Any metarepresentations we attribute to the monkey that make use of the monkey's representation of objects, using the symbols of the apparatus, will have a "wide" content, i.e., a content not solely individuated in the monkey, but individuated by the monkey + apparatus system.

**Daniel Sperber**'s talk, *"Culture and the Epidemiology of Metarepresentations"* suggested that metarepresentations play a major role in human cognition in three areas: "mental representations attributed to self and others, public (linguistic) representations, and abstract representations (e.g., mathematical theorems)." Sperber argues that, in terms of evolution, although it might be thought that linguistic metarepresentation came first, linguistic representation presupposes the ability to attribute representations to others, and thus psychological representation must have evolved prior to linguistic representation. Hence psychological metarepresentation must also be seen as having evolved prior to linguistic metarepresentation.

All in all, I felt the conference was very successful. It was large enough to draw significant talent from diverse fields, but small enough to allow those attending to personally engage the speakers. My only qualm was that there was not enough time allocated for questions after each talk (on average, five minutes).

**Note**
Unless otherwise cited, all quoted material came from the talks or their respective handouts.

## *A Personal View of The Ciba Open Meeting*

**Margaret Clegg**
*Department of Anthropology,*
*University College London.*

Psychological adaptations have evolved for the same reason as physical adaptations, because they enhance the fitness of the individual. Research in cognitive evolution needs a more integrated approach if the under lying structure and function of such adaptations is to be understood. The Characterising Human Psychological Adaptation Open Meeting at the Wellcome Centre for Medical Science addressed some of these concerns. The meeting presented highlights from the Ciba Foundation symposium on human psychological adaptation to a wider audience. The approach was broad bringing together psychology, anthropology, evolutionary biology and cognitive science. Many of the speakers' work embraced concepts from all or many of these disciplines. Three recurring themes wound their way through the talks. The first was the complementary nature of the research across these fields and the importance of learning from other disciplines. The second was the importance of human social life in shaping cognitive development and finally the need to look further than our own species, to other animals, to help reveal the hidden nature of our own behaviour and cognition.

Paul Gaulin stressed the complementary nature of the questions asked, the how of psychology and the why of evolutionary biology. He advocates using cross cultural comparisons as this gives a broad perspective of the characteristics of human behaviour that uncover what human behaviour was designed to do. He looks at two types of behaviour: Firstly human universals, behaviour that is expressed regardless of cultural differences, which are probably fixed in the species by evolution. Secondly, conditional universals, behaviour that varies predictably with some environmental parameter. This approach has much in common with that taken when looking at physical adaptations. It acknowledges the plasticity of behaviour while retaining the species specific nature of other adaptations. To understand fully what is happening researchers need to formulate questions, as Gaulin maintains, from many different viewpoints.

Leda Cosmides also argues for the need to use methods developed in other disciplines if they fulfil the criteria for examining human cognitive adaptation. She believes that the adaptionist programme of evolutionary biology gives the theoretical standards by which to dissect human cognitive architecture, specifically social cognition, which can be characterised as a suite of integrated mechanisms. She likened this to a Swiss army knife, the mind as a multifunction tool. She sees the brain as composed of a large number of expert systems. These systems are specialised for reasoning and making decisions about different aspects of the social world. The reasoning humans are specifically good at is that involved in a social context. Problems that are difficult to solve in an abstract way are easily solved when a social dimension is added. She maintains that there are dedicated cognitive systems for social situations. However, although there is evidence for such systems there is also evidence for a larger degree of environmental input. A variety of animal experiments (for example Wiesel & Hubel 1990, Killackey et al 1995, Rakic & Riley 1983) have shown that different areas of the brain can be made to assume functions other than that expected, depending on the sensory input to them . It seems likely that cognitive systems have a similar plasticity and present use may not divulge original function.

Other speakers also felt that social cognition was at the root of many psychological adaptations. Margo Wilson suggested that many psychological adaptations have evolved for social living. Social life largely comprises interactions with relatives and friends. But most investigations in psychology have focused on interaction with strangers. Evidence of various types shows examples of relationship specific social psychologies. One area that illustrates this is interpersonal conflicts. She has looked at victim-killer patterns of homicides for context, nature and demography. Male sexual jealousy as the number one motive for murder appears to be cross cultural. Children too are most at risk from unrelated males with whom they live. In evolutionary terms this makes sense, the male is safe guarding his genetic future by eliminating those who would supplant his own offspring. This does not of course condone such behaviour but does show our mammalian heritage, where such behaviour occurs when a new male takes control.

John Tooby proposed that human social life has formed pathways to the evolution of altruism that are alternate to kin selection and reciprocal altruism in explaining human love and friendship. Some distinguishing characteristics of altruism, specifically expecting a pay back, are seen in all cultures as a lack of friendship. He therefore maintains that humans have evolved risky behaviour which damages short term fitness but demonstrates long term commitment. Although his ideas are tempting, many of the behaviours he describes may still be explained by game theory (Lahno 1995) and extensions of kin selection (Rushton, Russell and Wells 1984). However, his proposals appear to be at an early stage of formulation and it will be interesting to see how Tooby develops his ideas.

The need for a comparative approach in understanding behaviour and cognition was proposed by both Marc Hauser and David Sherry. The former stressed the need for methods that could look across species, making comparisons direct and relevant. He made the point that not only do other species have different minds so too do human infants. He has used a technique developed to assess the cognition of human pre-linguistic infants on non-human primates. The technique makes use of the ability to count that is evident in primates in the wild. The subject looks preferentially longer at objects that are not the outcome expected, for example two pieces of fruit appearing when only one was seen going in. This expectancy violation is seen in both non-human primates and human infants older than one year. Therefore both have the ability to store representations. This type of comparison may allow us to distinguish between cognitive domains shared with other primate and those specific to humans presumably shaped by language. Hauser's work is refreshing in that it looks at primates in the wild. He is therefore examining behaviour and abilities inherent in these animals, not that induced by captivity or training.

David Sherry also stressed the need to use comparative studies. He illustrated this through research into spatial memory in both birds and mammals. Food sharing birds have an increased hippocampus and related brain areas. Rodents show sex differences in the same brain areas. Polygenous male rats, who need to find females in the environment, having the larger hippocampus. Humans also show sex differences in spatial ability. A comparative approach suggests that sex differences in human spatial ability are sex specific adaptations and may provide tools for determining the adaptive significance of sex differences in human cognition. There is one problem with Sherry's sex differences in cognition. Human culture has often restricted female experiences, therefore any differences may not show the adaptive differences in human cognition only differences imposed in the very recent past.

Steven Pinker focused as one would expect exclusively on language evolution. His talk was for the most part a reprise of his now famous book "The Language Instinct". His approach too embraced the cross cultural and complementary approach to cognitive research and the social aspects of cognition. However, his position in accepting no homologues for language in other primates restricts any comparative approach to language evolution, although work in several areas including; within group close calls (Harcourt, Stewart & Hauser 1984, among others),the representational abilities of primates (for example Hauser's work presented at the meeting), and the increase in both, cortex size (Dunbar and Aiello 1993) and the development of discrete cortical areas for communication (Gannon pers. comm.) as primate group size increases, may have comparative applications for language evolution.

The meeting was stimulating and thought provoking. Although many speakers asked more questions than they answered the questions need to be asked and the methods proposed may provide the means by which to understand human psychological adaptations. Such meetings bring together researchers from different backgrounds and interests. This enables them to talk to one another, encouraging researchers to look across subject boundaries, explore new literature and find and use techniques novel in their field. This cross fertilisation moves forward the boundaries of scientific research allowing us to build the 'new and better mouse trap'.

## References

Aiello LC & Dunbar RIM (1993): 'Neocortex size, Group Size and The Evolution of Language' *Current Anthropology* 184-193

Harcourt AH, Stewart KJ & Hauser M (1993): 'Function of Wild Gorilla Close Calls I Repertoire, Context and Interspecific Comparison' *Behaviour* 124 (1-2) 89-122

Killackey HP Rhoades RW & Bennett-Clarke CA (1995): 'The Formation of a Cortical Somatotpic Map Special Edition Evolution and Development of the Cerebral Cortex' *Trends in Neuroscience* Vol. 18 (9) 402-408

Lahno B (1995): 'Trust, Reputation and Exit in exchange Relationships' *Journal of Conflict Resolution* vol. 39 (3) 495-510
Rakic P & Riley KP (1983): 'Regulation of Axon Number in Primate Optic Nerve by Prenatal Binocular Competition' *Nature* Vol. 305 (5930) 135-137
Rushton JP, Russell RJ & Wells PA (1984) 'Genetic Similarity Theory: Beyond Kin Selection' *Behaviour-Genetics* Vol. 14 (3) 179-193
Wiesel TN & Hubel DH (1990): 'Brain Mechanisms of Vision' in *The Perceptual World: Readings From Scientific American Magazine* ed. Rock I WH Freeman & Co New York

---

# *The State of the Evolutionary Art in Psychology:*
## *A review of the "Evolutionary approaches to psychology" symposia at the 1997 Annual Conference of the British Psychological Society, 3 - 6 April.*

**Thomas E. Dickins**
*Department of Human Communication Sciences*
*University of Sheffield*

This year's annual British Psychological Society conference was held in the pleasant surroundings of the Heriot-Watt University in Edinburgh. This is a campus university situated a short bus ride away from the centre of Edinburgh and has marvellous views of the countryside. Unfortunately, the conference schedule was busy and there was not much opportunity to explore.

The conference was divided into three keynote themes: evolutionary approaches to psychology, organisational psychology and feminist psychology. I attended the conference with the intention of attending the first of these three sessions which lasted for two days.

Day one was devoted to the evolution of social reasoning and the audience was addressed by an impressive range of speakers including **David Perret**, **Rob Barton**, **Andrew Whiten**, **Simon Baron-Cohen** and **Leda Cosmides**. The predominant thrust of this section was that our cognitive capabilities and capacities were fashioned by natural selection within a highly social environment. What is more, our socially evolved brain has been structured with a modular architecture that allows specific "problems" to be dealt with in specific ways. This view of cognitive evolution is the current zeitgeist within the field.

Leda Cosmides gave an invited lecture at the end of the day, on *"Evolutionary Psychology",* that drove the social modularity point home through an analysis of human reasoning ability using the Wason Selection Task (WST) as a bench mark test. This is a tool that Cosmides has written much about, most notably in *"The Adapted Mind" (Barkow, Cosmides & Tooby, 1992, Oxford University Press)* but it was interesting to have it explained in public. The WST consists of variants about a core problem that focuses upon four marked cards, only one side of which can be seen. In the original version two of the cards had letters marking them, one vowel and one consonant, and two had numbers, one even and one odd. The task was to say which of the cards it was essential to turn over to check the rule: "if a card has a vowel on one side then it has an even number on the other". The logical way to do this is to attempt to falsify the rule by turning over the vowel card and the odd number card. If the vowel-card has an odd number on the back then the rule is falsified, if the odd-card has a vowel on the back the rule is again falsified. Cosmides has noticed that performance on the WST, which is typically poor in its original "abstract" form, can be enhanced when the task is rephrased so that people are looking for cheaters. Cosmides has manipulated the narrative around the task without altering the underlying logic such that the rule becomes part of a social scene and violations have direct ramifications within the group, be they trivial or important. Where people once were unable to spontaneously look for falsifying instances they now seem able to make the right choices based on a search for cheaters.

Cosmides argues that this work shows two things: First, that human reasoning, which was once thought to be a generalised ability, is actually specialised and not particularly generalisable. Second, that our reasoning mechanisms are predominantly social and that, judging from these WST findings, we have a specific module to deal with cheater detection. This latter hypothesis makes good intuitive evolutionary sense. Undoubtedly the social environment within which our hominid ancestors found themselves was a fiercely competitive one in terms of resources and much of the tactical deception of the sort seen in contemporary non-human primate groups went on. However, rampant cheating would cause groups to disintegrate and this would be disastrous, thus social censure of some sort is highly likely to have been selected for and an ability to reason well about such social violations would be an equally useful companion adaptation. Of course, the evolutionary stories of the genesis of these mechanisms needs to be told but for now we can see that Cosmides' putative adaptive story is persuasive.

Despite the good evolutionary principles at work in theories such as Cosmides' I am a little uncomfortable. The reason for this is two-fold. First, I am not sure I want to buy this modular story wholesale. It seems possible that a good case can be made for at least some generalised reasoning, or thinking ability by reflecting briefly on the personal and subpersonal computation division. It is quite clear that personal level interventions can make a difference to processes where subpersonal computations would normally carry on. As David Papineau pointed out in the recent meeting of the *Hang Seng Centre for Cognitive Studies (22-23 February 1997, University of Sheffield)* it is still the case that we can do the abstract version of the WST - all we have to do is learn a little logic and then we can reason about the task in whichever format it is presented. What is more, we can extend that logic beyond WST tasks to many others. It may well be that we have an evolved predisposition to initially socially reason, it may even be that this predisposition is the seed of our subsequent "rationality", but it is hard to deny that we can beat such biases at least some of the time.

When I raised this with Cosmides she made a comment that leads me to my second concern: how modular are these modules? Her way of resolving the above issue is to argue that her modules are not totally encapsulated, as in Fodor's *("The Modularity of Mind", 1983, MIT Press)* model, and that they are not overly content specific. For example, the ability to count is not dependent upon what is being counted. But this seems to be running the danger of giving away what has been gained. There must be some constraints on the type of information processed by cheater detection modules for the WST argument to stand. If these constraints are loosened then the model begins to look like a generalised reasoning machine argument again. How is a line to be drawn? Personally, I would favour an argument that incorporated a generalised and higher order cognitive architecture that is underpinned by Cosmides-type modules. Of course, this higher order system could itself be modular, but the crucial thing is where it gets its information from, and this has to be from the other modules. With this suggestion I am admitting the possibility that the data from the WST work may not indicate anything more than a cognitive bias rather than a hard and fast architectural constraint.

The second day dealt with *"Evolution, rationality and social games"* and again incorporated an august array of speakers including **Henry Plotkin, David Shanks**, **Geoff Miller** (who organised the day) and a contingent from the Max Planck Institute, Germany. This meeting was heralded by Geoff Miller as "the new stuff" to be viewed against the background of the previous day's introduction. However, I think it would only be regarded as new by those not already "dabbling" in the area or by the more narrow and well defined psychologists. One of the plus points of evolutionary psychology is that it is not only psychology but also biology, ethology, cognitive science and many other things, and this section gave the audience a taste of this variety. The papers ranged from game theory speculations, through mating strategies, to technical models of information processing and they were all fascinating.

To highlight the diversity, and to set the project, Henry Plotkin gave an invited lecture entitled *"Evolutionary psychology: Marrying the biological and social sciences"*. In this paper Plotkin argued that the challenge for the new century is to make this marriage work by encouraging a two-way relationship. Evolutionary theory must inform psychological theory but psychological theory can also inform evolutionary theory. Plotkin pointed to the need to have good theories of intelligence and culture that allow us to subsequently model human evolution. Such a position, though highly polemic, provided a good shake up for those of us (including myself) who have felt the one thing evolutionary theory is going to do is save psychology from itself. For theorists such as myself evolutionary theories enable us to better understand the function of cognitive mechanisms, as Cosmides' work on the WST has attempted to do, by placing them within a relevant selective and adaptive landscape. By realising what, and how, mechanisms were selected for we can better model their architecture. Evolutionary thought enforces a naturalistic parsimony on psychological modelling. Plotkin was not denying this but simply arguing that there is a more complex interaction between evolution and its products. Culture constrains possible routes within the design space open to natural selection in much the same way as early limb-like structures constrained the development of the bat's wing.

I still regard evolutionary theory as being best placed to save psychology from unparsimonious modelling whilst realising that Plotkin's central point about cultural constraints is true in its mild form.

What worries me is the claim that Plotkin made that the age-old nature-nurture issue is not a non-problem. The reason he said this was to force the point that culture, and general human activity has had an effect upon human evolution. This is true, but I do not think muddying the water with the nature-nurture issue will help generate good theory. It is quite clearly a non-problem as it misunderstands the basic fact that any "nurturing" behaviour that is undertaken is the act of a biological organism driven by mechanisms of varying complexity in order to maximise its inclusive fitness. In short, no distinction between "natural" and "unnatural" makes sense and in this light culture is to be seen as a biological adaptation. If we allow such dichotomies to stand we endanger our theories by opening up the possibility of cultural determinism, which is largely a backlash against a misunderstood biological determinism. What I term cultural determinism is a form of essentialism that argues for the prevalence of a human "essence" over biological restriction. Differences between human groups are pointed to as evidence for this, but I would maintain that these differences are best thought of as variance, or noise, and what is interesting are the huge similarities to be found between the people of the world. I suspect Plotkin would agree with my polemic, but I worry that encouraging such dualisms will lead to a jilted partner.

The take-home message from Plotkin should simply be *"the study of humankind is entirely biological"*.

In conclusion, the most important thing to say about the conference was that it was highly stimulating and generated much thought and discussion. Judging from the excited and talkative behaviour of the delegates attending the other sections this seems to have been a general rule rather than a domain specific one.

## *The Second Hang Seng Workshop on Evolution of Mind, The University of Sheffield, 22-23 February, 1997*

**Jennie Hawcroft**
*Research School of Archaeology,*
*University of Sheffield.*

Keen as ever to engage in cross-disciplinary discussion, we in the Archaeology Department approached this latest Hang Seng workshop with excitement. This was fuelled largely by our enthusiasm for the work of Thomas Wynn, which formed the central discussion topic of this meeting.

As the meeting got underway we were rather disappointed to find ourselves the only representatives of archaeology for most of the first day. Clive Gamble was delayed by those bungling amateurs South West Trains but, fortunately he arrived in time to give his paper. This was a coherent and comprehensive overview of the current picture of human evolution, presumably requested to help the non-archaeologists present put Wynn's paper on handaxes into context.

Wynn's paper itself was as insightful as we had hoped, but it was after this that we felt the lack of some more cognition-oriented archaeologists, such as Foley and Dunbar. To have such cross-disciplinary experts engaging in debate with Thomas Wynn would have been very exciting. That said, the discussion which followed was of a high standard in terms of depth and ideas, and was in no way unsatisfactory - but it could have been improved by the presence of someone whose research formed a bridge between the archaeologists, philosophers and psychologists.

Of course, the conference was aiming for a discussion on cognition, not archaeology, so it was not inappropriate that we were outnumbered. The papers were diverse and ranged widely around the given theme. In some of the papers I struggled to keep up, while in others I was already familiar with the research - this is probably unavoidable in a multi-disciplinary meeting.

On a selfish note, I found the diverse nature of the discussants present very valuable on two levels. First, much of the work presented in the papers was new to me and gave me ideas to pursue that I would not otherwise have had. Second, I was led to reconsider some of my previous assumptions when I realised how attitudes and priorities differ in other disciplines. The best example of this was Patty Cowell's paper on sex differences in abilities. That these differences exist as a biological function seemed to be already accepted by most discussants, but this would be extremely controversial in archaeology. Luckily, ample "socialising" breaks were provided and I was able to take this further with Patty.

Patty's paper described how males and females differ in their accomplishment of certain abilities such as language and spatial awareness. I had heard of this before, but had assumed the difference was caused by conditioning factors in life, rather than having a biological basis (which suggests a selective advantage for differing male and female roles in the past). Patty's talk gave me the impression that neuroanatomists and psychologists accept that the difference is caused by biological factors, something that I found hard to swallow. She went on to describe her evidence which suggests the difference is caused by hormonal factors, which are dependent in turn on environmental influences, thus placating me a bit. During break, I told her how surprised I had been at the seeming lack of controversy stirred by the idea of differential abilities for men and women. In archaeology, where feminism seems to have won the day in recent years, this idea would be outrageous, and would only be tolerated if it were explained as being a result of the way young boys and girls are treated and educated. Patty agreed that this could be a causal factor of the sex difference, but in turn, expressed surprise that a biological basis for what seems to be a well-substantiated effect could be controversial in other disciplines. We concluded that our opposing assumptions were the result of our material-centred and biology-centred backgrounds respectively, agreed to differ and commented again on how much you can get out of talking to people with different perspectives.

I left the conference with the a lot of new ideas and a fresh perspective on some old ones. While I would have appreciated having someone like Dunbar or Foley to act as "translator" between the diggers and the thinkers, in many ways it did me a lot of good to have my assumptions challenged by non-archaeologists, and I would be keen to attend such a gathering again.

Helped by comments from Dr Andrew Chamberlain

## *Themes From Davidson Conference at Bolton Institute* (10th May 1997)

**K.A.Wilkinson**
*Department of Philosophy*
*University of Sheffield*

This proved to be an enjoyable and stimulating conference with much to interest both Davidson scholars and those less familiar with his work. The conference started at 10am with everyone being welcomed with copious amounts of coffee and biscuits which were available throughout the day.

The first paper was presented by E.Harcourt (Lady Margaret Hall, Oxford) entitled "Two Problems for Interpretationism". This was a relatively technical paper concerning the incompatibility of three views attributed to Davidson namely the (re)expression thesis, the resticted availability thesis and the interpretationism with regard to first person utterances. The second problem raised was finding a suitable account of the "I" pronoun in self-ascription once the Fregean account is given up.

After a short break with yet more coffee, J.Hornsby (Birkbeck College, London) delivered her paper - "Anomalousness in Action". She distinguished two Davidsonian claims that form the basis to his account of Anomalous Monism - rationalization as a species of causal explanation and the stronger claim that reasons are causes. She claimed only the former claim is warranted. The claim that reasons are causes was dismissed due to its inability to successfully deal with reasons which explain deliberate omissions. Consequently, J. Hornsby argued that the weaker claim, taking rationalization to be a type of causal explanation generates a correspondingly weaker account of Anomalous Monism than Davidson outlines.

The final paper of the day, after an excellent lunch, was "In Defence of Convention" by T.Baldwin (University of York). This was a particularly interesting and accessible paper. This began with general discussion of the nature of convention and the respective attitudes of Lewis, Davidson and Chomsky. It went on to address the three main problems that Davidson raises against Lewis' conventionalism, namely the problem of insincerity or deception, a lack of seriousness (giving rise to Davidson's Autonomy of Meaning) and the human capacity to understand language abuse, including malapropisms. T.Baldwin identified all three phenomena as parasitic on the general community's use of language. Thus, he claimed they could occur against a backdrop of conventions as opposed to resorting to Davidson's Radical Interpretationalism. General conventions explain our capacity to understand one another, despite problematic cases identified by Davidson, and our confidence in being able to converse with strangers. The subsequent discussion initiated by S.Perry (Sheffield) focussed on the conceptual difference between Davidson's radical interpretationism with rules of thumb and a fully-blown conventionalist account.